NUCLEAR POWER – CONTROL, RELIABILITY AND HUMAN FACTORS

Edited by **Pavel V. Tsvetkov**

INTECHWEB.ORG

Nuclear Power – Control, Reliability and Human Factors
Edited by Pavel V. Tsvetkov

Published by InTech
Janeza Trdine 9, 51000 Rijeka, Croatia

Publishing Process Manager Petra Zobic
Technical Editor Teodora Smiljanic
Cover Designer Jan Hyrat
Image Copyright Ensuper, 2010. Used under license from Shutterstock.com

First published September, 2011
Printed in Croatia

A free online edition of this book is available at www.intechopen.com
Additional hard copies can be obtained from orders@intechweb.org

Nuclear Power – Control, Reliability and Human Factors, Edited by Pavel V. Tsvetkov
p. cm.
ISBN 978-953-307-599-0

Contents

Preface

Due to economic growth and increasing population, energy demands must be satisfied in a sustainable manner assuring inherent safety, efficiency and minimized environmental impact. Nuclear power has long posed a dilemma for environmentalists and scientists alike. On the one hand it is seen as a cheap, clean energy source whilst on the other some have concerns over its ability to dispose of radioactive waste. Whichever viewpoint one may assume, nuclear power is at the forefront of clean energy technology and can be made available on a large scale to meet energy needs of the rapidly growing world.

Today's nuclear reactors are safe and highly efficient energy systems that give electricity and a multitude of co-generation energy products ranging from potable water to heat for industrial applications. Meanwhile, a catastrophic earthquake and a tsunami in Japan led to the nuclear accident that forced us to rethink our approach to nuclear safety and design requirements. It also encouraged the growing of interest for advanced nuclear energy systems and next generation nuclear reactors, inherently capable of withstanding natural disasters, avoiding catastrophic consequences and leaving no environmental impact. Advances in reactor designs, materials and human-machine interfaces assure safety and reliability of emerging reactor technologies, eliminating possibilities for high-consequence human error, such as those which have occurred in the past. New instrumentation and control technologies based in digital systems, novel sensors and measurement approaches facilitate safety, reliability and economic competitiveness of nuclear power options.

Autonomous operation scenarios are becoming increasingly popular to consider for small modular systems designed for remote regions with limited industrial infrastructure or regions with no such infrastructure but with human population whose safety, prosperity and growth depend on a reliable energy supply.

This book is one in a series of books on nuclear power published by InTech. It consists of four major sections and contains twenty-one chapters on topics from key subject areas pertinent to instrumentation and control, operation reliability, system aging and human-machine interfaces.The book opens with the section on instrumentation and control aspects of nuclear power. The following sections and included chapters address selected issues in reliability and failure mechanisms, component aging, plant

operation and human factors. The book shows both advantages and challenges emphasizing the need for further development and innovation.

With all diversity of topics in 21 chapters, the issues of nuclear power control, reliability and human factor represent a common thread that is easily identifiable in all chapters of the book. The "systems thinking" approach allows synthesizing the entire body of provided information into a consistent integrated picture of the real-life complex engineering system – nuclear power system – where everything works together.

The goal of this book and the entire book series on nuclear power is to present nuclear power to our readers as a promising energy source that has a unique potential to meet energy demands with minimized environmental impact, near-zero carbon footprint, and competitive economics via robust potential applications.

The book targets a broad potential readership group - students, researchers and specialists in the field - who are interested in learning about nuclear power. The idea is to facilitate intellectual cross-fertilization between field experts and non-field experts taking advantage of methods and tools developed by both groups. The book will hopefully inspire future research and development efforts, innovation by stimulating ideas.

We hope our readers will enjoy the book and will find it both interesting and useful.

Pavel V. Tsvetkov
Department of Nuclear Engineering
Texas A&M University
United States of America

Part 1

Instrumentation and Control

1

Sensor Devices with High Metrological Reliability

Kseniia Sapozhnikova and Roald Taymanov
D.I.Mendeleyev Institute for Metrology, Russia

1. Introduction

At present, a great number of embedded sensor devices provide monitoring of operating conditions and state of equipment, including nuclear reactors of power plants. The metrological reliability of measuring instruments built in equipment determines the validity of measurement information. The quality of production, operating costs, and the probability of accidents depend on the validity of measurement information coming to control systems. The validity is particularly important in such fields as nuclear power engineering, cosmonautics, aviation, etc. For some products in definite periods of their operation, even a short-term loss of confidence in measurement accuracy is unacceptable.

The key problems of the measurement information validity are related to the sensor metrological reliability, since their components age and their parameters drift with time. Sudden failures can also happen. All this can result in control errors. The sensor devices used to monitor the condition of technological equipment and the parameters of a technological process, are, as a rule, subject to a variety of influencing quantities. Possible consequences of these influences are, for example, depositions, magnetization, and so on. In some cases, the effect of the influence quantity can be weakened by a careful design of the sensor. For example, the rate of fouling of a sensor surface can be reduced by polishing the surface. However, it is not always possible to develop a sensor device immune to influencing factors over a long period of operation. Sometimes, economic reasons may play a role as well.

At present, the traceability of measurements is provided by periodic calibrations or verifications (hereinafter both of these procedures will be referred to as calibrations). Accordingly, within the period of operation the probability of a metrological failure depends on the length of the calibration interval (CI). The state of a secondary transducer can be verified by supplying electrical signals of reference values to its inputs. As demonstrated in (Fridman, 1991), between 40% and 100% of all measuring instrument failures are due to metrological failures. Improvements in production quality result in decrease of the number of failures, the share of metrological failures being increased because with the technology improvement the share of sudden failures decreases. It is not expedient to apply fundamental assumptions of the classical reliability theory (e.g., mutual independence of failure rates and stability of a failure rate) to measuring instruments. Usage of methods based on these assumptions leads to crude errors in the CI determination.

To decrease the risk of getting unreliable information, usually the CI is no more than 2-3 years. However, the cost of a sensor device calibration is typically 35–300 euro, and the

number of sensor devices is growing year by year. If a CI duration is constant, the proportion of operating costs spent on calibration will rise to an unacceptable level. In many cases, it is necessary to disrupt a technological process in order to carry out sensor device calibration. Such interference leads to additional costs.

The standard (ISO/IEC, 1999) states that it is ''the responsibility of the end-user organization to determine the appropriate calibration interval under the requirements of its own quality system". The guidelines (OIML, 2007) state that ''the initial decision in determining the CI is based on the following factors:

- instrument manufacturer's recommendation;
- expected extent and severity of use;
- the influence of the environment;
- the required uncertainty in measurement;
- maximum permissible errors (e.g., by legal metrology authorities);
- adjustment of (or change in) the individual instrument;
- data about the same or similar devices, etc."

Furthermore, it is recommended to adjust the initial CI for the process of operation ''in order to optimize the balance of risks and costs", due to a number of reasons, for example:

- the instruments may be less reliable than expected;
- the operating conditions may vary significantly from the manufacturer's recommended ranges, requiring an adjustment to the CI;
- the level of drift determined by instrument recalibration can demonstrate that longer CIs are possible without any increase of risk;
- it may be sufficient to carry out a limited calibration of certain instruments instead of a full calibration, etc.

However, in some cases, it is impossible to perform calibrations with a short CI, in order to obtain the data necessary for adjusting the CI value for the measuring instrument. For many modern complex technical processes, the mean value of continuous running grows. At present, for some processes the campaign duration has to be no less than 10 years.

Measuring instrument operation conditions can vary considerably over the course of several CIs. In industrial equipment, they can appreciably vary when upgrading the technological process, e.g., in case of production modernization. Operation conditions for sensor devices in ship nuclear power sets will depend on the intensity of the equipment use. For all the reasons given above, the end user does not want or has no possibility of affording the testing of each measuring instrument in order to provide grounds for optimizing the CI. Calibrations are expensive, but as the experience shows, the majority of measuring instruments (according to various estimates which vary from 60% to 80% for all instruments submitted for calibrating) does not need it. However, approximately 12% of measuring instruments have an error exceeding the permissible limits within the CI.

The contradiction is obvious. To reduce the costs associated with the interruption of a technological process and the calibration of built-in measuring instruments, it is desirable to calibrate as seldom as possible. However, unreliable information received by a control system from measuring instruments, can cause failures and large economic losses. To prevent this, it is necessary to check the measuring instrument state as often as possible.

It is impossible to settle this contradiction using trivial methods of calibration.

This chapter deals with non-conventional methods of improvement of the measuring information validity and possibilities to increase the sensor device metrological reliability on the basis of these methods.

2. Way to solve the problem. Self-check in biological and technical sensor systems

Various attempts have been made to decrease the labour costs of instrument calibration: in some cases calibration is performed without dismantling the sensor. For example, in (Karzhavin et al., 2007) it was proposed to design the thermocouple housing with an additional hole and to periodically insert a thin reference thermocouple into this hole. Such periodical sensor calibration procedures are an additional load for personnel. They may result in bending or damaging of the thermocouples or sensor displacement from the required location. All these undesirable outcomes may result in calibration errors.
In some sensors, a ''live zero" correction is made. For example, in a pressure sensor this can be performed by ''switching off" the pressure measured. However, such a procedure does not reveal and correct a multiplicative error.
Both of the above proposals cannot ensure checking the ''metrological health'' of a sensor within the CI.
One possible solution of the problem is suggested by taking into account the qualitative difference in the goals which metrologists try to achieve at various stages of the life cycle of a measuring instrument. The initial calibration of a sensor on completion of its manufacture is aimed at establishing and specifying the parameters corresponding to a ''normal condition" of the sensor. At the stage of online operation, the goal is different. It consists in revealing a metrological failure, i.e. deviation of the parameters from their initial values registered at the initial calibration, and, if possible, in correcting them. This goal can be achieved using unconventional methods.
In order to find possible ways to solve a problem it is necessary to have a criterion. Professor Wiener (Wiener, 1948) and the outstanding philosopher and writer Mr. Lem (Lem, 1980) demonstrated the value of applying an analogy between technical and biological systems. This analogy is very successfully being developed in evolutionary cybernetics (Red'ko, 2007; Turchin, 1977), as well as in bionics (biomimetics). As applied to sensors, a biomimetic approach enables to realize functions, structural elements and other features which mimic similar designs ''discovered" by nature in the past (Bogue, 2009; Stroble et al, 2009).
In our opinion, the similarity between the biological and technical evolutions forms not only a ''reference book" containing successful structural and functional decisions. In the case of metrology, it gives a necessary strategic criterion for development of measuring instruments and systems with high metrological reliability.
The time scales of biological evolution and of sensor device development are different. However, the tasks solved were similar: with time, the methods providing the survival in the nature and the methods providing metrological reliability of sensor devices built in the equipment intended to be used under changing environmental conditions, were improved.
In the early period, prolonging of life was achieved through the use of conservative methods of enhancing the "reliability". For example, protection of a turtle with a shell or protection of a sensor device with a sheath, certainly could prolong their life span.
However, the ambient conditions of measuring instruments as well as those of living organisms are characterized by a significant level of unpredictability. Adaptive methods

allow to take into account the variability of the environment (Taymanov & Sapozhnikova, 2010b). The adjustment of the insulating properties of an animal's pelt with the season increases the likelihood of survival under a changing environment, as does the active thermoregulation of measuring instruments.

The appearance and growth of biological intelligence relate to ensuring the survival under increasingly rapid environmental changes. Intelligence enables to forecast and take into account future changes of dangerous character, including those of an intelligence carrier state. ''Evolutionary change is not a continuous thing; rather it occurs in fits and starts, and it is not progressive or directional''. On the other hand, evolution ''has indeed shown at least one vector: toward increasing complexity'' (McFarland, 1999).

Developing the idea of the analogy between biological and technical evolution, it is possible to consider a direct analogy between the life span of a living organism and the lifetime of a measuring instrument, during which the measuring instrument is characterized by metrological serviceability and the absence of any maintenance requirement.

Then the purpose of artificial intelligence in a measuring instrument can be defined as ensuring the reliability of measurements for an extended lifetime. To achieve this ultimate purpose, it is necessary to analyze ''metrological health", to forecast future ''behavior", as well to provide a correction of an error and self-recovery of a measuring instrument.

The idea of applying ''intelligence" for increasing the reliability of measurement information formed by measuring instruments, appeared and started developing not long time ago. At this stage of technical evolution, it became necessary to extend considerably the lifetime of the ''weak points'' of measurement instruments. These ''weak points'' are sensor devices. At the same time, it became possible to solve this problem at the expence of the increase in the complexity of a sensor device.

Intelligence in the nature has developed in two ways: the formation of a ''collective mind" consisting of many living organisms, and the development of the intelligence (mind) of a separate individual. If the risk of extinction of individual living organisms is high, the ''collective mind" provides a way of preserving the experience gained and supporting the life of the species as a whole. Certainly, just such a type of intelligence started forming at the early stages of evolution, when the life span of an individual organism was short. A representative example is given by the swarming insects, i.e., the bees, which select reliable information by ''voting". The validity of information obtained by scout bees, depends on the number of bees obtained this information.

Formation of the ''collective mind" is the integration of a number of autonomous subsystems at a lower level (they can be different, to some extent) and the development of an additional control mechanism at a higher level (Red'ko, 2007). A similar approach to checking the reliability of information is applied in metrological practice. In nuclear power plants, a redundant number of sensors are integrated into a multichannel measuring system. In this case, the metrological malfunction of a sensor can be detected on the basis of deviation of its signal from the signals of the remaining sensors of the system, the readouts of the most part of sensor devices being considered to be reliable (Hashemian, 2006).

However, it is not always possible to form a significantly large ''swarm" of sensors to measure the same quantity. Information coming from a group of the sensors comprising a ''swarm", can also be distorted by external factors. Signals from a significant part of the ''swarm" may come with some delay that can lead to erroneous decisions, etc.

The results of ''metrological health" checks in a multichannel system of the sensor devices measuring different, but correlated values of the same or various quantities, have an additional error component caused by the measurement method. Its value depends on the accuracy of the relationship between the values of measurands. It is possible to decrease the above error component by making the values of measurand field characteristics closer to each other. The same is related to characteristics of the fields of influencing quantities. However, this procedure requires a considerable amount of costly time and effort, particularly if it is necessary to perform measurements at several points of a given measurement range.

In comparison with the ''collective mind", the intelligence of an individual has an advantage in searching for effective ways to survive under a changing environment. The ultimate check of ''metrological sensor parameters" deviation has been realized by humans and other creatures with a developed personality. In addition to the minimum of ''sensors" for the quantity to be measured, each sense organ is provided with supplementary sensors. The brain has a special mechanism for testing the stability of essential activity characteristics. This mechanism, known as an ''error detector", has been discovered by the famous Russian Academician Bekhtereva (Bekhtereva, et al., 2005). A person diagnoses a ''malfunction" of a sense organ such as an eye or ear, initially, through an unpleasant sensation caused by signals coming from these supplementary sensors. It should be noted that to provide the selection of video- or audio- information, these supplementary sensors are not required and in this implication they are redundant.

Similarly to the sense organs of intelligent living creatures, a measuring instrument with artificial intelligence distinguishes by the following features:

- it contains one or more basic sensors, as well as additional elements, e.g., additional sensors;
- these sensors and elements enable the generating and processing a measurement signal as well as a number of additional signals which carry information about the ''metrological health" of a measuring instrument.

Besides the method discussed, a living organism applies auxiliary ways for detecting deterioration in the functioning of individual sense organs, namely, analysis of video-, audio- and other information, coming through all organs of sense as well as analysis of the response of other members of a society.

In biological evolution, both types of intelligence considered above are developing in parallel. Sometimes, they supplement each other, but the intelligence of an individual has gained the priority and greatest pace of improvement. By the analogy, in technical set, it is the most perspective to apply the sensors with the individual "intelligence", joint in the system with the ''collective mind".

3. Metrological self-check

In a number of publications, the experience has described, which demonstrates the methods of sensor device serviceability monitoring on the basis of joint processing of the data obtained in multichannel measuring system (e.g., GOST R, 1996; Hashemian, 2005, 2006).

Development of sensor devices with a structure that enables, to some extent, to control their metrological serviceability within the process of operation has been started in Russia since 1980s (Druzhinin & Kochugurov, 1988; Sapozhnikova, 1991; Sapozhnikova et al., 1988; Tarbeyev et al., 2007). Later on, such activity was also expanded in the UK and USA as well

as in Germany, China and other countries (Barberree, 2003; Hans & Ricken, 2007; Henry & Clarke, 1993; Henry et al., 2000; Feng et al., 2007, 2009; Reed, 2003; Werthschutzky & Muller, 2007; Werthschutzky & Werner, 2009). In general, the above works are of an heuristic character. This circumstance impedes estimation of their efficiency.
Below, the methodology concept is considered, which enables developing the measuring instruments with a high metrological reliability provided by metrological serviceability checks and additional measures realized on the basis of such checks.

3.1 General terms and their definitions

The terms and definitions given below were presented at a number of international conferences, discussed with specialists of many Russian organizations, and included into the national standard of Russia, developed by the authors (GOST R, 2009). The most part of these terms was given in (Taymanov & Sapozhnikova, 2009, 2010b). Taking into account the above and availability of those papers, there are no detailed explanation of the terms and their definitions here.
Sensor: an "element of a measuring system that is directly affected by a phenomenon, body, or substance carrying a quantity to be measured" (VIM, 2008).
Sensor device ("datchik"): a constructively isolated (separate) unit that contains one or a number of sensors. (Further in this chapter, the term "sensor device" is applied. The term "datchik" is given here to facilitate understanding of the content of some publications referred below.) The sensor device can include secondary measuring transducers and material measures.
Adaptive sensor device: a sensor device the parameters and/or operative algorithms of which can change in the process of operation subject to signals from sensors, secondary transducers and material measures it contains.
Metrological serviceability of a sensor device in the process of operation: a state for which an error specified for this sensor device under operating conditions lies within some specified limits.
Critical error component: the most "dangerous" error component, i.e., a predominant error component or component tending to rise quickly. This component determines mostly a risk of getting an unreliable result of measurement. It can be revealed by analysis of the experimental investigations results as well as of scientific and technical information.
Metrological self-check of a sensor device: an automatic procedure of testing the metrological serviceability of a sensor device in the process of its operation, which is realized using a reference value generated with the help of an additional (redundant) embedded element (a sensor, secondary transducer, or material measure) or additional parameter of an output signal. The term "reference value" corresponds to the term given in (VIM, 2008). The reference values are determined and specified at the stage of a previous calibration. (Often, they use the term "self-monitoring" instead of the term "self-check". At the same time, the metrological self-check accompanied by evaluation of error or uncertainty is usually called "self-validation").
Correction of the sensor device characteristics can be made on the basis of the metrological self-check results if an error nature (multiplicative or additive) is known. The results of the metrological self-check can be applied as a basis for increasing or reducing the value of a calibration interval as well as for making forecasts of a remaining life time. The metrological self-check is realized in a continuous or test mode and performed in two forms, i.e., in the form of a direct or diagnostic self-check.

Test mode of metrological self-checking of a sensor device: connection of a sensor or secondary transducer to an additional built-in sensor or built-in material measure, or injection of a test signal the relationship between which and a measurand or its variation is known with a required accuracy. The test mode assumes an interruption of an applied measurement procedure for a time interval within which the procedure of testing is performed.
Metrological direct self-check of a sensor device: a metrological self-check performed by evaluating the deviation of a measured value from a reference value generated by an additional (redundant) element (by a sensor or material measure) of a higher accuracy. Therefore, it provides an automatic check of the total error of a sensor device under operating conditions.
Metrological diagnostic self-check of a sensor device: a metrological self-check performed by evaluating the deviation of a diagnostic parameter characterizing the critical error component from the reference value of this parameter, established at the stage of a previous calibration. The metrological diagnostic self-check is performed without any embedded elements of a higher accuracy.
Data-redundant sensor device: a sensor device enabling to generate a diagnostic parameter on the basis of an additional output signal parameter or of a built-in means. In case of connecting such a sensor device to a signal processing unit, the latter can provide realization of the metrological self-check function.
Intelligent sensor device: a sensor device with the function of the metrological self-check.
Reliability of metrological self-check of an intelligent sensor device: a qualitative estimate indicating the level of a risk to get the results of metrological self-check of a sensor device, which are inconsistent with an actual condition of this sensor device.

3.2 Metrological self-check methods

In (Henry & Clarke, 1993), a group of methods which can be applied to detect and identify the faults of a sensor or system of sensors and actuators, is given. These methods are based on analytical redundancy, a priori knowledge, or measurement aberration detection. Analytical redundancy exploits the implicit redundancy in the static and dynamic relationships between measurements and actuator inputs using a mathematical model. A priori knowledge relate to information "concerning operational conditions and associated fault modes, patterns of signal behaviour characteristic of particular faults, or historical fault statistics". Measurement aberration detection permits to reveal faults, taking into account how they change the behaviour of the signal (e.g., bias, noise, etc.). A self-validating Coriolis flow meter developed on the basis of the above sources of additional information provides the self-diagnostics and diagnostics of corresponding actuators, the result of measurements being accompanied by a value of uncertainty (Henry et al., 2000).
In (Feng et al., 2007), sources of information intended for diagnostics of sensor device faults are classified in the following way:

- hardware redundancy (e.g., combination of a thermocouple and resistance thermometer);
- analytical redundancy taking into account a known relationships between the signals of several sensors or the signals of sensors and parameters of a technological process model;
- information redundancy of a sequence of sensor device signals which is revealed with the help of mathematical methods.

In the publications of the authiors of this chapter (Taymanov & Sapozhnikova, 2009, 2010a, 2010b), it is emphasized that the metrological self-check can be realized only on the basis of the redundancy that can be just of temporal, spatial, and informational type. The redundancy of the above types can be used separately as well as in any combination. Correspondingly, the methods permitting to organize the metrological self-check are subdivided in accordance with the types of redundancy.
The spatial redundancy is provided by usage of additional sensors, secondary transducers, and/or material measures which occupy in a sensor device housing or directly in a measurement zone an additional space that is comparable with a minimally required one. Since these additional elements complicate the intelligent sensor device structure (as compared with its trivial non-intelligent analogue), the spatial redundancy is often named as structural. This term is accepted by the authors and will be used bellow.
The temporal redundancy is provided by additional measurement operations, which should be carried out at a frequency or within the range of frequencies exceeding a corresponding minimal value required to perform measurements with a specified accuracy.
The informational redundancy is provided by using an additional dependence of a certain parameter of a sensor device signal on a measurand. Since all types of redundancy provide additional information, necessary to perform the metrological self-check, hereinafter the term "informational redundancy" is replaced by the term "functional redundancy". "External" information which is known for a specific equipment and mode of its operation (variation limits for parameters of a sensor device signal, correlation between the signals of various sensors and actuators, etc.) should be considered as a kind of the functional redundancy too. It should be noted that the accuracy of this "external" information is usually low. In a number of cases, it is subjected to interferences due to the remoteness of various sensor devices and actuators from each other. Application of these data is laid on customers. The competence of the customers in the field of the metrological self-check organization gives basis for doubts. The functional redundancy based on the " external" information can help to detect crude faults of sensor devices and to perform the equipment trouble-shooting. However, it is not very efficient for evaluating the sensor device metrological serviceability.
The metrological self-check (self-monitoring or self-validation) can be also organized in a multichannel measuring system containing a set of sensor devices on the basis of the types of redundancy considered above. However, in (Taymanov & Sapozhnikova, 2010a) it is shown that the metrological self-check of autonomous sensor devices is the most efficient means, and the self-check provided in a system of sensor devices can be a useful addition to it.

4. Direct metrological self-check. Essence and specific features

In organizing the direct metrological self-check, it is necessary to have arguments according to which a sensor or material measure, used as a reference measurement standard, is more reliable as compared with a sensor, the metrological serviceability of which is checked. As a rule, the attempts to realize self-checking in such a form meet some limitations concerning a kind of measurand, a speed and range of its variation, and others.
A typical example of the direct self-check realized in the test mode on the basis of both structural and temporal redundancy, is the eddy current sensor device of distance to a conducting surface of a target (Druzhinin & Kochugurov, 1988). This sensor device contains a drive inductance coil, sensor coil and target simulator made in the form of a switched flat

inductance coil. The simulator is fixed between the drive coil and target. The piece fixing the distance between the drive coil and simulator serves as a length measure. The simulator coil being open (disconnected), the distance to the target is measured using a signal received by the coil. In the test mode, the simulator coil is closed, and it becomes a shield for the target. The output signal in this situation is assumed to be the diagnostic parameter. It is possible to estimate the metrological serviceability of the sensor device on the basis of the deviation of the diagnostic parameter determined in the process of sensor device operation from the reference value of the diagnostic parameter measured at the stage of a previous calibration.

In a temperature sensor device, the direct self-check is realized in the continuous mode on the basis of redundancy of structural and temporal types (Bernhard et al., 2003). In this sensor device, there is an embedded cell (capsule) with a metal, the fixed point of which is known with a high accuracy. This fixed point is taken as the reference value. When the environment, the temperature of which is measured, is heated or cooled and the metal melts or hardens in the capsule, the speed of measured temperature changes significantly decreases, forming a "plateau" in a diagram "temperature – time". When the speed of temperature variation does not exceed a certain permissible minimum value that allows to register the "plateau", then it is possible to estimate the metrological serviceability of the sensor device on the basis of the deviation of the temperature value measured in the metal fixed point from the reference value. It is possible to apply correction of the sensor device characteristic on the basis of the evaluated deviation only if the type of the originated error (additive or multiplicative) is known. Nevertheless, the application of the direct self-check in such cases enables the time interval between calibration procedures to be increased, since the metrological reliability of the applied additional material measures can be assumed to be greater than that of the sensors contained in the sensor device.

Direct self-check can be useful for increasing the reliability of measuring the parameters of dynamic processes. In particular, if a range of the temperature to be measured is significantly wide and there is a need to trace its relatively fast changes, the reliability of measurements can be increased by applying a sensor device with structural redundancy. In this case in a sensor device body, in addition to a low-inertial sensor (thermocouple), it is necessary to place a more inertial platinum resistance thermometer (PRT) being at the same time more precise (Barberree, 2003). The latter will act as a reference measurement standard. Within the time intervals when the speed of measured temperature variation is so low that it does not significantly affect the error of temperature measurements made with PRT, the values obtained with the PRT are used as the reference ones. As a result, a sensor device of such a type can provide measurements with a time lag close to the thermocouple one, but with the accuracy typical for the PRT.

Sometimes, direct self-check can be realized on the basis of functional redundancy. For example, in any PRT both the resistance and parameters of its output noise spectrum depend on temperature. The temperature measurements made in terms of noise spectrum parameters are more accurate than those made in terms of resistance. That is why a value of the temperature measured in terms of a noise spectrum can be taken as a reference value. However, measurement of temperature on the basis of the noise spectrum requires for much more complicated unit for data processing.

The direct self-check is analogous to the conventional calibration procedure. The only difference consists in performing the procedure with the help of the built-in reference measurement standards available in an apparent or implicit (as in the last example) form.

5. Metrological diagnostic self-check. Essence and specific features

The self-check of such a type is a qualitatively new procedure in providing the traceability of measurements. To select a diagnostic parameter characterizing the critical error component, it is necessary to measure two or a number of original parameters, the values of which depend on the value of a measurand to be determined by a sensor device. At the same time, these original parameters depend on factors causing the growth of the critical error component in different ways. Additional procedures of measurements are organized on the basis of structural, temporal, or functional redundancy revealed or introduced in a device that should be capable to perform the metrological diagnostic self-check (MDSC). The MDSC does not assume any usage of reference measurement standards of a higher accuracy. The additional sensor or material measure can have the metrological reliability that is close to that of the sensor, metrological serviceability of which is under checking. The same statement can be related to the accuracy of them.

When the critical error component of duplicate sensors is drift, which for a group of sensors is characterized by a random distribution of the level and sign, the MDSC can be organized by arranging these sensors in a sensor device. In this case, it is possible to use a mean value of the deviation of output signal values from a mean value of the output signals as a diagnostic parameter. The metrological diagnostic check can be performed by estimating the difference between the diagnostic parameter values estimated in the process of operation and at the stage of a previous calibration.

However, in addition to the random deviation of the duplicate sensor parameters combined in a sensor device, a monodirectional drift of the same parameters, which cannot be revealed, may take place. In this case, the efficiency of the MDSC can be increased by application of the sensors similar in the accuracy but differing in their design, production technology and /or principles of operation. The probability that an error of such sensors will change equally, is very small.

The MDSC based on the structural redundancy in the test mode is realized, for example, in a pressure sensor device suggested in (Lukashev et al., 1984). In this sensor device, a diaphragm is rigidly connected with a plunger, the displacement of which inside of an inductive measuring transducer generates an output signal. At the same time, the plunger moves inside an electromagnet coil. Supplying the current of a fixed value to the electromagnet coil, it is possible to move the diaphragm simulating a certain increase of pressure, i.e., realizing the test mode. Within the time period of the test mode, the variation of an error value is assumed to be negligible in comparison with the permissible measurement error. If the conversion function "displacement of the diaphragm – pressure" is linear, then the self-check of such a type can detect a metrological fault for a case when the critical error component is multiplicative. The variation of an output signal as a result of the fixed variation of the electromagnet current can be used as the diagnostic parameter.

The MDSC of the same type can be realized in a data redundant capacitive sensor device measuring a distance to a flat conducting body. Usually, the critical error component in such a sensor device is caused by fouling of the surface of electrodes. To perform the self-check in the continuous mode it is possible to use a two-channel sensor device with the electrodes shifted relative to one another in a direction perpendicular to their surface (Sapozhnikova & Taymanov, 2010a). As the diagnostic parameter, it is possible to use the difference of voltage values at the shifted electrodes at a distance measured with the help of one of them.

The MDSC can be organized also on the basis of structural redundancy by combining the sensor under check and additional sensor of different principles of operation in one unit. This method of self-check is applied in an ultrasonic vortex gas flow meter (Hans & Ricken, 2007). A ratio of the value of vortex flow velocity to that of vortex frequency can be used as the diagnostic parameter.

The above examples show that realization of the MDSC on the basis of structural redundancy assumes the combination of the sensor under check and additional sensor into one unit. The main requirement for them and their packaging is reduced to the following. Under the impact of influence quantities generating a critical error component, the variation of conversion functions of the additional sensors should significantly differ from the corresponding variation of the conversion function of the sensor under check.

Realization of the MDSC on the basis of temporal redundancy assumes application of more wideband or fast-response sensors than it is necessary for a non-intelligent analogue. The MDSC of such a type is used in a tachometer sensor device of flow rate. Its critical error component is caused by the wear of a bearing. The growth of the critical error component is accompanied by increase of vibration, which results in increase of the period and amplitude dispersion of the sensor signal. Usually, the variation of the flow rate measured for some revolution periods of a rotating element of the flow meter is distinctly less than the permissible error of measurement. In this case, the period and amplitude dispersion of the sensor signal can be used as the diagnostic parameters.

If the critical error component of the temperature sensor device is caused by probable damages of a contact in a network of the sensor device, then the MDSC can be performed by using the temporal redundancy too. To achieve this, speed of a sensor device signal change can be used as the diagnostic parameter. In the process of sensor operation it should be compared with the maximum possible (limited by the equipment time lag) speed of the environmental temperature variation (Taymanov et al., 2010a).

The MDSC basing on the functional redundancy implies usage of an additional dependence of a certain output signal parameter on the measurand. This additional conversion function can be revealed in a sensor, introduced artificially or formed using a modulation of a measurand.

The MDSC based on the functional redundancy can be realized in an eddy current sensor device, which determines the distance to a metal non-magnetc target. The critical error component often arises due to a variation of the impedance of the inductance coil parameters in the process of operation, e.g., wind short-circuits or core parameter variation. The critical error component can be evaluated by measuring the active and reactive components of the output signal and comparing their relationship with the reference value obtained at the stage of a previous calibration.

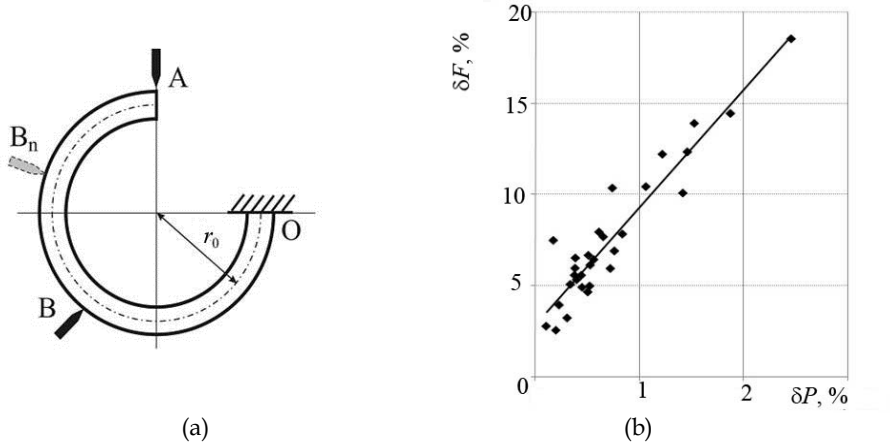
The MDSC in a capacitive or eddy current sensor device measuring a distance to the conducting target surface, can also be provided on the basis of functional redundancy by a modulation of the measurand. Such an approach is possible when the surface of the target cyclically moves with regard to the sensor device, e.g., the target is a surface of a rotating shaft. In this case, a step should be cut out on a section of the shaft or a strap has to be fixed there. When the shaft rotates, the distance to the target, which should be measured, is cyclically changing by a known value, e.g., by the value of a step depth. This can be used to estimate metrological serviceability of the sensor device.

The examples considered give a general picture of the possibility to organize the metrological diagnostic check.

6. Metrological diagnostic self-check of pressure sensor device

An analysis has demonstrated that the main sourse of error of the pressure sensor devices with elastic sensors is the residual deformation of sensors (Baksheeva et al., 2010). The method of metrological diagnostic self-check is illustrated below by an example of a sensor device with the Bourdon tube. As it is known, one end of such a tube is rigidly fastened in the device construction and the second end is free. Measuring the displacement of the free end, the pressure supplied to the tube is determined (Andreeva, 1981; Bera et al., 2009; Hashemian, 2005). The main reason of the error of such a sensor is the residual deformation of the tube (Baksheeva et al., 2010). When supplying the pressure, two differently directed deformation processes arise in the tube.

The first process is the non-uniform variation of curvature of a tube middle line. The tube does not return into its initial position. Longitudinal fibers undergo some increase of their length. Residual deformation of such a kind results in displacement of ''zero" of the conversion characteristic and increase of the Bourdon tube sensitivity, which in its turn, generates an additive and multiplicative components of error. The second process is the non-uniform variation of transverse cross-sections along the tube. The size of a small semiaxis of the tube cross-sections is increased and consequently, the form of the cross-sections approaches to a circle one. The residual deformation of this type results in displacement of ''zero" and decreases the Bourdon tube sensitivity. Thus, both the additive and multiplicative components of error have place here too. The degree of influence of the above processes on the state of the tube depends on distinguished features of a particular tube, as well as on its operating conditions. The multiplicative components of error partly compensate each other, and the additive components of error are summed up. As a result, the additive component of error should be considered to be the critical one.



(a) (b)

Fig. 1. Pressure sensor device with the metrological self-check
(a)Location of the points of the Bourdon tube displacement measurement
(b)Relationship between the relative variation of the diagnostic parameter (δF) and relative error δP of pressure measurement

The suggested method of the metrological diagnostic self-checking of an intelligent pressure sensor device on the basis of the Bourdon tube (Baksheeva et al., 2010), is based upon the diagnostic parameter (DP) in the form of a ratio of output signal values determined in various points of the sensor (Fig. 1a). Performed mathematical modelling has enabled the location of additional points of displacement measurement to be optimized.
A prototype was made, a displacement signals in which were taken from two points of the Bourdon tube with the help of mechanical transmissions and two differential transformer transducers of displacement. Dependence of relative DP variation on relative error of pressure measurement is shown in Fig. 1b. The diagram indicates that the dependence of the relative variation of the DP on the relative error of pressure measurements is close to a linear one.
The method considered provides the principal possibility for developing an intelligent pressure sensor device with the functions of self-correction and forecast of its metrological state (with gradual accumulation of measurement data).

7. Metrological diagnostic self-check of temperature sensor device

Sensors of a resistance sensor device measuring a temperature are usually made on the basis of metal wire or thin film. For the temperatures up to 150 ^{o}C , they usually apply copper, while for the temperatures to 600 ^{o}C, platinum is used. If one sums up the known data (Berry, 1982; Crovini et al., 1992; Hashemian & Petersen, 1992; Li et al., 2010; Mangum, 1984), then within the working temperature range up to T_{max}<(0.25-0.3)T_m, where T_{max} is its upper limit and T_m is the melting point of platinum, which corresponds to T_{max} ≅(450-500) ^{o}C, the main processes affecting the change of the PRT error in its operation can be divided into two main groups.
The first group includes the processes leading to destruction of a thin surface layer of conductors and variation of its conductive properties with regard to all sensors of the sensor devices. They are the following: surface oxidation, sublimation of surface substances, contamination of the surface layer by the diffusion of oxides and mixtures from the ceramic fill-up, mechanical damage of the surface, and so on. Under the influence of error sources, which refer to the first group, in the course of time the specific resistance of the surface layer begins to exceed significantly the specific resistance of a conductor material. This can be represented as a certain equivalent decrease of an area of the cross-section of the platinum wire, which correspondingly increases its resistance.
Other processes (which are not connected with the destruction of the surface layer) leading to an unexpected change of resistance of some sensors, in particular caused by consequences of the technological spoilage, refer to the second group. When the processes of the first group dominate, a critical error component is the deviation of the platinum wire resistance, which is caused by variation of the properties of the surface layer.
In (Taymanov et al., 2011) it is proposed to use the structural redundancy for realizing the function of metrological diagnostic self-check, i.e., to make an intelligent sensor device on the basis of two or a number of sensors differently sensitive to factors influencing on the growth of the critical error component. To realize the function of metrological diagnostic self-check in the process of operation, the DP β, that depends on the values of signals coming from various sensors, is calculated, for example, $\beta = R_1/R_2$.

At the stage of the original calibration of the intelligent temperature sensor device, which corresponds to the start of operation, a nominal value β_0 of the DP is determined. A relative

deviation $\delta\beta$ of the DP from the nominal value is rigidly connected with the error. In the process of operation the check of the metrological serviceability is performed by determining $\delta\beta$ at a temperature measured and comparing it with a permissible relative deviation.

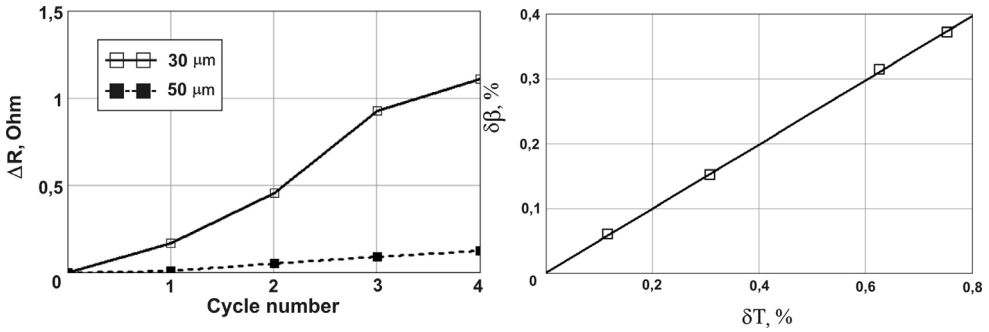
When the value $\delta\beta$ exceeds the permissible limit or approaches to such a limit, it is necessary to perform an unscheduled calibration of the sensor device even if the specified interval has not come to its end. When the specified calibration interval comes to the end and $\delta\beta$ does not exceed a permissible value, then this fact can become an argument for significant increasing the corresponding interval and using the sensor device further.

In the general case (for various designs of the PRT), it is necessary to provide a different ratio of the cross-section area to the perimeter of the cross-section of sensors included into sensor device. In each sensor the destruction rate (rate of resistance variation) of the conductor surface layer, weakly depends on the geometrical parameters of the conductors themselves. The depth of the destructed layer is small as compared to linear dimensions of the conductor cross-section. Then for the simplest version of the sensor device including two sensors

$$\delta\beta = \left|\frac{\beta-\beta_0}{\beta_0}\right| = \left|\frac{\beta}{\beta_0}-1\right| \approx \left|\frac{1-a\frac{P_2}{S_2}}{1-a\frac{P_1}{S_1}}-1\right| \tag{1}$$

where a is the assumed thickness of the surface layer subjected to the destruction; P_1, P_2 and S_1, S_2 are the perimeters and areas of the cross-sections of the sensors having different sensitivity to factors influencing on the growth of the critical error component, correspondingly.

Provided the relationship between the $\delta\beta$ value and error δT of the temperature measurement is known, then, using a value of $\delta\beta$ determined experimentally, it is possible to introduce a corresponding correction into a measurement result.



(a) (b)

Fig. 2. The characteristic dependencies obtained in the experiment at temperature of 450 °C.
(a) dependence of the resistance increment of the sensors on the cycle number;
(b) dependence of the relative DP deviation on the relative error of temperature measurements.

For preliminary evaluation of the efficiency of the considered metrological self-check method, there were carried out experimental investigations of the sensor device prototypes. This was done in the mode of a forced load, "heating up to 600°C - cooling", with a cycle time of 70 – 80 hours. The prototypes contained two spirals (each having a resistance of 100 Ohm) which were made of a platinum wire of 50 and 30 μm in diameter. In each cycle the resistance of each spiral was measured at temperature of 0 and 450 °C determined with a reference thermometer. Then the values of β, $\delta\beta$, δT and resistance deviation (increment) ΔR for each spiral were calculated (Fig.2). Fig. 2b indicates that the dependence of $\delta\beta$ on a relative error δT of temperature measurements is close to a linear one.
Thus, the method considered provides the principal possibility to develop an intelligent temperature sensor device with the functions of self-correction and forecast of its metrological state.

8. Combined methods of the metrological self-check. The system intended for measuring control rod position in a nuclear reactor

In Russian-built pressurized water reactor (WWER-1000), a control rod (CR) is moved by a linear stepping electromagnetic drive. A drive rack is connected with the CR. CR position is determined by a special measuring system that consists of a sensor device, a magnetic shunt mounted inside the rack cavity, and electronic processing unit. The CR position sensor device and some drive components (such as the rack and a transfer unit) are located in the coolant. The coolant temperature is up to 325 °C, and its pressure is up to 18 MPa. Also the coolant is under a strong radiation flux.
To improve currently used reactors and to build new ones, it is very important to increase sensor device lifetime to about 40-60 years (i.e., to the lifetime of a reactor vessel). In addition, it is necessary to improve the sensor device accuracy and metrological reliability, as well as provide extended diagnostics. All this measures would make nuclear reactor more effective and safe. Under the leadership of the authors, the intelligent system intended for measuring control rod position (IS) with the properties said above was designed and developed (Taymanov, et al., 2007). The originality of the applied technical decisions is confirmed by a number of corresponding patents.
A developed sensor device of the DPL-KV type is based on eddy current principle. In the sensor device, the combination of structural and functional redundancy is applied, which provides the possibility to perform metrological self-check and self-correction. The sensor device is a data-redundant one and includes 18 inductance coils inside a sealed housing (see Fig. 3a).
To provide the required lifetime and metrological reliability, the sensor coils are made of specially designed heat-resistant wire. The wire is based on Ni-Cr alloy covered by nonorganic insulation. The Ni-Cr wire was chosen because of its high stability under temperature variations and long lifetime (according to estimations based on experimental data and calculations, the lifetime of the wire at temperatures about 325 °C is expected to be more than 250 years). All the sensor coils are located much lower than the electromagnets of the drive. This decreases the effect of noise from the electromagnets.
The accuracy of CR position measurements was achieved by using multi-component magnetic shunt located inside the rack cavity (Fig. 3a). This shunt consists of a set of bushings made of magnetic and nonmagnetic steel. The coils and the set of bushings form a single-track code chain. The sequence of bushings in this shunt is quasi-random. The total

length of the shunt is slightly more than 2 m. For significantly longer shunts, the deeper rack boring is required to give enough space for the bushings. That would increase the risk of rack bending and decrease the rack life.

To increase the metrological reliability of the sensor device, the number of coils is redundant. Due to this redundancy, the sensor device keeps working even in case of wrong data (i.e., code bit distortion, which could happen if any signal wire breaks or if any coil fails). As a result, the sensor device has an ability to keep its measuring accuracy within reasonable limits in case of any single failure. That is the sensor device fault-tolerance (recoverability).

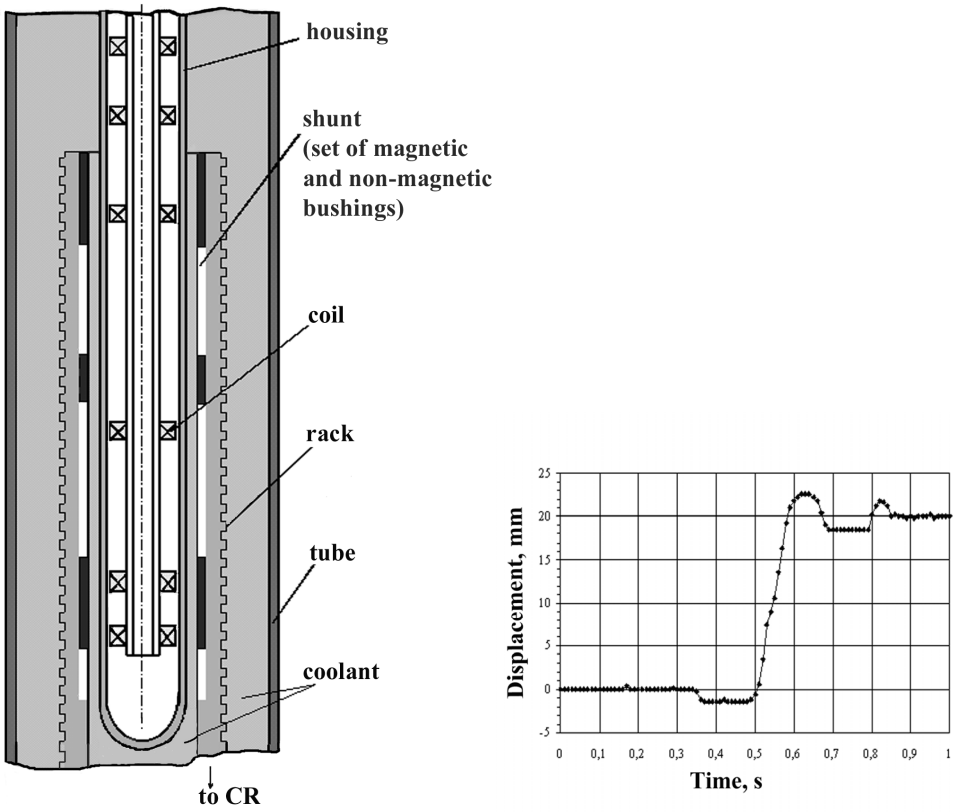
The electronic processing unit is located in a cabinet outside the containment. It consists of amplifiers, input and output transducers, a digital-analog power supply of the sensor device, and a microcontroller for data acquisition and processing (Sapozhnikova et al, 2005b). All the boards are galvanically isolated from outside devices. The power supply frequency was optimized to maximize signal-to-noise ratio. The microcontroller processes all available sensor device data. Data processing results from all the processing units included in several ISs, are transferred to a higher level controller. Such a design enhances the level of reliability and the effectiveness of control and protection system modernization. As a result, the IS provides:

- CR position measurement (accuracy is within 0.3% under all possible conditions);
- metrological diagnostic self-check of the reliability of CR position measurements (it is not necessary to calibrate the sensor device in the process of operation);
- maintenance of operating integrity if any signal wire breaks or any coil fails;
- automatic correction of a sensor device, shunt and processing unit parameters (this eliminates the influence of temperature variations, defects of joints, material and component aging); filtration of various noises;
- self-diagnostics of the IS as a whole with failure localization; generation of textual recommendations for malfunction elimination;
- assessment of the condition of main drive components (rack teeth, latches of the transfer unit and, partly, electromagnets) according to the step-by-step rack reciprocation diagrams (Fig. 3b); drive condition diagnostics (step missing or delay, as well as teeth slippage); checking of the CR and rack coupling;
- sensor device to processing unit connection diagnostics; control connection diagnostics;
- measuring and recording of CR drop time diagram (this allows diagnosing the condition of a guide sheath and rack curvature in case of CR emergency shut-down or spontaneous drop); determination of the top and bottom oscillation points during the CR damping process (this allows the condition of a rack damper to be diagnosed); checking whether the CR has fallen down on an arresting device.

The metrological and technical diagnostics of the sensor device and microcontroller unit generally consists in comparison between:

- the identified code combinations corresponding the CR position and the specified reference code combinations;
- the code combinations related to consequent control rod positions and the specified reference code combinations;
- the number of real steps made by the CR and the number of corresponding commands;

- current sensor coil parameters and their reference values determined at the original calibration.



(a) (b)

Fig. 3. System for measuring control rod position in a nuclear reactor
(a) simplified scheme of sensor device and rack with shunt
(b) diagram of drive rack: a step up

Fig. 3b illustrates the diagnostic capabilities of the IS on the basis of the displacement diagrams analysis. The diagram enables:

- determining the actuation time of the transfere unit latches,
- checking the correctness of the response to an electromagnet current cyclogram,
- checking the control rod and rack coupling.

The ability to obtain such diagrams is determined by both the high displacement sensitivity of the sensor device and the fact that the time interval between two consecutive control rod

position measurements is very short. In case of the drive fault, the shape of the diagram is changing. This makes it possible to find out the origin of the fault or to reveal the incipient malfunction (even before appearance of a significant failure). Information about all the CR moves, control commands, operation modes, occurred malfunctions or failures as well as operator's actions are logged in a "black box" recorder. At the same time, the IS estimates the drive operating time by accumulating the parameters like the number of drops, steps made, input control signals, etc.

The real time CR position is displayed on a front panel. Each IS can be connected to a local network. With the help of the network, the ISs can perform cross-system diagnostics. This improves the IS fault-tolerance. For instance, the local network gives an opportunity to inform operators about the wrong positions of CR, including the case of CR position mismatch in the control group as well as of any CR slipping down from the end switch. Based on diagnostic information obtained during system operation, an individual "registration certificate" is automatically issued for each drive. This certificate contains an assessment of the drive condition as well as recommendations for operators how to carry out a preventive maintenance.

Three ISs operated for many years at the power unit of the Kalinin NPP in Russia and were highly appraised by specialists. For that time interval, the first modification of the processing unit was replaced by a new one. The software parts related to diagnostics were improved. During the operation period, sensor signals varied insignificantly, and a tendency to stabilize the parameters was noticed. During the last years, the average change of resistance of sensor coils was less than 0.2% per year. Extrapolation of the resistance-time function for 60 years shows that the predicted sensor resistance variation is less than 3.5%. With the ability to automatically correct each individual sensor parameter variations within about 25%, the sensor device lifetime is much longer than it is required.

The use of the ISs improved the service effectiveness. It was more convenient for the stuff to work with textual recommendations from IS in case of malfunction. When the emergency shutdown of the power unit happened, the IS diagnostic capabilities helped to localize the failure even outside the ISs. Monitoring abilities are sufficient to extend the equipment lifetime by switching from pre-assigned lifetime to prediction of the state during future fuel cycle. As a result, the power plant can utilize equipment capability to the very end. In particular, the assessment based on the IS "black box" data at the Kalinin NPP gave the basis to increase significantly a projected lifetime of transfer unit and rack.

The additional study has shown that the electromagnet temperature can be decreased if a special inexpensive auxiliary component is added to the electromagnet.

Altogether, the developed technical solutions enable the lifetime of the equipment to become equal to the lifetime of the reactor vessel. Some additional information with respect to the IS considered has been given in the paper presented at the IAEA meeting (Sapoznikova et al., 2005b). The main ideas used in the IS can be applied to the control and protection systems of other reactor types.

9. Registration of self-check results. Status of measurement results

An estimate of the measurement error obtained in calibrating a given measuring instrument, cannot be transferred to the measurement results obtained with the help of

this instrument significantly later in the process of operation, since the instrument error component changes with time. The metrological self-check results are characterized by some error too.

It is not necessarily the case for the error to be determined quantitatively according to the metrological self-check data. For a significant part of applications, the qualitative estimate of the measurement reliability, by giving a certain "measurement value status" to the result of measurement, is expedient. For the first time, this concept was introduced in (Henry & Clarke, 1993). The following gradations of the status are recommended there: secure, clear, blurred, dazzled, blind. In the joint paper of Oxford and St.Petersburg scientists (Sapozhnikova et al., 2005a) a comprehensive reasoning of the necessity to introduce the measurement value status is given and some details of definitions and recommendations are proposed. It is noted that the number of status gradations should depend on the number of human operator's actions required in response to information about the measurement value status. The number of responses is usually no more than 5.

The status called "confirmed" indicates that a measurement result has been confirmed by additional information about the metrological serviceability of an intelligent sensor device or intelligent multichannel measuring system, and a risk to use an unreliable measurement result is negligible. This status is desirable in making very important decisions on equipment control. The status "confirmed" can be given to a measurement result obtained from a sensor device or measuring system when information at their output shows that they are in a "healthy" state.

The status called "normal" indicates that a risk to use an unreliable measurement result is small, which allows, for example, a decision on equipment control to be made in ordinary situations. This status can be given to the measurement result obtained within the calibration interval from a sensor device or multichannel measuring system, the metrological serviceability of which is not automatically checked in the process of operation.

The status called "orienting" indicates that a risk to use an unreliable measurement result increases due to a defect in a sensor device or multichannel measuring system, but the result of measurement can be applied for an orienting estimate of the equipment condition and that of the technological process under control. The "orienting" status is sufficient for making a decision in the case, for example, when parameters of the technological process are far from the borders allowed. Giving the status "orienting" to the measurement result, indicates the need to perform the maintenance of a sensor device or measuring system as well as to set the terms of this maintenance.

The status called "extrapolated" indicates that as a result of measurement they use the result obtained by extrapolating the data from the preceding time interval, since received information is unreliable during the known time interval that is rather short. The status "extrapolated" gives grounds, for example, to delay making a very important decision on equipment control before receiving reliable information or to make a certain cautious decision, orienting by a hypothesis that within this known time interval the condition of the equipment and flow of the controlled technological process do not change significantly.

The status called "unreliable" indicates that a risk to use an unreliable measurement result is great. The decision should be made to perform the maintenance of a sensor device or measuring system.

Status gradations can be joined into three groups which demonstrate the level of risk:

- status "confirmed" or "normal";
- status "orienting" or " extrapolated";
- status "unreliable".

Furthermore, the results of the metrological self-check can include:

- an estimate of the error (taking into account a correction when it was made) or critical error component;
- time when the corresponding estimate was obtained;
- an estimate of a residual metrological life;
- history of metrological self-check data.

10. Conclusion

The technological expansion has led to the situation, when the conventional methods of metrological assurance have ceased to satisfy the high requirements of nuclear power engineering, astronautics and a number of other fields of science and industry for the metrological reliability of measuring instruments. The measurement information validity becomes insufficient.

The similarity of the evolution of measuring instruments and biological sensor systems has created a basis for forecasting a significant complication of sensor devices and growth of the need for intelligent sensor devices and intelligent multichannel measuring systems with the metrological self-check.

This chapter deals with the general approach to the development of intelligent sensor devices. This approach is illustrated by a number of examples of the measuring instruments including those developed under leadership of the authors, namely, the temperature and pressure sensor devices as well as the intelligent system intended for measuring the position of control rod in a nuclear reactor.

It is shown that in the process of operation, the sensor devices with the metrological self-check can provide:

- practically continuous check of the measurement information reliability;
- forecast of the metrological state of a sensor device on the basis of the self-check results obtained in the previous period of time;
- automatic correction of the sensor device parameters (in a number of cases).

A growth of the need for intelligent and data-redundant sensor devices is confirmed not only by the examples showing that in various countries such devices and corresponding standards and guides (BSI, 2005; GOST R, 1996, 2009; MI 2021, 1989; VDI/VDE, 2005) were developed.

An increasing number of publications devoted to the topic considered, as well as organization of special sessions at international conferences and preparation of new standards (in particular, e.g., the Russian draft standard "State system for ensuring the uniformity of measurements. Intelligent sensors and intelligent measuring systems. Methods of metrological self-checking"), indicate the growth of this need too.

Under the conditions of economics globalization , the enhancement of requirements for the operating safety of various equipment, especially, nuclear reactors, obliges scientists and engineers to develop unified international requirements for standardizing the characteristics

of self-checked sensor devices and multichannel measuring systems as well as corresponding terms and definitions with respect to these instruments.
To our point of view, the development of intelligent measuring instruments is a natural stage of measurement technique evolution.

11. References

Andreeva, L.E. (1981). *Elastic Elements of Measuring Instruments*. Moscow: Mashinostroenie. (in Russian).

Baksheeva, Yu.; Sapozhnikova, K. & Taymanov, R. (2010). Metrological Self-Check of Pressure Sensors, *The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies,* Stratford-upon-Avon, England.

Barberree, D. (2003). Dynamically Self-validating Contact Temperature Sensors, *Proceedings of the Conference "Temperature: Its Measurement and Control in Science and Industry"*, No. 7, AIP Conference Proceedings, Melville, New York, pp. 1097-1102.

Bechtereva, N.P.; Shemyakina, N.V.; Starchenko, M.G.; Danko, S.G. & Medvedev, S.V. (2005). Error Detection Mechanisms of the Brain: Background and Prospects, Int. J. *Psychophysiol*, No. 58, pp. 227-234.

Bera, S.C.; Mandal, N.; Sarkar R. & Maity, S. (2009). Design of a PC Based Pressure Indicator Using Inductive Pick-up Type Transducer and Bourdon Tube Sensor, *Sensors & Transducers Journal*, Vol. 107, No. 8, pp. 42-51, ISSN 1726-5749.

Bernhard, F.; Boguhn, D.; Augustin, S.; Mammen, H. & Donin, A. (2003). Application of Self-calibrating Thermocouples with Miniature Fixed-point Cells in a Temperature Range from 500 °C to 650 °C in Steam Generators, *Proceedings of the XVII IMEKO World Congress,* Dubrovnik, Croatia, pp. 1604-1608.

Berry, R. J. (1982). Oxidation, Stability and Insulation Characteristics of Rosemount Standard Platinum Resistance Thermometers, Temperature, Its *Measurement and Control in Science and Industry*, AIP, New York, Vol.5, pp. 753-761.

Bogue, R. (2009). Inspired by Nature: Developments in Biomimetic Sensors, *Sensor Review*, Vol. 29, No.2, pp. 107-111, ISSN 0260-2288.

BSI (2005). *Specification for Data Quality Metrics of Industrial Measurement and Control Systems*, BS7986:2005 / British Standards Institute, 389 Chiswick High Rd, London W4 4AL.

Crovini, L.; Actis, A.; Coggiola, G. & Mangano, A. (1992). Precision Calibration of Industrial Platinum Resistance Thermometers, *Temperature: Its Measurement and Control in Science and Industry*, Vol. 6, edited by J. F. Schooley, New York: AIP, pp. 1077-1082.

Druzhinin, I.I. & Kochugurov, V.V. (1988) Check-up of Metrological Characteristuics of the Embedded Eddy-current Transducers, *Measurement Techniques*, Vol.31, No 11, pp. 1075-1091, 37-38, ISSN 0543-1972, ISSN 1573-8906.

Feng, Z.; Wang, Q. & Shida, K. (2007). A Review of Self-validating Sensor Technology, *Sensor Review*, Vol. 27, No.1, pp. 48-56, ISSN 0260-2288.

Feng, Z.; Wang, Q. & Shida, K. (2009). Design and Implementation of a Self-Validating Pressure Sensor, *IEEE Sensors Journal*, Vol. 5, No.3, pp. 207-218, ISSN 1530-437X.

Fridman, A.E. (1991). Theory of Metrological Reliability. *Measurement Techniques*, Vol. 34, No.11 1075-1091, ISSN 0543-1972, ISSN 1573-8906.

GOST R 8.673-2009. (2009). *State System for Ensuring the Uniformity of Measurements.* Intelligent *Sensors and Intelligent Measuring Systems. Basic Terms and Definitions.*

GOST R 8.565-96. (1996). *State System for Ensuring the Uniformity of Measurements. Metrological ensuring of atomic power stations exploitation. General principles.*

Hans, V. & Ricken O. (2007). Self-monitoring and Self-calibrating Gas Flow Meter, *Proceedings of the 8th International Symposium on Measurement Technology and Intelligent Instruments*, Sept 24-27, 2007, pp. 285-288.

Hashemian, H. M. & Petersen, K. M. (1992). Achievable Accuracy and Stability of Industrial RTDs, *Temperature: Its Measurement and Control in Science and Industry*, Vol. 6, New York: AIP, pp. 427-432, ISBN 1-55617-897-2, ISBN 1-55617-932-42.

Hashemian, H.M. (2005). *Sensor Performance and Reliability*, ISA, USA, ISBN-10 3-540-33703-2, ISBN-13 978-3-540-33703-4.

Hashemian, H.M. (2006). *Maintenance of Process Instrumentation in Nuclear Power Plants*. Berlin, Heidelberg, New-York: Springer.

Henry, M. P. & Clarke, D. W. (1993). The Self-validating Sensor: Rationale, Definitions and Examples. Control Engineering Practice, Vol.1., No. 4, pp. 585–610.

Henry, M.P.; Clarke, D.W.; Archer, N.; Bowles, J.; Leahy, M.J.; Liu, R. P. et al. (2000). A Self-validating Digital Coriolis Mass-flow Meter: an Overview, *Control Eng. Pract.*, Vol. 5, No.8 , pp. 487-506.

ISO/IEC 17025 (1999). *General Requirements for the Competence of Testing and Calibration Laboratories.*

Karzhavin, V.A. ; Karzhavin, A.V. & Belevtsev, A.V. (2007). About the Possibility to Apply Cable Nichrosil-nisil Thermoicouples as the Reference Ones, in: *Proc. of the 3rd All-Russian Conference "Temperature-2007"*, Obninsk, CD-ROM.

Lem, S. (1980). Summa Technologiae, Verlag Volk und Welt, Berlin.

Li, X.; Zhao, M. & Chen, D. (2010). A Study on the Stability of Standard Platinum Resistance Thermometer in the Temperature Range from 0 °C through 720 °C. http://www.hartscientific.com

Lukashev, A.P. ; Karlov, P.A. & Belyakov, A.E. (1984). SU1117472 (A1), *Pressure Pickup*, Priority Date: 1983-10-19, Pub. 1984-10-07

Mangum, B. W. (1984). Stability of small industrial PRTs, *Journal of Research of the NBS 89*, pp. 305-316.

McFarland, D. (1999). *Animal Behaviour. Psycology, Ethology, and Evolution*, Prentice Hall.

MI Recommendation 2021-89. (1989). *State System for Ensuring the Uniformity of Measurements. Metrological Assurance of Flexible Manufacturing Systems. Fundamentals*, Committee on Standardization and Metrology.

OIML D 10 (2007). *Guidelines for the Determination of Recalibration Intervals of Measuring Equipment Used in Testing Laboratories.*

Reed, R.P. (2003). Possibilities and Limitations of Self-validation of Thermoelectric Thermometry, *AIP Conference Proceedings, Temperature: Its Measurement and Control in Science and Industry*, Vol.7, p. 507, 2D. C. Ripple et al. eds., Melville, New York.

Red'ko, V.G. (2007). *Evolution. Neural Networks. Intelligence. Models and Concepts of the Evolutionary Cybernetics*, KomKniga, Moscow.

Sapozhnikova, K.V. Metrological Diagnostic Check, *Metrological Service in the USSR*, No.2, pp. 18-24, 1991.

Sapozhnikova, K.V.; Taimanov, R.Ye. & Kochugurov, V.V. (1988). Metrological Checking as a Component of Diagnostics of Flexible Production Systems and Robotics Complexes, *Testing, Checking and Diagnostics of Flexible Production Systems (from the materials of the seminar hold at the Blagonravov IMASH of the Academy of Science in 1985)*. – M.: Nauka, pp. 269-273.

Sapozhnikova, K.; Henry, M. & Taymanov, R. (2005a). The Need for Standards in Self-diagnosing and Self-validating Instrumentation, *Joint International IMEKO TC1+TC7 Symposium, September 21- 24, 2005*, Ilmenau, Germany (CD-ROM).

Sapozhnikova, K.; Taymanov, R. & Druzhinin, I. (2005b). About the Effective Approach to the Modernization of the NPP Control and Emergency Shutdown System, IAEA Technical Meeting on "Impact of the Modern Technology on Instrumentation and Control in Nuclear Power Plants" (621-12-TM-26932) 13-16 Sept. 2005, Chatou, France (CD-ROM).

Stroble, J.K.; Stone, R.B. & Watkins, S.E. (2009). An Overview of Biomimetic Sensor Technology, *Sensor Review*, Vol. 29, No.2 , pp. 112-119, ISSN 0260-2288.

Tarbeyev, Yu.; Kuzin, A.; Taymanov, R. & Lukashev, A. (2007) New Stage in the Metrological Provision for Sensors, *Measurement Techniques*, Vol. 50, No.3 , pp. 344-349.

Taymanov, R.; Sapozhnikova, K. & Druzhinin, I. (2007). Measuring Control Rod Position, *Nuclear Plant Journal*, 2007, No.2, pp. 45-47, ISSN 0892-2055.

Taymanov, R. & Sapozhnikova, K. (2009). Problems of Terminology in the Field of Measuring Instruments with Elements of Artificial Intelligence, *Sensors & Transducers journal*, Vol.102, 3, pp. 51-61, ISSN 1726-5749.

Taymanov, R. & Sapozhnikova, K. (2010a). Metrological Self-Check as an Efficient Tool of Condition Monitoring, *The Seventh International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Stratford-upon-Avon, England.*

Taymanov, R. & Sapozhnikova, K. (2010b). Metrological Self-Check and Evolution of Metrology, *Measurement*, Vol.43, No.7, pp. 869-877, ISSN 0263-2241.

Taymanov, R.; Sapozhnikova, K. & Druzhinin, I. (2011). Sensor Devices with Metrological Self-Check, *Sensors & Transducers journal*, Vol.10 (special issue), No.2, (February 2011), pp. 30-44, ISSN 1726-5749.

Turchin, V.F. (1977). *The Phenomenon of Science. A Cybernetic Approach to Human Evolution*, Columbia University Press, New York.

VIM. International Vocabulary of Metrology — Basic and General Concepts and Associated Terms, JCGM, 2008.

VDI/VDE Guideline 2650 (2005). *Requirements for Self-monitoring and Diagnostics in Field Instrumentation.*

Werthschutzky, R. & Muller, R. (2007). Sensor Self-Monitoring and Fault-Tolerance, *Technisches Messen*, Vol. 74, No.4, pp. 176-184.

Werthschützky, R. & Werner, R. (2009). Sensor Self-Monitoring and Fault-Tolerance, *Proceedings of the ISMTII'2009*, 29 June – 2 July, 2009, St.Petersburg, Russia, pp.4-061- 4-065.

Wiener, N. (1948). *Cybernetics: Or the Control and Communication in the Animal and the Machine*, MA, MIT Press, Cambridge.

2

Multi-Version FPGA-Based Nuclear Power Plant I&C Systems: Evolution of Safety Ensuring

Vyacheslav Kharchenko[1], Olexandr Siora[2] and Volodymyr Sklyar[2]
[1]*National Aerospase University KhAI,*
Centre for Safety Infrastructure-Oriented Research and Analysis,
[2]*Research and Production Corporation RADIY,*
Ukraine

1. Introduction

1.1 Problem of decreasing common cause failure probability for nuclear power plant instrumentation and control systems

To guarantee required level of dependability, safety and security of computer-based systems for critical (safety-critical, mission-critical and business-critical) applications it is used diversity approach. This approach implies development, choice and implementation of a few diverse design options of redundant channels for created system. Probability of common cause failure (CCF) of safety-critical systems may be essentially decreased due to selection and deployment of different diversity types on the assumption of maximal independence of redundant channels realizing software-hardware versions.

This circumstance calls forth that a lot of international and national standards and guides contain the requirements to use diversity in safety-critical systems, first of all, in nuclear power plant (NPP) instrumentation and control systems (I&Cs) (reactor trip systems), aerospace on-board equipment (automatic/robot pilot, flight control systems), railway automatics (signalling and blocking systems), service oriented architecture (SOA)-based web-systems (e-science) etc. (Pullum, 2001; Wood et al., 2009; Gorbenko et al., 2009; Kharchenko et al., 2010; Sommerville, 2011).

Application of the modern information and electronic technologies and component-based approaches to development in critical areas, on the one hand, improve reliability, availability, maintainability and safety characteristics of digital I&Cs. On the other hand, these technologies cause additional risks or so-called safety deficits. Microprocessor (software)-based systems are typical example in that sense. Advantages of this technology are well-known, however a program realization may increase CCF probability of complex software-based I&Cs. Software faults and design faults as a whole are the most probable reason of CCFs. These faults are replicated in redundant channels and cause a fatal failure of computer-based systems. It allows to conclude that, "fault-tolerant" system with identical channels may be "non-tolerant" or "not enough tolerant" to design faults. For example, software design faults caused more than 80% failures of computer-based rocket-space systems which were fatal in 1990 years (Kharchenko et al., 2003) and caused 13% emergencies of space systems and 22% emergencies of carrier rockets (Tarasyuk et al., 2011).

The CCF risks may be essential for diversity-oriented or so-called multi-version systems (MVSs) (Kharchenko, 1999) as well if choice of version redundancy type and development

of channel versions are fulfilled without thorough analysis of their independence and assessment of real diversity degree assessed by special metrics, for example, β-factor (Bukowsky&Goble, 1994).

1.2 Complex electronic components and FPGA technology for NPP I&Cs development

An analysis of development and introduction trends of computer technologies to NPP I&Cs has specified a number of important aspects affecting their safety, peculiarities of development, update and licensing. Such trends include, among others (Yastrebenetsky, 2004): introduction of novel complex electronic components (CECs); expanded nomenclature of software applied and increased effect of its quality to I&Cs safety; realization of novel principles and technologies in I&Cs development; advent of a large number of novel standards regulating the processes of I&Cs development and safety assessment. During recent decades the application of microprocessor techniques in NPP I&Cs design has substantially expanded. Microprocessors are used both in system computer core and in realization of intellectual peripherals – various sensors, drives and other devices with built-in programmable controllers.

Another contemporary trend is dynamically growing application of programmable logic technologies, particularly, Field Programmable Gate Arrays (FPGA) in NPP I&Cs, onboard aerospace systems and other critical areas. FPGA as a kind of CECs is a convenient mean not only in realization of auxiliary functions of transformation and logical processing of information, but also in execution of basic monitoring and control functions inherent in NPP I&Cs. This approach in some cases is more reasonable than application of software-controlled microprocessors (Kharchenko&Sklyar, 2008). In assessment of FPGA-based I&Cs it should be taken into consideration that application of this technologies somewhat levels the difference between hardware and software, whereas obtained solutions are an example of a peculiar realization of so called heterosystems – systems with "fuzzy" software-hardware architecture and mixed execution of functions. This circumstance and other features of FPGA technology increase a number of diversity types and enlarge a set of possible diversity-oriented decisions for NPP I&Cs.

1.3 Work related analysis

Known works, related to the current problem and taking into account features of NPP I&C systems, are divided into three groups: (1) classification and analysis of version redundancy types and diversity-oriented decisions; (2) methods and techniques of diversity level assessment and evaluation of multi-version systems safety in context of CCFs; (3) multi-version technologies of safety critical systems development.

1. A set of diversity classification schemes (general, software and FPGA-based) was analyzed in (Kharchenko et al., 2009). First one is based on NUREG technical reports and guides, samples two-level hierarchy and includes seven main groups of version redundancy (Wood et al., 2009): signal diversity (different sensed reactor or process parameters, different physical effects, different set of sensors); equipment manufacture diversity (different manufacturers, different versions of design, different CEC versions, etc); functional diversity (different underlying mechanisms, logics, actuation means, etc); logic processing equipment or architecture diversity (different processing architectures, different component integration architectures, different communication architectures, etc); logic or software diversity (different algorithms, operating system, computer languages,

etc); design diversity (different technologies, approaches, etc); human or life cycle diversity (different design organizations/companies, management teams, designers, programmers, testers and other personnel). Software diversity types are classified in according with following attributes (Pullum, 2001; Volkoviy et al., 2008): life cycle models and processes of development (for example, V-model for main version and waterfall model with minimum set of processes for duplicate version); resources and means (different human resources, languages and notations, tools); project decisions (different architectures and platforms, protocols, data formats, etc). Next one FPGA-based classification includes the following types of diversity (Kharchenko&Sklyar, 2008; Siora et al., 2009): diversity of electronic elements (different electronic elements manufactures, technologies of production, electronic elements families, etc); diversity of CASE-tools (different developers, kinds and configurations of CASE-tools); diversity of projects development languages (different graphical scheme languages, hardware description languages and IP-cores); diversity of specifications (specification languages) and others.

2. There are following methods of diversity level assessment and evaluation of MVS dependability and safety (Kharchenko et al, 2009). Theoretical-set and metric-oriented methods are based on: Eiler's diagram for sets of version design, physical and interaction faults (including vulnerabilities for assessment intrusion-tolerance); matrix of diversity metrics for sets of different faults (individual, group and absolute faults of versions); calculation of diversity metrics by use of Eiler's diagrams or other data about results of testing and faults of different versions. Probabilistic methods use reliability block-diagrams (RBDs), their modifications (survivability and safety block-diagrams), Markovian chains, Bayesian method, etc. Statistical methods include the following procedures: receiving and normalization of version fault trends using testing data; choice of software reliability growth model (SRGM) taking into account features of version development and verification processes and fitting SRGM parameters; metrics diversity assessment; calculation of reliability and safety indicators. Fault injection-based assessment consists of: receiving project-oriented fault profiles; performing of faults injection procedure; proceeding of data and metrics diversity calculation; calculation of reliability and safety indicators. Expert-oriented methods use two groups of metrics: diversity metrics for direct assessment of versions and MVS reliability and safety (direct diversity metrics); indirect diversity metrics (product complexity metrics and process metrics); values of these metrics may be used to assess direct diversity metrics. Expert methods are added other techniques founded on interval mathematics-based assessment of diversity metrics and MVS indicators, soft computing-based assessment (fussy logic, genetic algorithms), risk-oriented approach and so on.
3. Multi-version technologies (MVTs) of diversity types selection and application, development of MVSs as a whole are based on (Siora et al., 2009; Wood et al., 2009) use of diversity types and strategies table, a model of multi-version life cycle (MVLC), a special graph of diversity types and their modifications, and procedures of diversity type and volume choice according with different criteria. The set of diversity strategies developed in the (Wood et al., 2009) consists of three families of strategies: different technologies—Strategy A (digital vs analog), different approaches within the same technology—Strategy B (microprocessor vs FPGA) and different architectures within the same technology—Strategy C (IP-based vs VHDL). Each of the strategy families is characterized by combinations of diversity criteria that may provide adequate mitigation of potential CCF vulnerabilities according with metrics determined by expert way.

There are a lot of examples of multi-version systems and multi-version technologies application in different safety critical areas. Generalized results of MVS application analysis are presented by matrix "types of diversity – areas of multi-version I&Cs application" in Table 1 (Wood et al., 2009; Kharchenko et al., 2010).

Diversity types	Multi-version I&C systems application												
	Space		Aviation				Railways	Chemic. industry	Defense	Power Plants	NPPs		e-Commers
	Shuttle	ISS	MC JVC	FAA FCS	Air-bus A320	Boeng 777	SCB	CCPS	MICS	Electr. Grid	RTS	ESFAS	WSOA
Design													
Equip-ment													
Func-tion													
Human													
Signal													
Soft-ware													
Others													

Table 1. Matrix "types of diversity – areas of multi-version I&Cs application"

Types of diversity (diversity redundancy) are classified according to NUREG 6303 and painted by different colors. Last row of the matrix corresponds to other types of diversity. MVSs are used in space systems (Shuttle, ISS), aviation equipment (MC JVC, FAA FCS, Airbus and Boeing on-board systems), railway automatics (signaling, centralization and blocking systems SCB), chemical industry (CCPS), defense systems, power plants (electricity grid), NPPs (RTS and ESFAS), e-commerce and e-science (web-systems with diverse target web-services).

1.4 Goal and structure of the chapter

In spite of the intensive researches in area of multi-version systems and long-term experience of their application there are some problems of diversity approach implementation in context of FPGA technology application in NPP I&Cs, videlicet: specifying of concepts used; selection of diversity types and required volume of version redundancy; joint use of different diversity types taking into consideration state-of-the-art technologies; assessment of real diversity degree and effectiveness of MVSs, etc. Goal of the chapter is analysis of concepts in multi-version computing and diversity-scalable decisions for FPGA-based NPP I&Cs. Structure of the chapter is following. The section 2 elaborates the FPGA peculiarities in context of safety critical applications and evolution aspect of

FPGA-technology and diversity approach conformably to NPP I&Cs. The standards containing requirements to application of diversity approach in NPP I&Cs and key challenges in this area are analyzed in the section 3. The taxonomy of multi-version computing and models of MVSs and MVTs are represented in the section 4. General approach to assessment of diversity and MVS safety is described in the section 5. Features of FPGA-based platform RADIY™ and results of implementation of multi-version I&Cs in NPPs are analyzed in the section 6. Finally, the section 7 concludes the chapter and presents directions of future researches.

2. An evolution of FPGA technology and diversity application in NPP I&Cs

2.1 FPGA peculiarities in context of dependability and safety

FPGA architecture topologically originates from channeled Gates Arrays (GA) (Altera, 2001). In FPGA internal area a set of configurable logic units is disposed in a regular order with routing channels there between and I/O units at the periphery. Transistor couples, logic gates NAND, NOR (Simple Logic Cell), multiplexer-based logic modules, logic modules based on programmable Look-Up Tables (LUT) are used as configurable logic blocks. All those have segmented architecture of internal connections.

System-On-Chip architecture appeared due to two factors: high level of integration permitting to arrange a very complicated circuit on a single crystal, and introduction of specialized hardcores into FPGA. Additional hardcores may be: additional Random Access Memory (RAM) units; JTAG interface for testing and configurating; Phase-Locked Loop (PLL) – frequency control system to correct timing relations of clock pulses as well as for generation of additional frequencies; processor cores enabling creation of devices with a control processor and a peripheral.

Analysis of dependability assurance possibilities in FPGA-based systems allows to determine the following FPGA peculiarities (Kharchenko&Sklyar, 2008; Bobrek et al., 2009).

1. Simplification of development and verification processes: apparatus parallelism in control algorithms execution and realization of different functions by different FPGA elements; absence of cyclical structures in FPGA projects; identity of FPGA project presentation to initial data; advanced testbeds and tools; verified libraries and Intellectual Properties (IP)- cores in FPGA development tools.
2. There are three technologies of FPGA-projects development: development of graphical scheme with using of library blocks in CAD environment; development of software model with using of especial hardware describing languages (VHDL, Verilog, Java HDL, etc); development of program code for operation in environment of microprocessor emulators which are implemented in FPGA as IP-cores. It does allow increasing a number of options of different project versions and multi-version I&Cs.
3. Assurance of fault-tolerance, data validation and maintainability due to use of: redundancy for intra- and inter-crystal levels; diversity implementation; reconfiguration and recovery in the case of component failures; improved means of diagnostic.
4. Security assurance: FPGA reprogramming is possible only with use of especial equipment. Stability and survivability assurance due to: tolerance to external impacts (electromagnetic, climatic, radiation); possibilities of implementation of multi-step degradation with different types of adaptation.

2.2 FPGA technology application in safety-critical systems and NPP I&Cs

Due to these peculiarities area of FPGA technology application essentially has expanded. We can say about a affirmative answer to question "Expansion of FPGA-technology application in safety-critical systems for the last decades: evolution or revolution?" It is confirmed by (Bakhmach et al., 2009):

substantial increase of applying the technologies based on programmable logic (FPGA, CPLD, ASIC);

FPGA technology is improved and ensures new possibilities to develop more reliable and effective systems; application FPGA technology for development of military (B-1B, F-16, etc) and civil aircraft control systems (Boeing 737, 777, AN70, 140), space control systems (satellites FedSat, WIRE; the Mars-vehicle Spirit), etc;

application of FPGAs in NPP I&Cs (Ukraine, Russia, Bulgaria: 1999-start, 2002 – 1000, 2006 – 6000, 2008-2010 – more than 8000 chips every year).

Besides, the illustration of FPGA expansion is evolution of the NPP I&Cs produced by RPC Radiy during 2000-2008 years (Kharchenko&Sklyar, 2008).

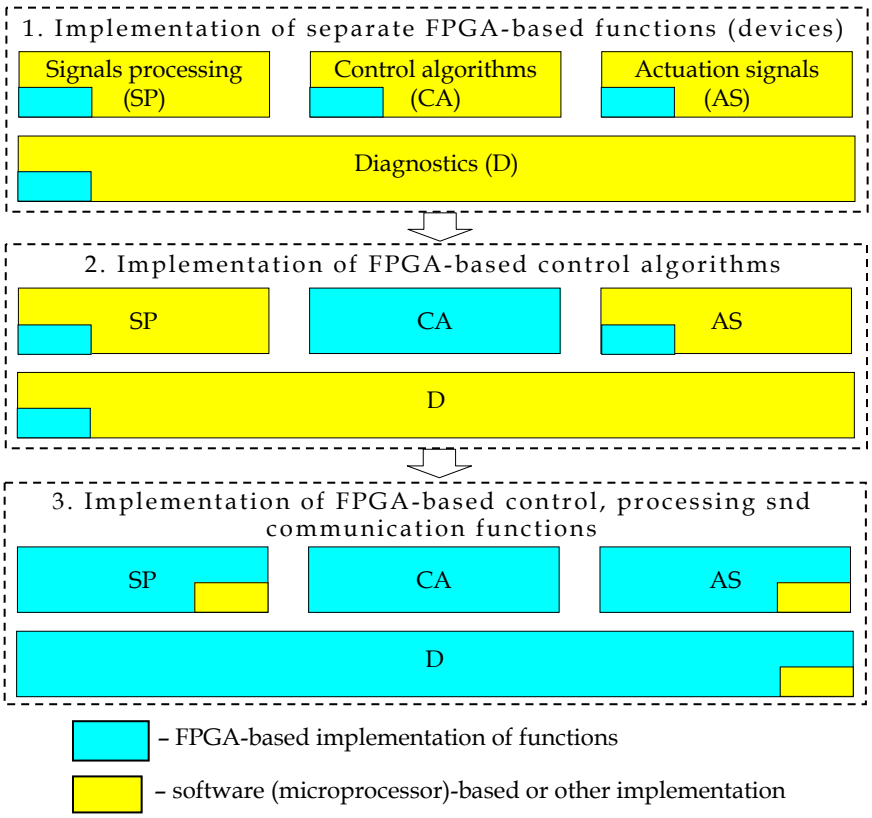


Fig. 1. Application of FPGA technology in the NPP I&Cs produced by RPC Radiy

There are three stages of the evolution (Fig.1): from implementation of separate FPGA-based functions in I&Cs (signals processing (SP), control algorithms (CA), actuation signals formation (AS) and diagnostics (D)), stage 1, and implementation of FPGA-based CA, stage 2, to preferred implementation of FPGA-based SP-, CA-, AS-, D- and communication functions, stage 3.
Analysis of industrial application experience of FPGAs in NPP I&Cs is described in technical report prepared by EPRI (Naser, 2009).

2.3 A law "negation of negation": Stages of diversity approach implementation evolution in NPP I&Cs

Interesting are the results of transformation of multi-version I&Cs for the last decades in context of hardware-software-FPGA technologies development. There are a few diversity implementation evolution stages in safety-critical NPP I&Cs, in particular, reactor trip systems. Analysis of these stages allows formulating (or demonstrating truth) a law "negation of negation" (Kharchenko et al., 2009) (Fig.2):

- stage 1 (1970-1980s) – use of hardware (hard logic, HL)-based one-version systems and transition from hardware (HW)-based systems with identical subsystems to systems with hardware (HL)-based primary subsystem and software (microprocessor, MP)-based secondary subsystem; it was the first "negation";
- stage 2 (1990s) – use of primary and secondary subsystems with software (SW) diversity (I&C platforms produced by Siemens, WH and other companies); example of multi-version systems with software diversity is two-version system consisting of subsystems developed using microprocessors Intel and Motorola (languages C and Ada); it completed the first cycle of "negation of negation";
- stage 3 (2000s, first half) – transition to FPGA-based primary and software-based secondary subsystems with equipment, design and software diversity (first generation of the I&C platforms produced by RPC Radiy); it was next "negation";
- stage 4 (2000s, second half) – application of FPGA-oriented soft processors for primary subsystem and FPGA project developed using HDL-oriented language (hard logic) for creation of secondary subsystem (next generation of the I&C platform produced by RPC Radiy); it completed the second cycle of "negation of negation";
- stage 5 (beginning of 2010s) – application of different FPGAs (hard logic) produced by different manufacturers (and other types of diversity) for primary and secondary subsystems correspondingly; it is next "negation".

What will be the next step? Probably, advancement of electronic technologies, in particular, nanotechnologies, naturally dependable, safe and secure chips will create new perspectives and possibilities for development of diversity-oriented decisions. Actel, Altera and others companies inform about creating first chips called nano FPGAs allowing to develop fault-tolerant projects using large-scale means.

3. Normative base and key challenges connected with diversity application in NPP I&Cs

3.1 Analysis of diversity related standards

There are the following standards and guides contained requirements to diversity:

- IEC 61513: 2001. NPPs - I&Cs important to safety – general requirements for systems;
- IEC 60880: 2006. NPPs - I&Cs important to safety - SW aspects for computer-based systems performing category A functions;

- IAEA NS-G-1.3: 2002. I&Cs important to safety in NPPs;
- IEEE std.7-4.3.2:1993. IEEE standard criteria for digital computers in safety systems of NPPs;

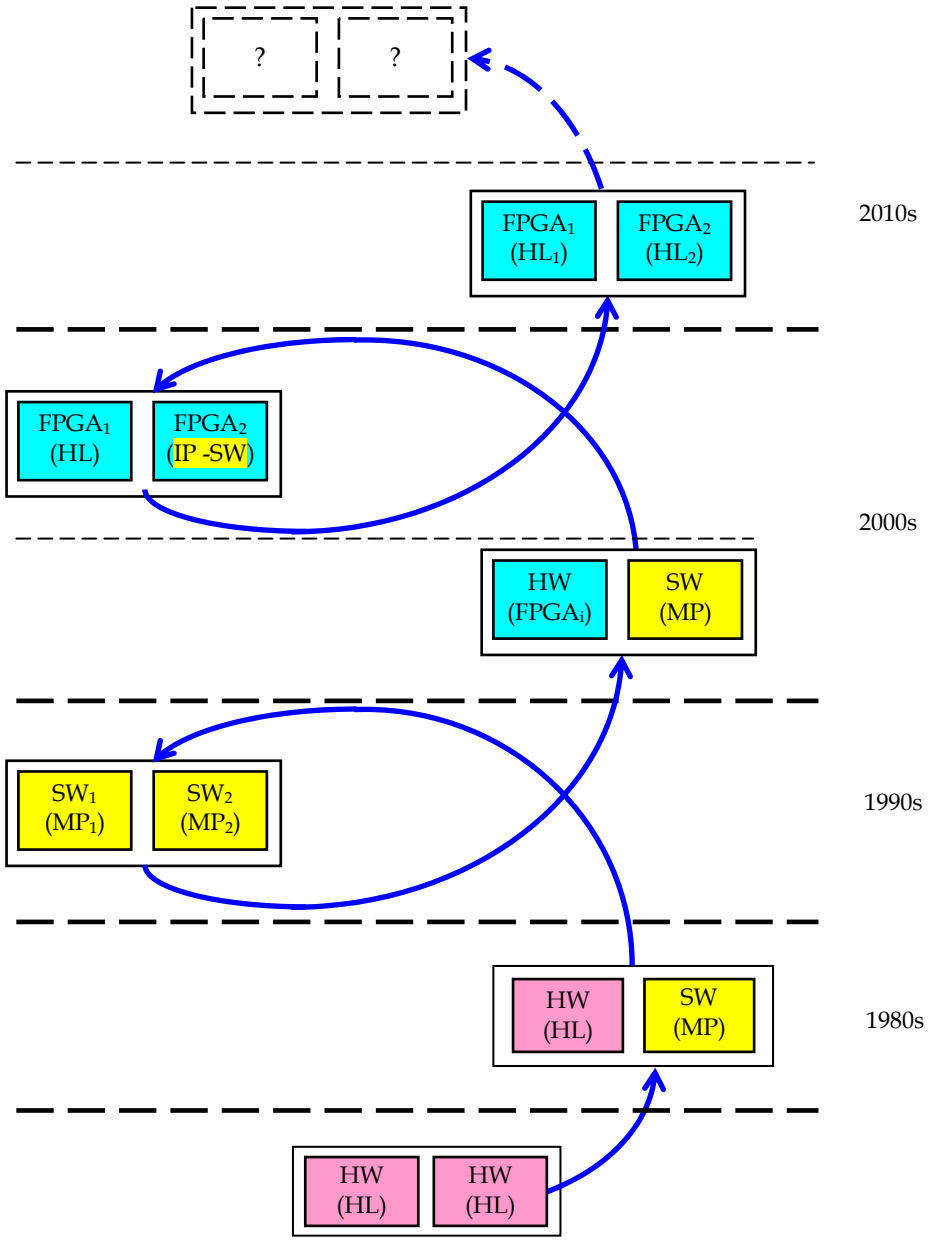


Fig. 2. Stages of diversity approach implementation evolution in safety-critical NPP I&Cs

- NUREG/CR-6303:1993. Method for Performing Diversity and Defense-in-Depth Analyses of Reactor Protection Systems;
- DI&C-ISG-02, Diversity and Defense-in-Depth Issues, Interim Staff Guidance, BTP 7-19, Guidance for Evaluation of D&DiD In Digital I&C Systems (USA);
- NP 306.5.02/3.035: 2000. Requirement on nuclear and radiation safety to I&Cs important to safety in NPPs (Ukraine), etc.

These standards contain general requirements concerning: systems which must/should be developed using diversity approach (Reactor Trip Systems); types of diversity used to develop NPP I&Cs and to decrease CCF probability; features of diversity implementation, determination of types and volume of diversity; assessment (justification) of real level of diversity in developed systems; drawbacks and benefits connected with the use of diversity.

The standards are not enough detailed to make all necessary decisions concerning diversity. It's important to develop additional detailed techniques of assessing diversity and choosing optimal kinds and volume of diversity according to criterion "safety-reliability-cost".

3.2 Key challenges

Main conclusions concerning FPGA-based MVS development and implementation experience are the following:

FPGA-based multi-version I&Cs are used in NPPs during 6-8 last years, i.e. these systems are new object of analysis and still more unique one;

FPGA technology gives additional possibilities to develop MVSs and ensure high safety and reliability;

processes of FPGA project development are similar to processes of SW-based project development. FPGA project product is similar to HW-based project product (hard logic);

there are not any international standards determined requirements to use of diversity for I&Cs development and application taking into account FPGA features.

Results of comparative analysis of challenges caused by development and application of software- and FPGA-based multi-version systems are presented in Table 2.

4. Main concepts and models of multi-version computing

4.1 Taxonomy scheme of multi-version computing

A set of concepts concerning diversity may be united by general term "multi-version computing" on the analogy with "dependable computing" (Avižienis et al., 2004). Multi-version computing is a type of dependable computing organization based on use of diversity approach. Taxonomy scheme of multi-version computing developed taking into consideration concepts in this area described in international standards includes the following elements (Kharchenko et al, 2009) (Fig.3).

Version is an option of the different realization of identical task (by use software, hardware or FPGA-based products and life cycle processes); identical versions of structure redundancy-based system are trivial. Version redundancy (VR) is a type of product and process redundancy allowing to create different (non-trivial) versions; product VR is realized jointly with structure, time and other types of non-version redundancy.

Challenges	Software-based multi-version I&C	FPGA-based multi-version I&C
Detailed standards	There are standards determining general requirements to use of diversity	There are no special standards
Experience of development and operation	More 20 years	6-8 years
Trustworthiness of diversity assessment	Methods of expert-based, metrical assessment, probabilistic methods using SRGMs	Methods of expert-based, metrical, probabilistic (RBD), deterministic methods
Development of MVSs	Choice of diversity kinds, generation of really diverse software versions	Number of diversity kinds increases
Verification of MVSs	Verification activities volume are significantly increased	Verification is more simple due to simplifying of version verification

Table 3. Key challenges for software-based and FPGA-based MVSs

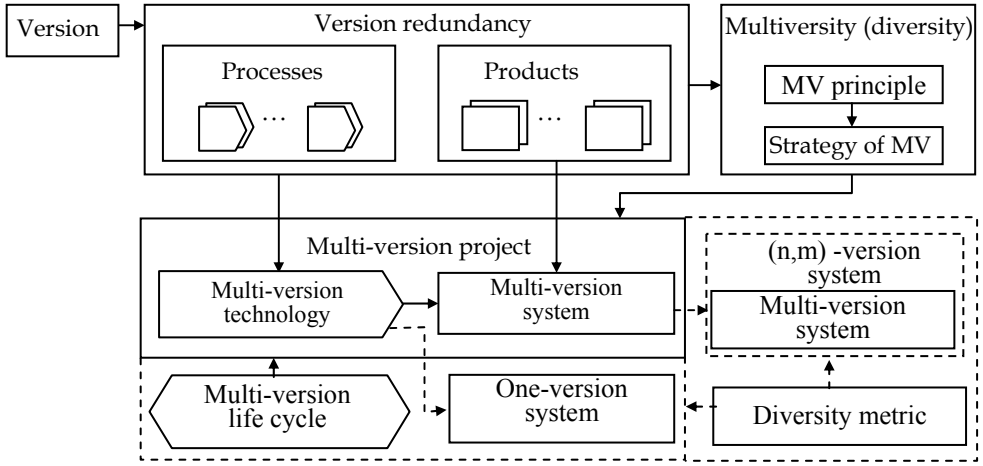


Fig. 3. Taxonomy scheme of multi-version computing

Diversity or multiversity (MV) is a principle providing use of several non-trivial versions; this principle means performance of the same function (realization of products or processes) by two and more options and processing of data received in such ways for checking, choice or formations of final or intermediate results and decision-making on their further use.

Multi-version system (MVS) is a system in which a few versions-products are used; one-version systems may be redundant but consists of a few trivial versions. Multi-diversion system (MDVS) is MVS in which two or more VR types are applied. Multi-version

technology (MVT) is set of the interconnected rules and design actions in which in accordance with MV strategy a few versions-processes leading to development of two or more intermediate or end-products are used; thus for development of MVS should be used MVT, for development one-version systems can be used both multi-version and one-version technology.
Multi-version project (MVP) is a project in which the multi-version technology is applied (version redundancy of processes is used) leading to creation of one- or multi-version system (realization of version redundancy of products). Strategy of diversity (MV) is a collection of general criteria and rules defining principles of formation and selection of version redundancy types and volume or/and choice of MVTs. Besides, important elements of multi-version computing are concepts "multi-version life cycle", "diversity metric". More detailed interpretation of these concepts will be done below.

4.2 Diversity type classification schemes

Different variants of diversity type classifications were described above. The analysis of the considered classifications allows approving that:

- they are presented by classifications of mixed facet-hierarchical or matrix (network) types;
- the NUREG-based classification presented in (Wood et al., 2009) is the most detailed and systematic, though the principle of attributes orthogonality is not sustained in full in it; for example, subsets of design and software, functional and signal version redundancy are crossed and dependent;
- variety of product (system, hardware and software components) and of process (technologies of development, testing and maintenance) version redundancy cause complexity of VR selection and MVS development.

More general diversity type classification scheme is so-called "cube" of diversity described by matrix MVR = $||\, vr_{ijk} \,||$ in three-dimensional space (Fig. 4). The scheme has coordinates: stage of LC (i); level of project decisions (PD, j) and type of VR (project decision).
Example of two-space matrix presented a cut of "cube" for FPGA-based systems is shown on the table 3. This table contains variants of joint application of one or two diversity types (items 1.4.2-1.4.4, 2.3.3-2.3.8, 3.3.3-3.3.8, 4.2.4-4.2.15; for example, last combinations correspond to 12 = 4 (kinds of EE diversity) x 3 (kinds of CASE-tool diversity) couples).

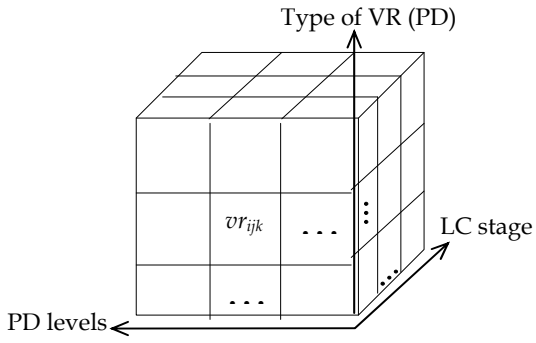


Fig. 4. Cube" of diversity-oriented decisions

Stages of FPGA-based I&C life cycle	Kinds of version redundancy			
	1 Diversity of electronic elements (EE)	2 Diversity of CASE-tools	3 Diversity of project development languages	4 Diversity of scheme specification (SS)
1 Development of block-diagrams according with signal formation algorithms		1.2.1 Different developers of CASE-tools 1.2.2 Different CASE-tools kinds 1.2.3 Different CASE-tools configurations		1.4.1 Different SSs 1.4.2-1.4.4 Combination of couples of diverse CASE-tools and SSs
2 Development of program models of signal formation algorithms in CASE-tools environment		2.2.1 Different developers of CASE-tools 2.2.2 Different CASE-tools kinds 2.2.3 Different CASE-tools configurations	2.3.1 Joint use of graphical scheme language and HDL 2.3.2 Different HDLs 2.3.3-2.3.8 Combination of diverse CASE-tools and HDLs	
3 Integration of program models of signal formation algorithms in CASE-tools environment		3.2.1 Different developers of CASE-tools 3.2.2 Different CASE-tools kinds 3.2.3 Different CASE-tools configurations	3.3.1 Joint use of graphical schemes and HDL 3.3.2 Different HDLs 3.3.3 – 3.3.8 Combination of couples of diverse CASE-tools and HDLs	
4 Implementation of integrated program model in FPGA	4.1 Different manufacturers of EEs 4.2 Different technologies of EEs production 4.3 Different families of EEs 4.4 Different EEs of family	4.2.1 Different developers of CASE-tools 4.2.2 Different CASE-tools kinds 4.2.3 Different CASE-tools configurations 4.2.4-4.2.15 Combination of diverse CASE-tools and EEs		

Table 2. Matrix of diversity-oriented FPGA-based decisions

4.3 Models multi-version systems

One-version W(1) and multi-version W(n) systems are defined by 4 and 6 variables (Kharchenko et al., 2010):

$$W(1) = \{X, Y, Z. \Phi\}, \quad (1)$$

$$W(n) = \{X, Y, Z. \Phi, V, \Psi\}, \quad (2)$$

where X, Y, Z – sets of input signals, internal conditions (states) and output signals correspondingly; $\Phi = \{\varphi_i, i=1, ..., a\}$ – a set of I&C functions (for examples, actuation functions or algorithms of reactor trip system); $V = \{v_j, j=1, ..., n\}$ – a set of versions with output signals $Z_1,\ldots, Z_n$ (or signals Z_{id} , $d = 1,\ldots, n_i$; n_i is a number of versions for function φ_i; $\forall\varphi_i \sim v_j = \{ v_{ij}, j =1,...,n_i\}$); $\Psi = \{\psi_s, s=1, ..., в\}$ – mapping $Z_i \rightarrow Z$.

If the function φ_i is performed, local mapping is true: $\psi_s:\{z_i(v_{i1}),..., z_i(v_{in_i})\}\rightarrow Z_i^{(S)}$. Taking into account formulas (1) and (2), multi-version system and one-version system are connected by relationship:

$$W(n) = \{W(1), V, \Psi\}. \quad (3)$$

System W(1) may be structure-redundant and contain usual means Ψ for signals processing from identical channels (versions). In this case card V=1. For system W(n) is true that: $\forall_j=\overline{1,a} : \exists_j: n_i>1$.

Mapping ψ_s is generally described by: a subset of versions $\Delta v_s \subset v_j$ for receiving output signal Z_i; a vector $\vec{t}_s$ of version v_{ij} initialization time ($\vec{t}_s = \{t(v_{i1}),..., (v_{in_i})\}$); a mean of transforming η_s values $z_i(v_{i1}),..., z_i(v_{in_i})$ in output signal Z_i^S. Hence,

$$\forall\psi_s\in\Psi:\psi_s = \{\Delta vs, \vec{t}_s, \eta_s\} \text{ and } Z_i^{(S)} = \eta_s [z_i(v_{ij}), \vec{t}_s], v_{ij}\in\Delta v_s.$$

There are the following means of transforming η_s: (a) the conjunctive, when $Z_i^S = Vz_i(v_{ij})$; (b) the time conjunctive, when $Z_i^S = Vz_i(v_{ij})\sigma_{ij}$, where $\sigma_{ij}=1$, if $t=t(v_{ij})$, and if not $\sigma_{ij}=0$; (c) the majority, when $Z_i^S = M[z_i(v_{ij})]$, where M is a majority function k out of *l* (or k out of n); (d) the majority-weighted, when weights of versions $\omega(v_{ij})$ are additionally defined on majorization; (e) the functional, when $Z_i^S = f[z_i(v_{ij})]$, where f - some function of transforming output signals of every version.

The model (2) describes system with n versions that, $n = \sum_{i=1}^{a} n_i$. This model does not take into account the possibility of applying several diversity kinds. A set of version redundancy kinds $R=\{r_d, d=1,..., m\}$ may be decomposed on subsets for versions of products $v_{prd}(t_j)$ and processes $v_{prc}(t_j)$: $R=(\bigcup_j \Delta R_{prdj}) \cup (\bigcup_j \Delta R_{prcj})$, where ΔR_{prdj} and ΔR_{prcj} – appropriate subsets.

Thus, different diversity kinds, $r\in R$, are accumulated in final versions of a multi-version system. It is described by special mapping $\Theta : R \rightarrow V$. Mapping Θ may be presented by Boolean matrix $\|\theta d_j\|$, $d=\overline{1,m}; j=\overline{1,n}$, where $\theta_{dj}=1$, if diversity kind r_p is used in version v_j, and if not $\theta_{dj} = 0$. Then multi-version system W(n,m) or multi-diversion system is described by formula:

$$W(n,m) = \{ X, Y, Z, \Phi, V, \Psi, R, \Theta \} = \{W(n), R,\Theta \} = \{W(1), V, \Psi,R, \Theta \}. \quad (4)$$

It is important to describe correspondence between a set of versions V and a set of redundant channels $C=\{c_q, q=1,...,l\}$. This correspondence may be defined by mapping $Q:V\rightarrow C$. This mapping is presented by Boolean matrix $Q = \|\omega_{jg}\|$, $d=\overline{1,m}$, $g=\overline{1,l}$, where $\omega_{gj} = 1$, if version v_i is realized by channel c_j, and if not $\omega_{gj} = 0$. Then model of multi-version (multi-diversion) system is the following:

$$W(n,m,l) = \{ X, Y, Z, \Phi, V, \Psi, R, \Theta, C, Q \}= \{ W(n,m), C, Q \}. \quad (5)$$

MVSs with temporal redundancy and p iterations of algorithms are indicated as W(n,m,n,p) dividing number of parallel (structural) versions n_c, and sequential versions realized by using one channel. Set X may be decomposed for different versions if

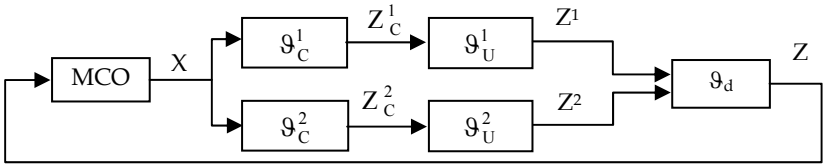
$$X = \bigcup_j X_j \,,\ \forall j_1 j_2 \in \overline{1,n} \,,\ j_1 \neq j_2 \colon\ X_{j1} \cap X_{j2} \,,\ X_{j1} \cap X_{j2} = \phi.$$

Such MVSs are called multi-version systems with naturally divided input alphabet:

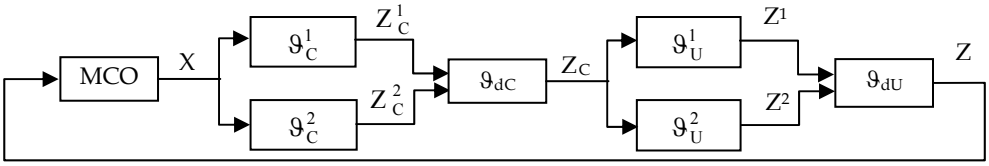
$$W_{NX} = \{ \{X_j\}, Y, Z, \Phi, V, \Psi, R, \Theta, C, Q\}. \tag{6}$$

If versions process data presented in different notations, such MVSs are called multi-version systems with artificially divided input alphabet WAX. A special function-transformer Πx (Πxj) should be specified in addition to alphabet X :

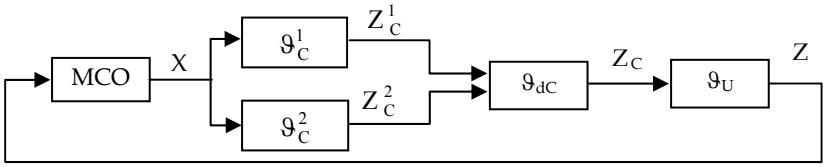
$$W_{NX} = \{X, \{\Pi xj\}, Y, Z, \Phi, V, \Psi, R, \Theta, C, Q\}. \tag{7}$$



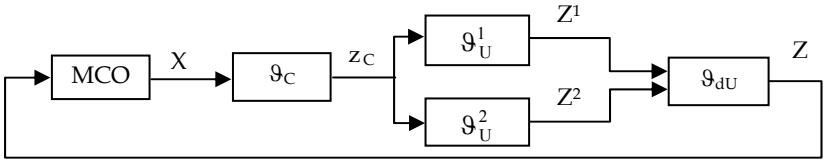
a) two-versions system with full common diversity, ϑ_{FO}



b) two-versions system with full separate diversity, ϑ_{FS}



c) two-versions system with partial diversity (for ϑ_C), ϑ_{PC}



d) two-versions system with partial diversity (for ϑ_U), ϑ_{PU}

ϑ_C^1 , ϑ_C^2 - the first and the second versions of a monitoring automaton;

ϑ_U^1 , ϑ_U^2 - the first and the second versions of a contril automaton;

ϑ_{dC}, ϑ_{dU}, ϑ_d - solver for union of two versions results.

Fig. 5. Architecture variants of two-version I&C systems

Besides, I&Cs performing safety-critical functions may be represented by a composition of two interconnected subsystems – monitoring (checking) subsystem and control subsystem (monitoring and control automata). Monitoring automaton ϑ_C analyses output signals X from monitoring and control object (MCO) and forms its status code Z_C.
Control automaton ϑ_U forms control signals Z in accordance with signals Z_C. Several options of MVS architectures are possible for a FPGA-based I&Cs. Those options may be classified according with such attributes (see Fig. 5):
degree of diversity coverage (I&Cs with a full ϑ_F and partial ϑ_P diversity);
diversity depth (I&Cs with a common ϑ_O and separate ϑ_S diversity); it should be noted that this feature is applicable only to full system diversity.

4.4 Models of multi-version life cycle and technology

A model of MVS life cycle (or multi-version LC model) is based on operations of version generation G, aggregation and selection U at various stages (Kharchenko et al., 2007). Example of the two-version life cycle model is shown on Fig. 6 taking to account some FPGA-oriented design features (V_{ij} are different versions obtained on different stage of development) (Prokhorova et al., 2008).

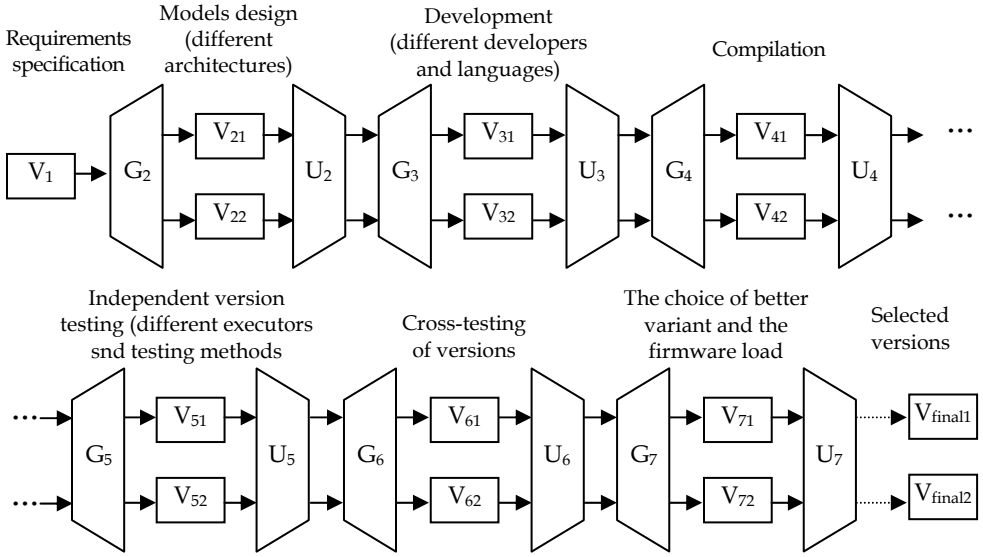


Fig. 6. FPGA-system multi-version life cycle

In general case I&C system LC is a sequence of N stages. At each i-th stage of a multi-version I&C system LC Mi of diversity types may be applied. From Mi, i = 1,...,N; diversity types only a single j-th type, j = 1,...,Mi, may be selected. Besides, at each i-th stage of LC a single-version development technology may be selected. Each j-th diversity type at each i-th LC stage is characterized by two indices: diversity metrics (depth) dij and cost of respective diversity type application (cost increase as compared to single-version option of each i-th LC stage).

Thus, a set of solutions on selection of diversity kind is described by two matrices: diversity metrics values D = || d_{ij} || and cost values C = || c_{ij} ||. Hence MVS LC may be presented as a bipolar N-level graph (Fig.7) called graph of multi-version technologies (Sklyar &Kharchenko, 2007). MVT corresponds to non-zero way in this graph.
Algorithms of MVT (optimal way in the graph) selection according with criteria "diversity (safety)-reliability-cost" are described in (Kharchenko&Sklyar, 2008).

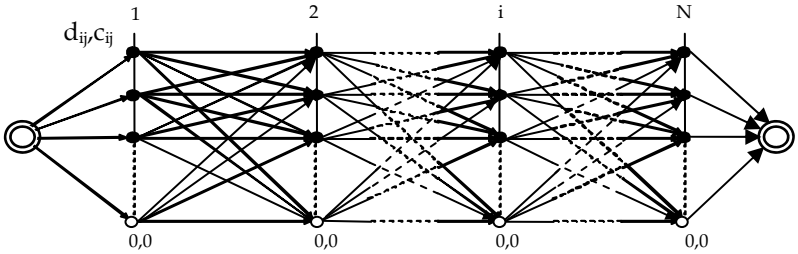


Fig. 7. Graph of MVTs

5. Assessment of multi-version FPGA-based systems safety

5.1 General approach to assessment

Assessment of diversity level and MVS safety is based on the following basic procedures analysis and evaluation:

- check-list-based analysis of applicable diversity types (CLD); initial data for the CLD analysis are I&C design and documentation, a table of diversity types (subtypes) was developed in advance; a result of the CLD analysis is a formalized structured information about used diversity types and subtypes in analyzed I&C system;
- metric-based assessment of diversity (MAD); initial data for the MAD procedure are results of the CLD analysis and values of metrics and weight coefficients for diversity types (subtypes) used in I&C systems; a result of the MAD assessment is a value of general diversity metric;
- Reliability Block Diagram (RBD) and Markovian model (MM)-based assessment taking into account results of MAD.

5.2 Stages of assessment

The main stages and operations of diversity analysis and MVS assessment depend on the type of the evaluated system. The first stage is a Check-list-based analysis of MVS design and documentation. This stage contains two operations:

1. Analysis of I&C specification and requirements to system, definition of system safety class; requirements to diversity (necessary for diversity application);
2. Analysis of I&C design and development process that involves activities: (a) identification of MVS types: which of the subsystems are FPGA-based and which are software and microprocessor-based; (b) identification of product diversity; for FPGA-based MVSs: manufacturer of chips; FPGA technology; FPGA families; FPGA chips, languages; tools, etc); (c) identification of process diversity kinds.

Results of analysis are entered in a check-list in accordance with rule Yes (if corresponding diversity type is used in a system) / No (in opposite case) and is presented as a n-bit Boolean vector.

The second stage is a metric-based assessment of diversity. This stage contains two operations:

1. Determination of metric values for different types of applied diversity, i.e. performing two activities: (a) determination of metric values (local diversity metrics μ_i for diversity type d_i and local diversity metrics μ_{ij} for diversity subtype d_{ij}); the metric values may be predefined; (b) correction of metric values in accordance with development and operation experience.
2. Calculation of general diversity metric μ for a system: (a) determination (correction) of weight coefficients ω_i (ω_{ij}) of metrics (taking into account multi-diversity aspect); sum of weight coefficients ω_i (ω_{ij}) is equal 1; (b) convolution (additive or more complex) of metrics and calculating value of general diversity metric $\mu = \Sigma\, \omega_i \Sigma\, \omega_{ij}\, \mu_{ij}$, i = 1,…, n; j = 1,…$n_i$.

Thus, result of this stage is a value of general diversity metric μ, which is some approximation of β, and can characterize the diversity effect on CCF probability.

The third stage is a probabilistic RBD- or MM-based (RDM) assessment of MVS reliability and safety. Initial data for the RDM procedure are I&C design and documentation, results of the CLD and MAD analysis; results of the RDM procedure are values of safety and dependability indicators. Detailed description of the RDM procedure is given in (Kharchenko et al., 2004).

6. Implementation of FPGA-based safety-critical NPP I&Cs

6.1 General description of the FPGA-based RADIY™ platform

The platform RADIY™ produced by RPC Radiy is an example of a dependable and scalable FPGA-based I&C platform ensuring possibility of development of multi-version systems. Dependability assurance feature of the I&C platform RADIY™ is multi-diversity implementation through the following diversity types: equipment diversity is provided by different electronic components, different programmable components (FPGAs and microcontrollers) and different schemes of units; software diversity is provided by different programming languages and different tools for development and verification; life cycle (human) diversity is provided by different teams of developers.

Scalability of the I&C platform RADIY™ permits to produce different types of safety-critical systems without essential changing of hardware and software components. The I&C platform RADIY™ provides the following types of scalability: scalability of system functions types, volume and peculiarities by changing quantity and quality of sensors, actuators, input/output signals and control algorithms; scalability of dependability (safety integrity) by changing a number of redundant channel, tiers, diagnostic and reconfiguration procedures; scalability of diversity by changing types, depth and criteria of diversity choice.

The FPGA-based I&C RADIY™ platform comprises both upper and lower levels (Kharchenko&Sklyar, 2008). The upper level has been created on purchased IBM-compatible industrial workstations. The software for the upper level RADIY™ platform was developed by RPC Radiy and is loaded on the workstations. The functions of the upper level workstations are the following: receipt of process and diagnostic information; creation of

man-machine interface in the Control Room; display of process information on each of the control algorithms relating to control action executed by I&C system components; display of diagnostic information on failures of I&C system components; registration, archiving and visualization of process and diagnostic information.
The lower level of the RADIY™ platform consists of standard cabinets including standard functional modules blocks). The RADIY™ platform comprises the following standard cabinets (Bakmach et al., 2009):

- Normalizing Converters Cabinets performs inputting and processing of discrete and analog signals as well as feeding sensors;
- Signal Forming Cabinets performs inputting and processing of discrete and analog signals, processing of control algorithms, and formation of output control signals;
- Cross Output Cabinets receives signals from three control channels (signal formation cabinets) and forms output signals by "two out of three" mode;
- Remote Control Cabinets controls 24 actuators on the basis of Control Room signals, automatic adjustment signals and interlocks from signal formation cabinets;
- Signalling Cabinets forms control signals for process annunciation panel at Control Room and others.

The platform includes the following main modules: chassis and backplanes; power supply modules; analog input modules; normalizing converter modules, thermocouples; normalizing converter modules, resistive temperature detector; discrete input modules; discrete information input modules, pulse; potential signals input modules, high voltage; protection signal forming modules (logic modules); analog output modules, voltage; analog output modules, current; discrete output modules; potential signal output modules; solid-state output modules; relay output modules; actuator control modules; fiber optic communication modules; system diagnostic modules; fan cooling modules etc.

6.2 Opportunites of the RADIY™ platform

Application of the RADIY™ platform with the use of FPGA technology provides the following opportunities:

- to implement control and other safety-critical functions in the form of FPGA with implemented electronic design, without software;
- to use software only for diagnostics, archiving, signal processing, data reception and transfer between I&C systems components; failures of those functions do not affect execution of basic I&C systems control functions, and an operation system is not applied at I&C systems lower levels;
- to process parallel of all control algorithms within one cycle, thus ensuring high performance of the system (for instance, a processing cycle of Reactor Trip System is 20 ms) and proven determined temporal characteristics;
- to develop the software-hardware platform in such a way that it becomes a universal interface to create I&C systems for any type of reactors;
- to assure high reliability and availability due to the application of industrial components as well as using the principles of redundancy, independency, single failure criterion, and diversity;
- to modify the I&C system after commissioning in a quite simple manner, including algorithm alterations, without any interference in I&C systems' hardware structure;

- to reduce by more than 10 times the number of contact and terminal connections which cause many operational failures of equipment on account of the wide use of integrated solutions and fiber optic communication lines, etc.

6.3 Licensing of the RADIY™ platform

The RADIY™ platform has been licensed for NPP application in Ukraine and in Bulgaria. The main idea for licensing FPGA-based NPP I&C systems lays in consideration of FPGA-chip as hardware and FPGA electronic design as a special kind of software with specific development and verification stages (Siora et al., 2009b).

Qualification tests of FPGA-based hardware in accordance with International Electrotechnical Commission (IEC) standard requirements include: radiation exposure withstand qualification; environmental (climatic) qualification; seismic and mechanical impacts qualification; electromagnetic compatibility qualification. Results of qualification tests confirmed FPGA-based hardware compliance with IEC safety requirements.

FPGA electronic design has a V-shape life cycle in accordance with requirements of standard IEC 62566 "NPP – I&C important to safety – Selection and use of complex electronic components for systems performing category A functions".

The safety assessments have been conducted by Ukrainian State Scientific Technical Centre on Nuclear and Radiation Safety (SSTC NRS), which is the supporting organization of Ukrainian Regulatory Authority. Experts of SSTC NRS have considerable experience in the area of FPGA-based systems safety assessment, as they have performed reviews of all thirty three FPGA-based safety systems supplied to Ukrainian NPP units since 2003.

6.4 Implementation of the RADIY™ platform-based I&Cs in NPPs

The RADIY™ platform has been applied to the following NPP I&Cs systems which perform reactor control and protection functions:

Reactor Trip System (RTS); these I&Cs were developed as two-version systems consisting of two triple module redundant subsystems;

Reactor Power Control and Limitation System; Engineering Safety Features Actuation System (ESFAS); Control Rods Actuation System; Automatic Regulation, Monitoring, Control, and Protection System for Research Reactors; these I&Cs were developed as one-version systems consisting of triple module redundant subsystems.

The first commissioning of the RADIY™ platform was done in 2003 for Ukrainian NPP unit Zaporozhe-1. In seven years since that time, more than 50 applications of RPC Radiy systems have been installed in 17 nuclear power units in Ukraine and Bulgaria. These systems are commissioned in pressurized water reactor (PWR) plants known as "VVER" reactors developed by design companies of the former Soviet Union. VVER reactors are used in Armenia, Bulgaria, China, Czech Republic, Finland, Hungary, India, Iran, Russia, Slovakia, and Ukraine.

The largest project realized by RPC Radiy is the modernization of six ESFASs for Bulgarian NPP Kozloduy (three ESFASs for Kozloduy-Unit 5 and three ESFASs for Kozloduy-Unit 6).

7. Conclusion

Development and implementation of multi-version FPGA-based systems is new stage of evolution in area of improving safety of NPP I&Cs. In this chapter we discussed basic concepts of diversity as a key approach to decreasing probability of common cause failure of

safety-critical I&Cs and the taxonomic scheme of multi-version computing as a part of dependable, safe and secure computing.
Known version redundancy classification schemes were generalized in three-space matrix ("cube of diversity") taking into account features of FPGA technology. It is unique technology allows to simplify NPP I&C development and verification, realize multi-reconfiguration (dynamical function- and dependability-oriented architecting, multi-parametrical space-structural adaptation, etc.), to propose decisions with different product-process version redundancy.
Key challenges related to diversity-oriented and FPGA-based systems are the following: existing standards are not enough detailed to make all necessary decisions concerning diversity (all the more FPGA-based decisions); multi-version I&Cs are still unique, failures occurred rarely and information about failures is not enough representative and accessible; methods of diversity assessment and kind selection, as a rule, are based on expert approach.
FPGA technology allows developing multi-version systems with different product-process version redundancy, diversity scalable multi-tolerant decisions for safety-critical NPP I&Cs.
Described models of multi-version systems and multi-version technologies (life cycle) may support selecting of cost-effective technique and optimal architecture according with requirements to diversity, safety, reliability and limitation of applied technologies. These theoretical issues were used on development of FPGA-based I&C RADIY™ platform. Main peculiarities of the platform are realization of control and other safety-related functions without software and ensuring dependability- and diversity-scalable decisions of safety-critical I&C. Experience of RPC Radiy has proved effectiveness of these decisions.

8. References

Altera Data Book (2001). *APEX II Programmable Logic Device Family. Data Sheet*, Ver.1.3.

Avizienis A.; Laprie J.-C., Randell, B. & Landwehr, C. (2004). Basic Concepts and Taxonomy of Dependable and Secure Computing. *IEEE Transactions on Dependable and Secure Computing*, vol.1, (2004), pp. 11-33.

Bakhmach, E.; Kharchenko, V., Siora, A., Sklyar, V. & Tokarev, V. (2009). Advanced I&C Systems for NPPS Based on FPGA Technology: European Experience. *Proceedings of 17th International Conference on Nuclear Engineering (ICONE 17)*. ISBN: 978-0-7918-3852-5, Brussels, Belgium, July, 2009.

Bobrek, M.; Bouldin, D; Holkomb, D. et al. (2009). *Review Guidelines for FPGAs in Nuclera Power Plants Safety Systems*, NUREG/CR-7006 ORNL/TM-2009/020.

Bukowsky, J. & Goble, W. (1994). An Extended Beta Model to Quantize the Effects of Common Cause Stressors, *Proceedings of ISAFECOMP,* London, October, 1994.

Gorbenko, A.; Kharchenko V. & Romanovsky A. (2009). Using Inherent Service Redundancy and Diversity to Ensure Web Services Dependability In *Methods, Models and Tools for Fault Tolerance*, M. Butler, C. Jones, A. Romanovsky, E. Troubitsyna (Eds.), pp. 324-341, LNCS 5454, Springer.

Kharchenko V.; Yastrebenetsky M. & Sklyar V. (2004). Diversity Assessment of Nuclear Power Plants Instrumentation and Control Systems, *Proceeding by 7th International Conference on PSAM and ESREL Conference*, pp.1351-1356, Vol.3, Berlin, Germany, July, 2004.

Kharchenko, V. & Sklyar, V. (Eds.). (2008). *FPGA-based NPP Instrumentation and Control Systems: Development and Safety Assessment.*, RPC Radiy, National Aerospace

University KhAI, State STC on Nuclear and Radiation Safety, ISBN 978-966-96770-2-0, Kharkiv & Kirovograd, Ukraine.

Kharchenko, V. (1999). Multi-version Systems: Models, Reliability, Design Technologies, *Proceeding of 10th ESREL Conference*, pp. 73-77, Vol.1, Munich, Germany, September, 1999.

Kharchenko, V.; Bakhmach, E. & Siora, A. (2009). Diversity-scalable decisions for FPGA-based safety-critical I&C systems: From theory to implementation. *Proceedings of the 6th Conference NPIC&HMIT*, Knoxville, Tennessee, American Nuclear Society, LaGrange Park, IL. ISBN: 978-0-89448-067-6, April, 2009.

Kharchenko, V.; Siora, A., Sklyar, V. & Tokarev, V. (2010). Diversity-Oriented FPGA-Based NPP I&C Systems: Safety Assessment, Development, Implementation, *Proceeding by 18th International Conference on Nuclear Engineering (ICONE18)*, Xi'an, China, May, 2010.

Kharchenko, V.; Siora, A., Sklyar, V., Volkoviy, V. & Bezsaliy, V. (2010). Multi-Diversity Versus Common Cause Failures: FPGA-Based Multi-Version NPP I&C Systems, *Proceedings of the 7th Conference NPIC&HMIT*, Las-Vegas, Nevada, USA, November, 2010.

Kharchenko, V.; Sklyar, V. & Tarasyuk, O. (2003). Risk Analysis of Accidents of Space-Rocket Technik: Evolution of Reasons and Tendencies. *Radio-Electronic and Computer Systems*, Vol.3, (May, 2003), pp. 135–149, National Aerospace University KhAI, Kharkiv, Ukraine.

Kharchenko, V.; Sklyar, V. & Volkoviy, A. (2007). Multi-Version Information Technologies and Development of Dependable Systems out of Undependable Components, *Proceedings of International Conference on Dependability of Computer Systems*, pp. 43-50, Szklarsla Poreba, Poland, July, 2007.

Kharchenko, V.; Sklyar, V., Siora, A., Tokarev, V. (2008). Scalable Diversity-oriented Decisions and Technologies for Dependable SoPC-based Safety-Critical Computer Systems and Infrastructures, *Proceeding of IEEE International Conference on Dependability of Computer Systems*, pp. 339-346, Szklarska Poreba, Poland, July, 2008.

Li, B.; Xu, Y. & Choi, J. (1996). Applying Machine Learning Techniques, *Proceedings of ASME 2010 4th International Conference on Energy Sustainability*, pp. 14-17, ISBN 842-6508-23-3, Phoenix, Arizona, USA, May, 2010.

Lima, P.; Bonarini, A. & Mataric, M. (2004). *Application of Machine Learning,* InTech, ISBN 978-953-7619-34-3, Vienna, Austria.

Naser (Ed.) (2009). *Guidelines on the Use of Field Programmable Gate Arrays (FPGAs) in Nuclear Power Plant I&C Systems*, EPRI, Palo Alto, CA: 2009 1019181.

Prokhorova, Y.; Kharchenko V.; Ostroumov, B.; Ostroumov, S. & Sidorenko, N. (2008) Dependable SoPC-Based On-board Ice Protection System: from Research Project to Implementation, *Proceeding of IEEE International Conference on Dependability of Computer Systems*, pp. 312-317, Szklarska Poreba, Poland, July, 2008.

Pullum L. (2001). *Software Fault Tolerance Techniques and Implementation*, Artech House Computing Library.

Siora, A.; Sklyar, V., Rozen Yu., Vinogradskaya, S. & Yastrebenetsky, M. (2009). Licensing Principles of FPGA-Based NPP I&C Systems, *Proceedings of 17th International*

Conference on Nuclear Engineering (ICONE 17), Brussels, Belgium, ISBN: 978-0-7918-3852-5, July, 2009.

Siora, A.; Krasnobaev, V. & Kharchenko, V. (2009). *Fault-Tolerance Systems with Version-Information Redundancy*. Ministry of Education and Science of Ukraine, National Aerospace University KhAI. ISBN 978-966-96770-7-5.

Sklyar, V. & Kharchenko, V. (2007). A Method of Multi-version Technologies Choice on Development of Fault-Tolerant Software Systems. *Proceeding of Workshop on Methods, Models and Tools for Fault Tolerance*, pp.148-157, Oxford, UK, July, 2007.

Sommerville, J. (2011). *Software Engineering*. 9th edition, Addison-Wesley, ISBN 9-780-13-703515, England.

Tarasyuk, O., Gorbenko, A., Kharchenko, V., Ruban, V. & Zasukha, S. (2011). Safety of Rocket-Space Engineering and Reliability of Computer Systems: 2000-2009 Years. *Radio-Electronic and Computer Systems,* Vol.11, (March, 2011), pp.23-45, National Aerospace University KhAI, Kharkiv, Ukraine.

Volkovij, A.; Lysenko, I., Kharchenko, V. & Shurygin, O. (2008). *Multi-Version Systems and Technologies for Critical Applications,* National Aerospace University KhAI, Kharkiv, Ukraine.

Wood, R..; Belles, R., Cetiner, M. & et al, (2009). *Diversity Strategies for NPP I&C Systems*, NUREG/CR-7007 ORNL/TM-2009/302.

Yastrebenetsky, M. (Ed.) (2004). *Safety of Nuclear Power Plants: Instrumentation and Control Systems*, Technika, Kyiv, Ukraine (Translated by NRC, USA, 2007).

3

Nuclear Power Plant Instrumentation and Control

H.M. Hashemian
Analysis and Measurement Services Corp.
United States

1. Introduction

Installed throughout a nuclear power plant, instrumentation and control (I&C) is an essential element in the normal, abnormal and emergency operation of nuclear power plants (International Atomic Energy Agency [IAEA], n.d.). Through their equipment, modules, sensors, and transmitters, I&C systems measure thousands of variables and processes the data to activate pumps, valves, motors, and other electromechanical equipment that control the plant. The I&C system senses basic physical parameters, monitors performance, integrates information, and makes automatic adjustments to plant operations to keep process variables within the plant design limits. By reacting appropriately to failures and abnormal events, I&C ensures the plant's safety and efficient production of power (U.S. Nuclear Regulatory Commission [U.S. NRC], 2011).

All of these roles can be reduced to three basic functions (IAEA, 1999). First, as the plant's nervous system, I&C provides plant operators with accurate and relevant information so they can make the appropriate actions during normal as well as abnormal operation. Second, I&C provides plant operators with the capacity to exercise automatic control over the plant and its associated systems so they can take whatever actions are needed to maintain efficient and safe operation. Finally, I&C serves the critical function of protecting the plant from faults in the system or errors made by the operator as well as abnormal or extreme external events that threaten the plant's operation. More specifically, I&C should enable the plant to operate safely for an extended period without operator intervention following an accident (IAEA, 1999).

Nuclear plant I&C systems must be *accurate* to properly sense and communicate the process variables and reasonably *fast* to provide timely display, adjustment, and protection against upsets in both the main plant and its ancillary systems. For example, temperature sensors such as resistance temperature detectors (RTDs), which are key elements in the safety system instrumentation of nuclear power plants, may be expected to provide 0.1 percent accuracy and respond to a step change in temperature in less than 4 seconds.

Nuclear plant I&C is more complex and varied than the control instrumentation in other industrial applications because of the special nature of nuclear power. A nuclear plant's production must remain continuous because of its high capital costs, direct access to and control over the nuclear plant's reactor is impossible, and the potential risks of nuclear energy production require greater redundancy and reliability in plants' control infrastructure (IAEA, 1999). Although I&C is a relatively small component in a typical plant's maintenance and

capital upgrade budget, its impact on the plant's safety, reliability, and performance is preeminent (Hurst 2007). For example, assuming that a 1000 MWe plant has a daily operating revenue of about $2 million per day, a loss in power production level of even 1 percent can quickly amount to millions of dollars in lost revenue.

Some 10,000 sensors and detectors and 5,000 kilometers of I&C cables—representing a total mass of 1,000 tons—comprise the I&C system of a typical nuclear plant unit, including up to 20 neutron detectors, 60 RTDs, as many as 100 thermocouples, and 500 to 2,500 pressure transmitters (IAEA, n.d.; Hashemian, forthcoming). Categorized by function, I&C components consist of:

- Sensors that interact with the plant's physical processes to measure process variables such as temperature, pressure and flow as well as control, regulation, and safety components that process the sensors' data.
- Communication infrastructure—wires and cables, fiber-optic and wireless networks, digital data protocols—that move sensor and control data through the I&C system.
- Human-system interfaces such as displays that enable human plant operators to monitor and respond to the continual flow of I&C data.
- Surveillance and diagnostic systems that monitor sensor signals for abnormalities.
- Actuators such as valves and motors that physically operate the plant's control and safety components to adjust physical processes so the plant's performance is optimized for efficiency and safety or, if needed, shut down.
- Actuator status indicators that visually reflect automatic or manual control actions, such as the switching on or off of a motor or the opening or closing of a valve (IAEA, n.d.).

2. Important I&C components

Nuclear plant instrumentation can generally be classified into the following four categories:

- *Nuclear:* instruments that measure nuclear processes or reactor power, such as neutron flux density.
- *Process:* instruments that measure non-nuclear processes such as reactor pressure, coolant or pressurizer level, steam flow, coolant temperature and flow, containment pressure, etc.
- *Radiation monitoring:* instruments that measure radiation, for example, in monitoring radiation in steam lines, gas effluents, and radiation at the plant site.
- *Special:* Instruments encompassing all other applications, such as for measuring vibration, hydrogen concentration, water conductivity and boric acid concentration or meteorological, seismic, or failed fuel detection applications (IAEA, 1999).

The variety of I&C components and applications notwithstanding, temperature, pressure, level, flow, and neutron flux remain the most important and safety-critical measurements for the control and safety protection of nuclear reactors. The heart of each of these measurements is the sensor itself--the most important component in an instrument channel and the one that usually resides in the harsh environment of the field (Hashemian, 2007). Despite the accelerating advances in I&C technology (to be discussed in the next section), the basic mechanism of measurement used by these sensors has not changed significantly since the earliest nuclear plants. Today, temperature, pressure, level, flow, and neutron flux are still primarily measured using conventional sensors such as resistance temperature detectors (RTDs), thermocouples, capacitance cells, bellows, force-balance sensors, and conventional neutron detectors although some advances have been made in developing new neutron detectors for nuclear power plants (Hashemian, 2009a).

The control and safety of nuclear power plants depend above all on temperature and pressure (including differential pressure to measure level and flow) instrumentation—the two most ubiquitous instrument types in a typical nuclear power plant process. In pressurized water reactor (PWR) plants, RTDs are the main sensors for primary system temperature measurement. RTDs are thermal devices that contain a resistance element referred to as the sensing element. Two groups of RTDs are typically used in nuclear power plants: direct immersion (or wet-type) and thermowell mounted (or well-type). The resistance of the sensing element changes with temperature, and therefore by measuring the resistance, one can indirectly determine the temperature. The number of RTDs in a nuclear power plant depends on the plant design and its thermal hydraulic requirements. For example, PWR plants have up to 60 safety-related RTDs while heavy water reactors such as Candu plants have several hundred RTDs.

Pressure transmitters are the next most common I&C component. A pressure transmitter may be viewed as a combination of two systems: a mechanical system and an electronic system. The pressure transmitter's mechanical system contains an elastic sensing element (diaphragm, bellows, Bourdon tube, etc.) that flexes in response to pressure applied. The movement of this sensing element is detected using a displacement sensor and converted into an electrical signal that is proportional to the pressure. Typically, two types of pressure transmitters are used in most nuclear power plants for safety-related pressure measurements. These are referred to as motion-balance and force-balance, depending on how the movement of the sensing element is converted into an electrical signal.

A nuclear power plant generally contains about 400 to 1200 pressure and differential pressure transmitters to measure the process pressure, level, and flow in its primary and secondary cooling systems. The specific number of transmitters used in a plant usually depends on the type and design of the plant. For example, the number of transmitters used in PWRs depends on the number of reactor coolant loops. Figure 1 illustrates a typical process instrumentation channel in a nuclear power plant.

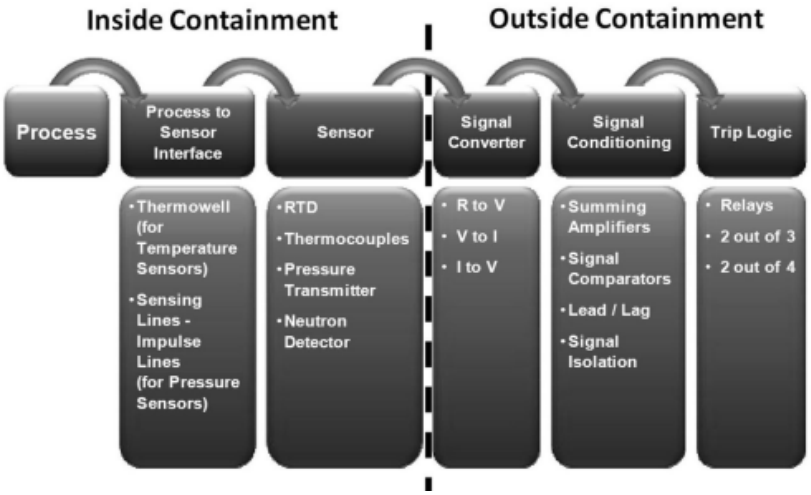


Fig. 1. Typical Instrumentation Channel in Nuclear Power Plant (R = resistance; V = voltage; I = current).

3. Evolution of I&C

The evolution of I&C has been marked by three generational shifts. In the first, analog technology was used for instrumentation, and mechanical relay-based equipment was used for control of discrete processes. The second generation of I&C was marked by the use of discrete or integrated solid-state equipment for both instrumentation and control. The emergence of the microprocessor in the late 1970s made possible the replacement of mechanical relays by programmable logic controllers (PLCs). PLCs were initially used in non-nuclear applications in nuclear plants, but their evolving ability to handle large volumes of data, perform mathematical calculations, execute continuous process control, and communicate with computers brought them into plants' nuclear applications. The third generation of I&C is digital, to be discussed in the next section.

One of the key forces driving the evolution of I&C has been the obsolescence of analog equipment. A second driver has been technological: new information, electronic, display, and digital technologies seem tailor made for the NPP I&C environment, where complexity rules, automation is essential, and high initial infrastructure cost can be rationalized (IAEA, 1999). Though sensor technology itself has not changed significantly, other I&C systems have—perhaps more so than any other area of nuclear power plant science, offering quantum functionality and performance improvements.

A third driver has been accidents, like Three Mile Island, Chernobyl, and Fukushima, which force I&C system designers to reevaluate operating principles, system robustness and safety margins, and accident probability assumptions. For example, both Three Mile Island and Fukushima underscored the critical role of I&C signals in enabling operators to understand the nature of the accident they are facing. On a general level, Three Mile Island helped stimulate new research and development into signal validation, ultimately spawning the discipline of on-line monitoring (to be discussed later in this chapter). Specifically, Three Mile Island led directly to the adoption of safety parameter display systems. Both Chernobyl and Fukushima forced I&C designers to focus more on analyzing the potential occurrence of very rare events that would once have been considered non-'design basis events' so their consequences might be mitigated.

A fourth driver of changes in I&C has been economic. Enhanced I&C means greater knowledge of and control over plant conditions and therefore greater leeway in pushing plant operating limits and extending uptime. More in-core instrumentation, redundant and diverse instrumentation providing deeper comparative operational databases, and enhanced qualification, calibration and maintenance have enabled plants to uprate their power profiles without sacrificing safety margins (IAEA, 1999).

Because the cost of building new plants is so high, regulatory hurdles are so substantial, and political resistance to nuclear power so significant, few new plants have been built. Instead, existing plants are relicensed for extended lives far beyond their original design assumptions. Nuclear power plants that operate for 60 years, for example, live through three generations of I&C evolution (the qualified life of most nuclear plant pressure transmitters and RTDs is typically about 20 years, although most properly maintained pressure transmitters last longer than 20 years) (IAEA, n.d.). In the mid-1980s, the nuclear industry began to talk about aging and obsolescence in analog I&C equipment (Hashemian, 2009a). In this plant-life extension climate, enhanced, digital I&C became a way to offset the plant's age by giving operators new eyes and ears for staying on top of the continuing aging-

induced degradation of the plant. New fatigue monitoring and 'condition limitation' systems have made it possible to minimize disturbances and smooth out transients (IAEA, 1999). Typically, plants will replace I&C in steps or modularly, swapping out a discrete analog control system with a digital one, but retaining the existing field cabling, sensors, and actuators (IAEA, 1999).

I&C system advances as a result of these drivers have produced a significant improvement in plant capacity factor, outage time duration, personnel radiation exposure, power uprates, and operational efficiency (Hashemian 2009b). However, it remains the case today that the bulk of I&C systems used to monitor and control existing NPPs use analog process technology developed in the 1950s and 1960s (IAEA, 1999).

4. Emergence of digital I&C

Digital I&C evolved from microprocessor-based PLCs and plant process-monitoring computers (IAEA, 1999). Because they can be programmed to perform complex tasks, microprocessors quickly replaced analog relays and spawned new applications in plant monitoring and control systems, including graphical display interfaces so human operators could observe and interact with the I&C system (IAEA, 1999). The first protection systems using digital technology, known as "core protection calculators," were implemented on combustion engineering designed reactors in the late 1970s (Bickel, 2009). In the 1980s, digital technology was integrated into control systems for NPPs' auxiliary subsystems. Digital relays and recorders, smart transmitters, and distributed control systems (DCSs) were implemented primarily in non-safety systems such as feedwater control, main turbine control, and recirculation control (U.S. NRC, 2011; IAEA, 1999).

By the 1990s, microprocessors were being used for data logging, control, and display for many nonsafety-related functions (U.S. NRC, 2011). In 1996, the first fully digitalized I&C system was integrated into Japan's Kashiwazaki-Kariwa Unit 6 advanced boiling-water reactor (ABWR), followed by Kashiwazaki-Kariwa Unit 7 in Japan (U.S. NRC, 2011; Hashemian 2009a). In the 2000s, all-digital I&C systems for both safety-related systems and safety-critical systems were implemented worldwide (IAEA, 1999). For example, France, the United Kingdom, Korea, and Sweden, among other countries, implemented digital I&C systems in their nuclear power plants (U.S. NRC, 2011; Hashemian 2009a). Today, about 40% of the world's operating power reactors in almost all of the thirty nations with operating NPPs have been upgraded to some level of digital I&C. Ten percent of such installations have occurred at new reactors, with the rest involving upgrades at existing reactors (IAEA, n.d.). Since 1990, all of the reactors under construction worldwide have some digital I&C components in their control and safety systems (IAEA, n.d.).

Today, control panel instruments such as controllers, display meters, and recorders are mostly digital. Most diagnostic and measuring equipment is digital, and increasingly common digital transducer transmitters now offer so-called smart features like automatic zeroing and calibration (IAEA, 1999). Similarly, digital I&C systems like Westinghouse's Eagle 21, Common Q, and Ovation systems, Areva Nuclear Power's Teleperm XS, the Triconix Company's TRICON system and Rolls Royce's Spinline are available for retrofitting implementation on existing plants' safety-related applications or in new all-digital plants (U.S. NRC, 2011; Hashemian 2009a; IAEA, 2008). The advanced boiling water reactor (ABWR) plants built in Japan for more than a decade all use fully integrated digital I&C systems for both safety-related and nonsafety-related plant control

and protection (Hurst 2007). Finally, the new reactor designs that have already won certification (including the AP1000, System 80+, and ABWR) will make extensive use of digital I&C (Oak Ridge National Laboratory [ORNL], 2007). To satisfy the demanding operational environments of new designs, ranging from high temperatures to high neutron flux (not to mention the post-Fukushima demands for I&C that can survive "beyond design basis" conditions), advanced and in many cases digital sensors, detectors, transmitters, and data transmission lines will continue to be needed (IAEA, n.d.).

4.1 Benefits of digital

The attractions of digital I&C are many. First, by minimizing the number of analog circuits required to perform an I&C measurement, digital processing reduces the potential interference (noise) and drift that result from using multiple analog circuits. This makes possible more accurate or precise measurements, which can be further refined through digital data processing programs (IAEA, 1999; ORNL, 2007; Lipták, 2006). Second, measurement parameters can be much more easily modified with digital systems than with analog systems. In contrast to the physical reconfiguration of an analog device, modifying digital I&C merely requires loading a different program, which greatly enhances versatility. Shifting functionality from hardware to software in this way means quicker installation of I&C components (IAEA, 1999; ORNL, 2007; Lipták, 2006). Third, the increasingly miniaturized integrated circuits in digital I&C offer substantial processing power relative to device size, greatly reducing the space required for I&C equipment. Fewer and smaller devices capable of transmitting higher concentrations of data using multiplexing also translates into minimized cabling needs. Both the number and quality of I&C links in a plant can be increased (IAEA, 1999; ORNL, 2007). Fourth, digital technology's processing power means more complex functional capabilities for I&C, from on-line power density limit computation and dead-time and temperature measurement correction to highly specifiable and versatile signal filtering (IAEA, 1999). Fifth, by offering greater automation possibilities, digital I&C minimizes the need for human intervention, thus minimizing the possibility of human error. Sixth, because digital I&C systems can perform automatic self-testing much more easily than analog systems, they reduce maintenance costs and improve reliability through continuous monitoring capability. Such self-testing functionality greatly aids in analyzing system faults (IAEA, 1999).

4.2 Emerging sensors for digital I&C

Although the core technology of nuclear plant sensors has remained largely unchanged since the inception of the industry, since the 1990s several new sensor technologies have been conceived, and some prototyped, that may find adoption in the next-generation nuclear power plants. The extreme high temperatures of next-generation reactors are probably the most significant driver of and technical challenge facing new sensor development today. While the current generation of industrial RTDs can accurately measure processes up to about 400°C, some Gen IV reactors are expected to operate at coolant temperatures three or four times higher than light water reactors—that is, up to about 1,000°C (Hashemian, forthcoming).

Emerging sensors fall into three main categories: (1) so-called next-generation sensors, (2) fiberoptic sensors; and (3) wireless sensors (Hashemian 2009a; Hashemian 1999).

4.3 Next-generation sensors

Next-generation sensors encompass advanced sensor designs that will only find application in the longer term, 20-30 years from the present (Hashemian, forthcoming). Solid-state and Silicon Carbide (SiC) neutron flux monitors, magnetic flow meters, hydrogen sensors, virtual sensors, Nanotriodes, gamma ray tomographic spectrometers, fuel mimic power monitors, and Quantum Cascade Laser infrared sensors that sniff emissions and detect overheating, odor, burning, and fumes are among the designs currently in the R&D stage at Oak Ridge National Laboratory (ORNL), Ohio State University, Idaho National Laboratory (INL) and other facilities (Hashemian 2008). One advanced sensor that is closer to actual implementation in nuclear power plants is the Johnson noise thermometer, which consists of an RTD whose open-circuit voltage is measured and related to temperature. This essentially drift-free sensor measures absolute temperature, and its reading is independent of RTD characteristics (Hashemian 2009a). The sensor was developed at ORNL and is ready for commercialization.

Because flow is an inherently difficult parameter to measure and most industrial flow measurement techniques have large uncertainties, flow measurement is another area where advanced sensor types may find application in the longer term (Hashemian 1999). For example, one conventional method, measuring differential pressure across venturi flow elements, is susceptible to fouling, which causes erroneous flow indication. Ultrasonic flow meters address this because they do not depend on venturi elements or other constrictions in the pipes. Rather, they measure flow by sending an ultrasonic signal through the fluid and measuring the time that it takes for the signal to travel through the fluid from the signal source to a downstream signal receiver and back again. Referred to as "transit time," the signal travel time depends on the fluid flow rate (Hashemian 1999).

Despite the long-term promise of advanced sensor types, in the short term--the next 10-15 years--advances in sensors and transmitters are expected to center primarily on fiber-optic and wireless sensors (Hashemian, forthcoming).

4.4 Fiber optic sensors

Fiber optic technologies are emerging as a potential near-term sensor class for future nuclear power plants (Hashemian 1999). Fiber-optic sensors offer driftless accuracy and high sensitivity, light weight and small size, ease of installation, low power requirements, immunity to electromagnetic interference (EMI), potential for multiplexing (several sensors can be used with a single transmission cable), large bandwidth, and reliability and environmental ruggedness. Moreover, since some fiber-optic sensing modulation techniques are digital in nature, fiber-optic sensing can easily be made compatible with digital control systems (Hashemian, forthcoming).

Fiber-optic sensors operate on the principle that environmental effects or displacements can be converted into measureable optical signals. Fiber-optic sensors can be divided into two broad categories based on the way in which the process measurement is applied to the fiber: extrinsic (or hybrid) and intrinsic (or all-fiber). In extrinsic or hybrid sensors, the sensing element itself is often similar to those in conventional sensors, but fiber optics are used to sense the movement of the sensing element (as with a strain gage) and then convert it into an electrical signal.

In contrast, in intrinsic or all-fiber sensors, the fiber itself senses the environmental effect and itself transmits the affected light beam to a device that converts it into a measurement. The three most advanced fiber-optic sensor technologies—those most likely to replace the

functionality of conventional non-fiber-optic sensors now installed in nuclear power plants—are single-point interferometry, distributed fiber Bragg grating, and optical counter and encoder techniques.
Fiber-optic temperature sensors are the most mature fiber-optic sensor types, with some commercially available types able to withstand operational temperatures of up to about 450°C (Hashemian, forthcoming). Longer term, new sensor principles based on the transmission modes of fiber optic devices may also emerge (IAEA, 1999).

4.5 Wireless sensors

While sensor technologies change slowly, rapid advances have been made in networking technology to wirelessly transmit sensor data to a monitoring system (IAEA, 2008). So-called wireless sensors usually consist of a conventional sensing device such as a thermocouple, resistance temperature detector (RTD), or strain gauge as well as circuitry to convert the sensor output into an electrical signal (voltage or current), filter the signal, digitize it, and transmit it to a receiver. If fast data acquisition is required, the data is sometimes processed at the sensor, and the results are then transmitted. For example, averaging and fast Fourier transform (FFT) can be performed at the sensor. Faster data rates consume more battery power, and data processing at the sensor places additional demands on any battery (Hashemian 2008).
In nuclear plants, equipment is typically spread over a large footprint, and data is gathered through wires that are drawn through conduits buried in trenches. Moreover, much of the cost of adding new instrumentation to existing equipment in a nuclear plant lies in the cabling. Wireline networks usually impose high cabling and installation costs, which can exceed $1000 per linear foot in typical nuclear power plants. A recent project funded in part by the Electric Power Research Institute (EPRI) concluded that adding cabling in existing nuclear plants costs approximately $2000 per foot (Hashemian 2009b). In addition to cost, over time rust, corrosion, steam, dirt, dust, and water degrade the wires and cause maintenance issues (IAEA, 2008). The extension of older plants' licenses necessitates more instrumentation to monitor age, but installing wired sensors on all the equipment of an aging plant that needs monitoring would be prohibitively expensive (AMS, 2010b). Fortunately, the cost of wireless systems can be less than 1% of the cost of wired systems in a nuclear plant environment. These cabling costs alone represent a substantial incentive for plants to explore wireless systems. Moreover, the wireless industry is aiming to reduce wireless costs from $20/foot to $2/foot over the next few years (AMS, 2010b).
Wireless sensors facilitate difficult measurements in processes where wiring is a weak link, in hazardous environments, and in applications where space for wiring installation is limited. Wireless sensors can also be added as needed, without laying more cabling, and they can be moved from one location to another without having to move wires. Wireless sensors can usually be installed and operational very quickly and offers immediate off/on availability, minimizing communication complexity, promoting system modularity, and facilitating the interconnection of devices within an I&C system. Wirelessly networked devices can be monitored for anomalies and quickly reconfigured (via software) much more easily than wirelined or cabled devices (IAEA, 2008).
Furthermore, with wireless sensors, data can be collected from anywhere and routed on to the Internet where it can be easily accessed and analyzed (Hashemian 2008). The return on investment of wireless systems is often only several months, versus the years that wired/cabled systems require (IAEA, 2008). Wireless technologies do not suffer from a number of critical weaknesses to which wired technologies are susceptible. For example, one

intrinsic benefit to using wireless sensors is that the communication link between the sensor and destination is largely unaffected by moisture. For instance, in a loss-of-coolant accident (LOCA) the containment building of a nuclear reactor can be inundated with water, which can damage sensitive equipment cabling. On the other hand, a wireless sensor would likely be unaffected by this connection issue and continue to provide reliable and important reactor health information throughout the accident and subsequent investigations (AMS, 2010b).

Though wireless technologies do not completely eliminate all wiring needs, they reduce it by one to two orders of magnitude. For example, at Comanche Peak Nuclear Power Station--currently, the largest installation of wireless sensors in the world--more than 10,000 feet of cable were used to develop the foundation for implementing wireless technologies. The wireless infrastructure put in place there provides 100% communications coverage throughout the site and gives the plant the ability to add wireless sensors to monitor and analyze various plant processes and equipment (Hashemian 2009b). This installation has demonstrated that wireless sensor networks can be cost efficient, reliable and secure (IAEA, 2008).

In nuclear power plants, wireless sensors can provide a simple, cost-effective path to improved redundancy without compromising safety. Wired sensors would continue to be designated as the primary element and wireless sensors as a substitute if the wired sensor fails, such as during a LOCA, in which cables become wet or damaged and provide compromised signals (AMS, 2010b).

Many sensor manufacturers have partnered with companies that make wireless transmitters, receivers, and network equipment to produce an integrated network of wireless sensors that can measure process temperature, pressure, vibration, humidity, and other parameters (Hashemian 2008). In addition, wireless community leaders, users, and producers are working on common terminology, a unified platform, and a new standard to facilitate the use of wireless sensors. For example, in 2009 the Instrumentation, Systems, and Automation Society (ISA) approved and released a new standard, referred to as ISA100, to harmonize the use of wireless technologies in industrial applications such as nuclear plants (Hashemian 2008). Including wireless communication capabilities based on a standard protocol such as ISA 100 or IEEE 802.11 in the design plans of the next generation of nuclear power plants can not only provide the necessary means to transmit much-needed sensor data; it can also provide an infrastructure for plant-wide communications (Hashemian 2009b).

Wireless sensors are gaining popularity in plant monitoring in non-nuclear plants and radio frequency identification (RFID)–based sensors, coupled with small-scale, distributed, device-specific "energy harvesting" systems (Hashemian, forthcoming). Though wireless sensors may eventually find their way into nuclear plant process measurement and control, today, they are mainly useful for condition-monitoring applications (Hashemian, 2008). Indeed, on-line condition monitoring (to be discussed later in this chapter) is emerging as the first opportunity for wireless technology to prove itself in the industry (IAEA, 2008).

Because of the potential offered by wireless networking, sensors are rapidly evolving from information devices to communication devices, with substantial implications for the management of security and configuration control in nuclear plants (IAEA, 2008). New wireless sensors from Eaton, Honeywell, General Electric, and others are expected to offer improved reliability and security in monitoring process conditions in real-time or near-real-

time. Not only will they likely find application in nuclear condition monitoring applications; they may even one day be used in nuclear control applications.

Future applications of wireless technologies will include distributing intelligence along the I&C network (which IAEA calls "the convergence of sensing, computation and communication"), thereby reducing the need for high data rates along wireless links, and reductions in sensor size and power requirements (IAEA, 2008). Already, the author is working with the Department of Energy on a project to extend wireless sensors and networks inside the reactor containment for equipment condition monitoring, auxiliary measurements during plant outages, and improved capability for post accident monitoring of the plant. Phase III of this project, which started in the fall of 2010, is designing and qualifying a wireless sensor network for use in the reactor containment building of nuclear power plants, where wiring costs can be as high as $50,000 per foot (Analysis and Measurement Services Corp. [AMS], 2010b).

5. Challenges of digital I&C

Although digital I&C technology has been successfully applied outside the U.S., the U.S. nuclear power industry has been slow to adopt digital I&C, and even then mostly for only non-safety-related applications, such as feedwater control systems, recirculation control systems, demineralizer control systems, main turbine controls, etc. (U.S. NRC, 2011; IAEA, 1999; Hashemian, 2009a). This is largely the result of regulatory concerns over the unique question marks raised by digital I&C technology (Hashemian 2009a).

One critical concern—and the primary reason why digital instrumentation is subject to stringent licensing requirements for use in process safety systems (Lipták, 2006)--is digital I&C's dependency on software. Although analog I&C may have higher overall failure rates, its failure mechanisms and modes are perceived as better understood and more easily reproducible (ORNL, 2007). Repeatability gives confidence that periodic testing can minimize future failures. In contrast, software programs' high number of discrete logic steps and inputs and algorithmic complexity means that I&C programs could potentially generate a unique, potentially infinite range of operating characteristics. To verify the reliability of such systems would require testing each line of code for every conceivable combination of inputs and at all possible rates of change—a monumental task (IAEA, n.d.; European Nuclear Agency [ENA], 2008). As a concrete example, in 2009 the UK Nuclear Installations Inspectorate reviewed the European Pressurized Reactor I&C architecture developed by AREVA and EDF and concluded that it "appears overly complex" and contains too many connections with less safety-critical systems (Hirsch, 2009).

Common mode failure—failures resulting from errors or 'bugs' shared by identical software programs running on multiple I&C systems--is a second concern stemming from digital I&C's dependence on software (Lipták, 2006). Specifically, calibration errors, errors in generating setpoints, and hardware and sensor failures are the types of common mode failure most feared from shared flaws in I&C software (Bickel, 2009). According to the U.S. Nuclear Regulatory Commission, in the past twenty years, 38 of about 100 operating plants have reported "potential and actual" common-mode failures, some affecting single plants, but others affecting multiple plants using the same digital system (U.S. NRC, 2011). The more software is integrated into every layer of I&C—from large platform computer systems and microprocessor-driven control systems to software embedded in primary instrumentation and controllers—the greater the potential challenge posed by common mode failure (IAEA, 1999).

A second challenge posed by digital I&C is cyber security--the protection of data and systems in a network, both wired and wireless, from unauthorized access or attack, whether from business espionage, technology theft, or disgruntled employee interference or from recreational hacking, cyber activism, or the probing of a foreign state or terrorist organization. Wireless is the least secure of the physical layers (IAEA, 2008). Wireless transmissions are inherently open, meaning that access can potentially be obtained anywhere within the transmission zone, so they are more vulnerable to such intrusions and threats as non-directed, damaging attacks by software viruses and worms; data network nonperformance from denial-of-service attacks and network spoofing; loss of data privacy and confidentiality from eavesdropping and network packet sniffing; and directed threats involving network packet modification, mimicking, and data tampering (Hashemian, 2009b; AMS, 2010b). These threats can generally be grouped into four categories: loss of confidentiality (unauthorized access to data), loss of integrity (data or software/hardware changed by the intrusion), loss of availability (data transmissions interrupted or systems shut down), and loss of reliability (potential changes made to I&C data systems or computers) (IAEA, n.d.).

There are two major cybersecurity concerns related to the use of wireless technologies in nuclear power plants: being able to satisfy regulatory requirements and employing sufficiently robust methodologies to protect data transmissions across wireless networks (e.g., encryption, authentication, intrusion prevention) (AMS, 2010b).

A final challenge posed to digital I&C is electromagnetic and/or radio frequency interference (EMI/RFI). For wireless devices to be safely used in nuclear power plants, they must first be deemed electromagnetically compatible with the surrounding environment. A device is said to have electromagnetic compatibility (EMC) if it does not interfere with surrounding electronics and is not itself susceptible to interference from the other devices (AMS, 2010b). Aside from the EMI/RFI effects of wireless devices on surrounding plant equipment and vice versa, EMI/RFI issues can also exist between wireless devices.

In industrial applications, most interference results from intermittent bursts of narrow-band signals, random electromagnetic interference (e.g., background noise) and deterministic EMI (e.g., radio stations; AMS, 2010b). The sources of EMI are many and varied, ranging from welders to managers with radio sets (IAEA, 1999). The range and fidelity of wireless signals can also be influenced by implementation issues such as multipath and signal attenuation resulting from proximity to metallic structures, which can limit deployment (IAEA, 2008; AMS, 2010a). Although, EMI/RFI issues have largely been addressed with respect to implementing wireless sensors and networks for equipment condition monitoring in nuclear plants, using wireless for equipment or process control is another matter. Much more secure EMI/RFI safeguards are required for wireless to find use in safety or control applications, which is why NRC standards specifically prohibit the use of wireless technology on "critical digital assets" (AMS, 2010b; Hashemian, 2008).

5.1 Addressing the challenges posed by digital I&C

The challenges posed by the application of digital and wireless I&C in nuclear power plants can partly be addressed by continued application of the nuclear power community's longstanding "defense-in-depth" strategy. This strategy's basic principle is that safety risks can be met by designing in multiple, distributed barriers and layers in I&C systems so that no abnormal event, error, or failure—external, electronic or mechanical, or human--can completely interrupt the system's functioning (IAEA, 1999).

Broadly speaking, defense in depth takes three different forms: diversity of components, redundancy of components, and independence of components. Diversity can involve design diversity (the use of different technologies such as digital versus analog or different architectures, etc.), equipment diversity (the use of different equipment manufacturers, different equipment versions, etc.), functional diversity (applying different mechanisms such as rod insertion versus boron injection or different response times), human diversity (the use of different designers, engineers, programmers, testers), signal diversity (relying on different process parameters sensed by different physical effects or sensor types), and software diversity (different algorithms, logic, programming languages) (U.S. NRC, 2011; Hashemian 2009a). Such diversity diminishes the likelihood that an error or failure in one I&C element will be duplicated in another. As such, diversity is a specific form of protection against common mode failure in I&C software and systems (IAEA, 1999).

The redundancy aspect of defense in depth complements the diversity aspect: not only are diverse systems available to perform functions should one system fail, but multiple components of the same systems are also available. If one component fails, an identical component is available to take its place (IAEA, n.d.; IAEA, 1999).

The third aspect of defense in depth, independence or separation, minimizes the risk of I&C failure by ensuring that each element in an I&C system is truly independent of the others, through, for example, electrical isolation, physical separation (e.g., barriers, distance), and/or independence of system intercommunication (IAEA, n.d.; IAEA, 1999; Hirsch, 2009).

The Fukushima Daiichi emergency of 2011 illustrates the principle and limits of the defense-in-depth strategy. The Tokyo Electric Power Co. (TEPCO), the plant's operator, believed it had sufficient *diversity* of electrical supply to provide the plant with ongoing electrical power during an emergency: it had primary electrical supply from TEPCO's regional grid, it had backup generators in case grid power failed, and it had 8-hour emergency batteries in case the generators failed. However, in one stroke the earthquake and tsunami knocked out the primary grid power and rendered the generators unusable. The backup batteries worked but not long enough to enable TEPCO to reinstitute continuous power to prevent a LOCA. In other words, the plant's electrical plan lacked true *independence* (both grid and backup generators were knocked out by the same factor, the tsunami) and true *redundancy* (no second-line generators or batteries were available to replace the first-line-of-defense generators and batteries).

Of course, adding diversity, redundancy, and independence also increases a system's complexity, expanding, in other words, the range of possible error or failure scenarios that plants must track. The nuclear power community has attempted to address the complexity issue through stringent regulation of proposed new I&C, by requiring the use of hardware and software I&C components that have been thoroughly verified and validated for nuclear plant environments (IAEA, n.d.), and by requiring that the complexity of I&C components be graded such that a safety-essential I&C component may have only limited, specific functionality to ensure that it will more reliably perform its design task. (Thus, I&C elements controlling non-safety tasks are allowed to have more complexity since less is at stake should that complexity produce unanticipated errors or failures) (IAEA, 1999).

For example, field-programmable gate array (FPGA) technology has emerged as an answer to the risks posed by overly complex I&C software. An FPGA is a device made up of thousands or millions of logic gates on integrated circuit chips that can be programmed after manufacture by the customer to perform various tasks, ranging from simple logic

operations to complex mathematical functions (U.S. NRC, 2011; Hashemian, 2009a). Because, once programmed, an FPGA executes only that program repetitively and link only the functions needed for a given I&C application, they are substantially simpler than microprocessors, minimizing the risk posed by complexity (U.S. NRC, 2011).

The cybersecurity concerns posed by the use of wireless in nuclear plant I&C are being addressed by the application of experience gained in military, national security, banking, and air-traffic sectors (IAEA, n.d.). On a technological level, intrusion detection, virus scanning, and encryption tools can identify and block cyber threats. In technical terms, security in wireless is no different from security in wired infrastructure. Wireless can be made more secure than wired by including security in the physical layer, thus providing no access to record or tap into the bit stream (AMS, 2010b). On an administrative level, security zones, security management systems, passwords and biometric identification can limit cybersecurity concerns (IAEA, n.d.). At least four sets of standards are relevant to cybersecurity in nuclear applications: IEC Security Standards ISO/IEC 27000 series, IEEE P1711 and IEEE P1689 for Cyber Security of Serial SCADA Links, ISA99 Security for Industrial Automation and Control Systems, and North American Electric Reliability Corporation (NERC) Critical Infrastructure Protection (CIP) standards CIP-002 through CIP-009 (Hashemian, 2009b).

Finally, concerns over EMI/RFI arose partly because many nuclear plants discovered their security personnel's 'walkie-talkie' radios inadvertently affected plant systems. As a result, they established exclusion zones for such wireless devices around sensitive or critical equipment. However, the radios typically used by plant security personnel transmit at a much higher power level (several watts) than do wireless sensor technologies and in the megahertz (MHz) region. Wireless systems' operate at the 100 milliwatts (mW) power level and in the gigahertz (GHz) range of frequencies. In general, modern wireless devices' lower power and higher frequency levels significantly decrease the chances of interference with nuclear power reactor equipment (AMS, 2010a). Moreover in new plants, the plant EMI/RFI design should allow for other wireless sensor networks to be deployed side-by-side for various applications. This will enable the wireless sensors from various manufacturers to be used in the plant without interference (Hashemian, 2009b).

Recent R&D work performed by the author under a Department of Energy Small Business Innovation Research grant has demonstrated that concerns such as cyber security, EMI, and wireless signal impact on plant equipment can be easily managed. Wireless technology can be implemented successfully and practically in industrial nuclear power plants for condition monitoring of safety-related equipment (AMS, 2010b). However, although wireless sensors and networks are well suited for equipment condition monitoring in nuclear power plants, they are not yet ready for control applications nor is it yet safe to attempt to use wireless sensors for equipment or process control. A hacker cannot cause much damage through wireless technologies used for condition monitoring, but he/she can cause problems in control (AMS, 2010b). The full application of digital I&C to safety-essential control will depend on further advances in nuclear plant I&C design, technology, and regulation.

6. On-line monitoring

The evolution of digital I&C is making possible the development of holistic, integrated systems for automatically verifying the performance of I&C sensors and assessing the health of nuclear power plant equipment and processes while the plant is operating. These so-

called online condition monitoring (OLM) systems can be used for on-line I&C maintenance, predictive maintenance, and troubleshooting of reactor components, aging equipment, and to support life extension objectives (EPRI 2008). OLM can be used in PWRs, BWRs, and other reactor types. The system can be built into the design of new plants or deployed as an add-on feature to the existing generation of plants (Hashemian, 2009b).

Applications that can be performed using OLM include in-situ response-time testing of process instrumentation; instrument calibration monitoring; cross-correlation flow measurement; online detection of venturi fouling; online detection of sensing-line blockages, voids, and leaks; fluid and gas leak detection; equipment and process condition monitoring; core barrel vibration measurement; online measurement of temperature coefficient of reactivity; aging management of neutron detectors and core exit thermocouples; and measurement of vibration of in-core flux monitors, core flow monitoring, or N-16 flow measurement.

One of the important applications of OLM is in monitoring the performance of pressure, level, and flow transmitters (AMS, 2010b). In the simplest implementation, redundant channels are monitored by comparing the indicated measurement of each individual channel to a calculated best estimate of the actual process value. Each channel's calibration status can be made by monitoring each channel's deviation from the calculated best estimate (IAEA, 2008). Figure 2 shows the data acquisition signal path for an OLM system.

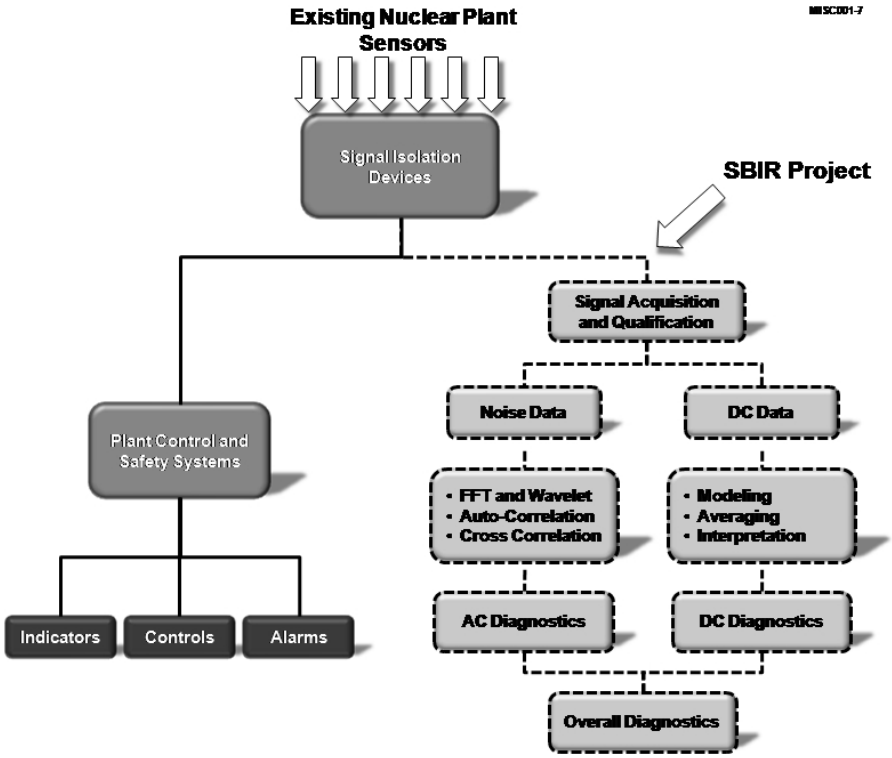


Fig. 2. Data Acquisition Signal Path for OLM

An OLM system is made up of a data acquisition module involving hardware and software and a data processing module involving software implemented on a fast computer. The data acquisition module includes signal isolation devices as well as fast sampling capabilities (e.g., 1000 Hz). If the data is sampled fast, it can be used for both calibration verification by DC signal analysis using averaging and modeling techniques, and response- time testing by AC signal analysis using the noise analysis technique (Hashemian, 2009b). Dynamic analysis of nuclear plant sensors and equipment uses AC signal analysis to determine how sensors and equipment react to fast-changing events such as temperature or pressure steps, ramps, spikes, etc. (Hashemian, 2009b).

OLM originated from reviews of equipment performance data from a variety of industries. These reviews showed that a majority of process equipment performs well for long periods of time and that frequent hands-on maintenance is not needed and is sometimes counterproductive. For example, research performed by Emerson Company's Rosemount Division--which manufactures process sensors such as pressure, level, and flow transmitters for a variety of industries--has shown that these sensors perform well for periods of ten to twenty years and need little hands-on maintenance (Hashemian, 2008). Based on such research, it is now known that over 70% of maintenance work on pressure transmitters in industrial processes does not reveal problems, and maintenance intervals can therefore be extended (AMS, 2010b).

Building on such findings, over the past twenty years, the nuclear power community has made substantial strides to establish OLM technologies in the industry. Numerous academic, government, and industry institutions (as well as private companies) have sponsored R&D efforts in this area. As a result, the feasibility of OLM technologies has been successfully demonstrated for implementation in the existing nuclear fleet (AMS, 2010b).

Moreover, the NRC has approved the OLM concept for in-situ determination of the calibration status of pressure, level, and flow transmitters in nuclear power plants. That is, nuclear power plants can use OLM to establish when a pressure, level, or flow transmitter must be calibrated.

Although OLM provides substantial benefits to the safety and economy of nuclear power plants and has been approved by the NRC, OLM use in nuclear power plants is fragmented and sporadic (Hashemian, 2009b; Electric Power Research Institute [EPRI], 2008; AMS, 2010b). For example, the Sizewell B plant in the United Kingdom was able to extend the calibration intervals of 70% of the transmitters that were eligible for calibration extension using OLM techniques. Similarly, as noted, TXU Comanche Peak nuclear power plant currently has the largest installation of wireless sensors in the world in its $14 million wireless network. Most of existing nuclear power plants have the capabilities and equipment needed for implementing many of the OLM technologies. However, for most plants, these capabilities are not used to their fullest extent for OLM applications (Hashemian, 2009b).

One reason is that the implementation of OLM techniques depends on the availability of data from a large network of sensors deployed on equipment such as motors, fans, pumps, etc. While many nuclear power plants have an OLM or predictive maintenance program for equipment outside of their containments, none have OLM programs for equipment inside the containments due to the sensor wiring costs and penetration space limitations (AMS, 2010b). Today, a majority of industrial equipment does not benefit from OLM technologies

partly because no sensors exist to provide the necessary data, and installing wired sensors is often cost prohibitive and impractical.
Wireless sensors will help fill this gap, enabling condition-monitoring technologies to flourish (Hashemian, 2008). As a result, wireless sensors promise to experience explosive growth over the next decade in OLM. Incorporating a wireless infrastructure will help new plants to provide the necessary means of communicating OLM data to plant engineers at low cost, and provide a means for the future expansion of OLM capabilities (Hashemian, 2009b). Inevitably, research in OLM methods will continue, and there will be a need to measure and analyze parameters that are not being considered now.
The application of wireless sensors for equipment condition monitoring in industrial processes has left open a critical gap in the handling of data from wireless sensors, in the guidelines that define which parameters must be measured, in the type and number of sensors to be deployed for measuring these parameters, and in the methods for ensuring that optimum data is gathered to monitor the health and condition of various equipment.
Furthermore, over the next few years, the use of wireless sensors will generate an enormous amount of data from industrial processes. Although much thought has been focused on developing wireless sensors, little or no effort has been expended on data qualification and data processing techniques for these sensors. Moreover, little effort has been spent in determining the type of parameters that should be measured and what the correlation should be between these parameters and the actual condition of the equipment being monitored (Hashemian, 2008).
In the next generation of reactors OLM systems should be built into the design so as to provide automated measurements, condition monitoring, and diagnostics to contribute to optimized maintenance of the plant (Hashemian, 2009b). Reactor designs for next-generation plants will typically incorporate an integrated digital infrastructure including highly integrated control rooms, fault-tolerant control systems, and monitoring systems with large amounts of available information and data. Most of these digital systems will be designed to monitor their own performance continuously, self-correct for identified changes, and function more reliably than previous designs (Hashemian, 2009b).
To develop OLM for future needs, considerations will be needed for increased availability of process sensor data in the plant computer, higher sampling frequency and resolution data acquisition capabilities, increased redundancy for critical process sensors, and more flexible infrastructure to accommodate future data acquisition needs. Utilities will have to adapt to continuous 24-hour monitoring of instrumentation (AMS, 2010b).

7. Conclusion

Today, OLM technologies and techniques have evolved to the point where in many cases equipment failures and/or maintenance needs can be adequately predicted days, weeks, or even months in advance of a system or equipment failure (AMS, 2010b). In general, a wireless system provides the lowest overall cost for large-scale OLM applications (IAEA, 2008). In the years ahead, future I&C will be fully digital (software based), distributed, bus connected, amenable to OLM, and qualified to industrial standards (IAEA, 1999).

8. References

Analysis and Measurement Services Corp. (February 2010a). "On-Line Monitoring of Accuracy and Reliability of Instrumentation and Health of Nuclear Power Plants," Final Project Report, DOE Grant No. DE-FG02-06ER84626.

Analysis and Measurement Services Corp. (November 2010b). "Implementation of Wireless Sensors for Equipment Condition Monitoring in Nuclear Power Plants, SBIR Phase II Final Report, DOE Grant No.: DE-FG02-07ER84684.

Bickel, J. (December 11, 2009). "Digital I&C Is Safe Enough," *Nuclear Engineering International*.

Electric Power Research Institute. (November 2008). "Requirements for On-Line Monitoring in Nuclear Power Plants," Final Report, EPRI, Palo Alto.

European Nuclear Agency. (July 2008). "Inspection of Digital I&C Systems – Methods and Approaches," Proceedings of a CNRA Workshop, Garching, Germany, 24-26 September 2007, OECD ENA.

Hashemian, H.M. (1999). "Advanced Sensor & New I&C Maintenance Advanced Sensor and New I&C Maintenance Technologies for Nuclear Power Plants," Paper presented at POWID conference, International Society of Automation.

Hashemian, H.M. (2006). *Maintenance of Process Instrumentation in Nuclear Power Plants*, Springer Verlag, ISBN 978-3-642-07027-3Berlin, Heidelberg.

Hashemian, H.M. (2008). Predictive Maintenance of Critical Equipment in Industrial Processes, dissertation for Lamar University.

Hashemian, H.M. (2009a). "State of the Art in Nuclear Power Plant I&C," *International Journal of Nuclear Energy Science and Technology*, Volume 4, No. 4, page 330-354.

Hashemian, H.M. (2009b). On-Line Monitoring Applications in Nuclear Power Plants, doctoral dissertation, Chalmers University of Technology.

Hashemian, H.M. (Forthcoming). "Sensors for Next-Generation Nuclear Plants: Fiber-Optic and Wireless," *Nuclear Science and Engineering*.

Hirsch, Dr. H. (November 05, 2009). "Statement on the Separation of Safety I&C and Operational I&C: Expanded Version," Greenpeace.org.

Hurst, T. (January 2007). "Tow nuclear power I&C out of the 'digital ditch,'" *Power* magazine.

International Atomic Energy Agency. (1999). Modern Instrumentation and Control for Nuclear Power Plants: A Guidebook, IAEA, Vienna.

International Atomic Energy Agency. (2008). *On-Line Monitoring for Improving Performance of Nuclear Power Plants, Part 1: Instrument Channel Monitoring*, NP-T-1.2, and *Part 2: Process and Component Condition Monitoring and Diagnostics*, NP-T-1.2, IAEA, Vienna.

International Atomic Energy Agency. (n.d.). "Instrumentation and Control (I&C) Systems in Nuclear Power Plants: A Time of Transition," IAEA, Vienna.

Lipták, Béla G. (2006). "Safety Instrumentation & Justification of Its Cost," *Instrument Engineers' Handbook*, 4th ed., Taylor & Francis, ISBN o-8493-1081-4, Boca Raton, FL.

Oak Ridge National Laboratory. (May 2007). "Industry Survey of Digital I&C Failures," ORNL, Oak Ridge, TN.

U.S. Nuclear Regulatory Commission. (2011). "Digital Instrumentation and Controls," U.S. NRC, Washington, DC.

4

Design Considerations for the Implementation of a Mobile IP Telephony System in a Nuclear Power Plant

J. García-Hernández[1], J. C. Velázquez- Hernández[1],
C. F. García-Hernández[1] and M. A. Vallejo-Alarcón[2]
[1]*Electric Research Institute (IIE)*
[2]*Federal Commission of Electricity (CFE)*
Mexico

1. Introduction

IP telephony, also called voice over Internet protocol (VoIP), is rapidly becoming a familiar term and technology that is implementing in the enterprise, education, government organizations and industry. Mobile IP telephony is the new generation of communications networks that makes possible the convergence of voice and data over wireless local area networks (WLANs). This technology combines data networks with mobile technologies to support voice and data applications over a common integrated network. A key advantage of IP telephony is that it allows the transmission of voice signals from conventional telephones over an IP data network, being either a public network (Internet) or a private network (Intranet). Figure 1 shows a general IP telephony system. IP telephony is designed to replace the legacy TDM (time division multiplexing) technologies and networks by an IP-based data network. Digitized voice will be carried in IP data packets over a LAN and/or WAN network. A major aspect involved in a voice conversation using mobile IP telephony is the conversion of analog or digital voice signals from conventional phones to IP packets for further transmission either to a fixed or mobile phone, over an IP network. One of the most important recommendations that can be made is to pay close attention to the infrastructure

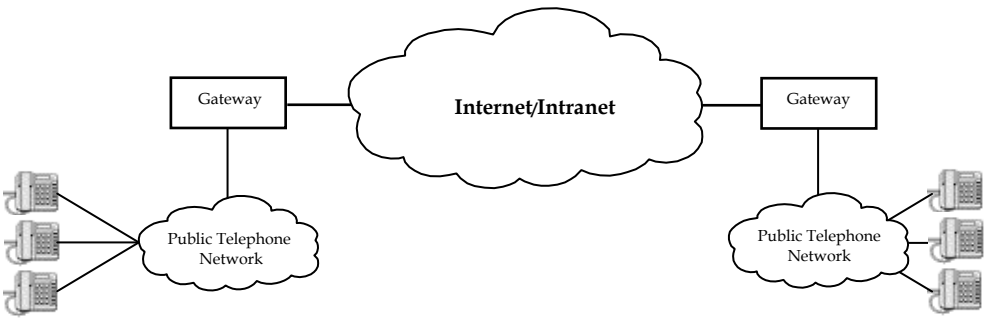


Fig. 1. General IP telephony system

that the IP telephony network is built on. The design considerations must be solid otherwise; there will be ongoing quality issues until the network design issues are resolved. Even though IP telephony have made some vast reliability and quality improvements in recent years, customers and network designers still struggle with implementing the technology in a multi-vendor network. There are many reasons for this, such as interoperability issues and proprietary protocols. In addition, the use of new wireless technologies in nuclear power plants is growing fast. The WLAN technology based on the IEEE 802.11 standard has a very promising future for its use in nuclear power plants, due to features like mobility, reliability, security, scalability and compatibility with other technologies in order to provide new services such as voice over IP (VoIP) and IP video (Shankar, 2003).
In this work, the design considerations for the implementation of a mobile IP telephony system in a nuclear power plant based on national and international standards are presented, as well as, technical requirements that commercially available equipment must meet. In addition, this work gives an analysis of the most relevant wireless technologies currently available that can be implemented in nuclear power plants and also identifies nuclear regulatory guidelines, wireless networks standards, electromagnetic and radio-frequency interference standards. In the next sections, an analysis of the most relevant wireless technologies currently available that can be implemented in nuclear power plants is presented.

2. Wireless LAN standards

The Institute of Electrical and Electronics Engineers (IEEE) has been produced a series of standards for wireless networks referred as 802.11x, for wireless LAN (Local Area Networks). The original standard used to implement wireless LANs was 802.11 (IEEE, 1999a). It was first published in 1999 and designed to support a maximum data rate of 2 Mbps in the 2.4 GHz band. This standard uses two modulation techniques: frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS). Figure 2 shows the IEEE 802.11 standard architecture.

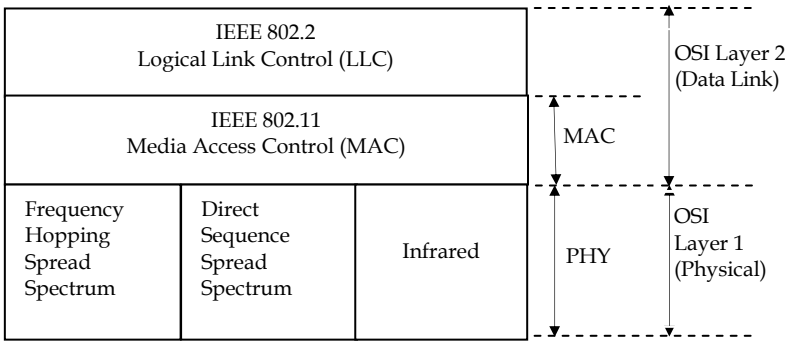


Fig. 2. The IEEE 802.11 Standard Architecture

2.1 The IEEE 802.11b standard

Wireless Fidelity (Wi-Fi) networks are intended to be implemented more in enterprises and in industry. Wi-Fi is commonly used as the abbreviation of 802.11b standard. It supports

bandwidth up to 11 Mbps, comparable to traditional Ethernet. The 802.11b standard also uses DSSS as modulation technique in the 2.4 GHz band as the original 802.11 standard, where Wi-Fi devices communicate to each other at data rates up to 11 Mbps (IEEE, 1999b). In case of any radiofrequency or interference signal cause errors in data transmission, Wi-Fi devices will automatically reduce their data rate to 5.5 Mbps, to 2Mbps and finally to 1 Mbps (Martínez, 2002). These occasional reductions in the data rate are aimed to keep the network very stable and highly reliable. Some advantages of Wi-Fi standard are: high data rates, reliable, wide signal ranges and supports of all 802.11 devices using DSSS.

2.2 Other IEEE 802.11x standards

Along with the 802.11, other standards have been developed (Martínez, 2002): 802.11a, 802.11g, 802.11e and 802.11i. The 802.11a standard (IEEE, 1999c) was published in 1999 and uses Orthogonal Frequency Division Multiplexing (OFDM) to support bandwidth up to 54 Mbps and signals in the unlicensed frequency spectrum around 5.8 GHz. This higher frequency compared to 802.11b limits the range of 802.11a networks. The higher frequency also means 802.11a signals have more difficulty penetrating walls and other obstructions. To solve this problem and to cover a similar range that 802.11b, a greater number of access points must be installed. Because 802.11a and 802.11b utilize different frequencies, the two technologies are incompatible with each other. The 802.11g standard (IEEE, 2003) attempts to combine the best of both 802.11a and 802.11b standards. It uses OFDM to support bandwidth up to 54 Mbps, using the 2.4 GHz band for greater range. 802.11g is backwards compatible with 802.11b, meaning that 802.11g access points will work with 802.11b wireless network adapters and vice versa. Also, 802.11g produces less multipath losses than 802.11a. The 802.11e standard (IEEE, 2005) has been developed to support applications with quality of service (QoS) requirements. It supports a wide range of real-time applications such as voice, audio, video and data over WLAN. It adopts time scheduling and polling mechanisms to cope with delay sensitive traffic. The 802.11i standard (IEEE, 2004) has been developed to improve the security of data provided by Wired Equivalent Privacy (WEP) protocol in 802.11b. 802.11i incorporates an entirely new privacy algorithm and authentication mechanism based on the Advanced Encryption Standard (AES).

3. Electromagnetic interference standards

Electromagnetic interference (EMI), radio-frequency interference (RFI), and power surges have been identified in (NUREG, 2003) as environmental conditions that can affect the performance of safety-related electrical equipment in a nuclear power plant. A series of comprehensive commercial EMI/RFI immunity standards have been issued by the following international organizations:

- International Electrotechnical Commission (IEC)
- European Committee for Electrotechnical Standardization (CENELEC)
- International Special Committee on Radio Interference (CISPR)

These international organizations have produced standards for general application. National organizations in countries like United States of America, Canada, Australia and Europe have their own standards that regulate EMI/RFI immunity of electronic equipment. In the U.S., the Nuclear Regulatory Commission (NRC) has produced the most comprehensive guide known as Regulatory Guide 1.180 (NRC, 2003) with the aim of

developing the technical basis for regulatory guidance to address electromagnetic interference, radio-frequency interference, and surge withstand capability in safety-related instrumentation and control systems in a nuclear power plant.

3.1 Regulatory guide 1.180

This regulatory guide endorses design, installation, and testing practices acceptable to the NRC staff for addressing the effects of EMI/RFI and power surges on safety-related instrumentation and control systems in a nuclear power plant environment. This regulatory guide is based on the standards: IEEE 10-50 (IEEE, 1996), MIL-STD-461E (DoD, 1999) and IEC 61000 series (IEC, 2002). For instance, the design and installation practices described in the IEEE 1050 standard are endorsed for limiting EMI/RFI subject to the conditions stated in the Regulatory Position. Electromagnetic compatibility (EMC) testing practices from military and commercial standards are endorsed to address electromagnetic emissions, EMI/RFI susceptibility, and power surge withstand capability. The MIL-STD-461E standard contains test practices that can be applied to characterize EMI/RFI emissions. In addition, selected EMI/RFI test methods from MIL-STD-461E and the IEC 61000 series are endorsed to evaluate conducted and radiated EMI/RFI phenomena for safety-related control and instrumentation systems. The IEC electromagnetic compatibility (EMC) standards include IEC 61000-3 (part 3: Limits), IEC 61000-4 (part 4: Testing and measurement techniques), and IEC 61000-6 (part 6: Generic standards). This regulatory guide also endorses electromagnetic operating envelopes corresponding to the MIL-STD-461E test methods. These operating envelopes were tailored from the MIL-STD-461E test limits to represent the characteristic electromagnetic environment in key locations at nuclear power plants. Comparable operating envelopes for the IEC 61000 test methods are also endorsed. The Regulatory Guide 1.180 was updated in the year 2003 to provide additional acceptable methods for validating the performance of instrumentation and control systems and includes guidance on testing to address signal line susceptibility and very high frequency (>1GHz) phenomena.

4. Wireless LAN technologies

The use of wireless technologies in industrial and power utility environments, including their use in nuclear power plants is growing fast. The wireless technologies that have been implemented in nuclear power plants include paging systems, digital mobile radio and cellular systems. Currently, WLAN technology is been installing and evaluating in nuclear power plants, due to it provides enhanced features compared to traditional wireless communications technologies such as conventional mobile radio in two key aspects:

- Higher operation frequencies
- Lower output power

WLAN technology is based on the 802.11 standard and generally operates at higher frequencies (2.4/5 GHz) and at significantly lower effective power output levels (20 mW-1W) than UHF/VHF (890/450 MHz) communication systems operating at power output levels between 4-5 W. This feature of more modem wireless devices, including WLANs, general1y requires that the end user be closer to a potentially sensitive device before interference is noted. So, it can be inferred that modem wireless devices are less of a threat and less like1y to interfere with nuclear power plant equipment than older devices; that is,

more modern devices tend to be less intrusive. Those features make that the electromagnetic interference generated by devices based on WLAN technology does not affect significantly to safety-related instrumentation and control equipment (EPRI, 2003). In 2002, the Electric Research Institute (EPRI) published a report in which EPRI developed guidelines for the use of wireless technologies in nuclear power plants (EPRI, 2002). The purpose of this report was the evaluation of wireless technologies in nuclear power plants for integrated (voice, data and video) communication, remote equipment and system monitoring, and to complement an electronic procedures support system. The guidelines effort focuses on the development of a rules structure to support the deployment of wireless devices in a nuclear power plant without compromising continuous, safe, and reliable operation. For these reasons, spread spectrum appears to be the most adequate technology for the nuclear power environment.

4.1 Frequency spectrum regulation

The Federal Communications Commission (FCC), organism that manages and regulates the electromagnetic spectrum in U.S. assigned in 1985 to ISM band the 900 MHz, 2.4 GHz and 5.8 GHz frequency ranges. These regulations are specified in the CFR-47 section 15.2.47 (FCC, 2004). The ISM band is a license-free band and it is used by WLAN technology. In Mexico, the Consultative Committee for Standardization in Telecommunications (CCNNT) manages and regulates the electromagnetic spectrum (CCNNT, 2001). Table 1, shows the frequency ranges and bandwidth reserved for the ISM band for their use the U.S. and Mexico.

ISM Band	Bandwidth
902 – 928 MHz	26 MHz
2.4 – 2.4835 GHz	83.5 MHz
5.725 – 5.850 GHz	125 MHz

Table 1. Reserved frequencies for the ISM band

The FCC set these ISM bands for license-free and low power radio transmission over short to medium distances (Meel, 1999). The FCC requires that the signal be distributed over a wide swath of bandwidth using a spread spectrum technology originally developed by the military for anti-jamming applications. Wireless devices that operate in these license-free bands can allow immediate, real-time commissioning of a network, avoiding the delays associated with installing wiring or cables. By spreading data transmissions across the available frequency band in a prearranged scheme, spread spectrum encoding technology makes the signal less vulnerable to noise, interference, and snooping. The significant amount of metal often found in industrial settings can cause signals sent over a single frequency to bounce and cancel other signals arriving at the same time. Spread spectrum technology helps overcome this problem and allows multiple users to share a frequency band with minimal interference from other users.

4.2 Radiated power regulation

The power level radiated by an antenna in a WLAN is specified by the rules stated in the FCC section 15.247 for operation in the U.S. (FCC, 2004). The FCC also limits the increase in

the output power of the antenna to a maximum of 6 *dBi*. So, the radiated power is limited to a maximum of 1W for the 2.4 GHz band. In Mexico it is limited to a maximum of 650 mW (CCNNT, 2001). In Europe and Japan, the radiated power levels are different to those allowed for U.S. and Mexico. Table 2, shows the maximum radiated power levels allowed for different countries (Meel, 1999).

Maximum Transmit Power	Geographical Location	Compliance Document
1000 mW	U.S.A.	FCC 15.247
650 mW	Mexico	CCNN-T
100 mW	Europe	ETS 300-328
10 mW/MHz	Japan	MPT ordinance for Regulating Radio Equipment, article 49-20

Table 2. Maximum transmit power

4.3 Bandwidth regulation

Data throughput is adversely affected by distance and the amount of noise or interference in the area. If too many wireless devices are operating in the same vicinity, they can interfere with each other, restricting network capacity. In terms of protection from interference, the FCC and CCNN-T specify that WLANs operating in the three ISM bands, use spread spectrum (SS) as the encoding technique to comply with regulation requirements (Meel, 1999; DoE, 2002; Pearce, 2001). Spread spectrum technology is based on two interference avoidance techniques: frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS). Both modulation schemes have been defined in (IEEE, 1999b) to operate in the 2.4 GHz band, using a bandwidth of 83 MHz (from 2.400 GHz to 2.4835 GHz). Also, the CCNNT specifies that the bandwidth of the transmitted signal depending on the modulation scheme (FHSSS or DSSS) employed (CCNNT, 2001), as shown in table 3. FHSS technique permits the fast movement or "hopping" to any channel within the total allocated spectrum. Here, the carrier frequency hops from channel to channel in some pre-arranged sequence. The major drawback to this technique is a limited data rate.

SS Modulation scheme	Bandwidth
Frequency Hopping (FHSS)	1 MHz (maximum)
Direct Sequence (DSSS)	500 KHz (minimum)

Table 3. Bandwidth of transmitted signal

By contrast, DSSS technique provides much higher data rates. Here, the carrier frequency does not jump from frequency to frequency, but instead spreads the information across a much wider bandwidth. Also, DSSS can provide many users to be on the same channel at the same time and be distinguished from each other by a digital code. DSSS uses the phase shift keying modulation technique known as differential BPSK (DBPSK) and differential BPSK (DQPSK). FHSS uses the frequency shift keying modulation technique known as 2

and 4 level Gaussian FSK (2GFSK and 4GFSK). The data rate for both FHSS and DSSS schemes is defined in (IEEE, 1999a; IEEE, 1999b). Table 4, shows the maximum data rates for the physical layer supported by the most commonly used 802.11 standards. The most common problem when applying an unlicensed wireless system in an industrial environment is radio frequency interference or better known as "radio noise". Both FHSS and DSSS handle noise differently and can have certain advantages depending on the type of interference experienced.

Standard	Frequency band (GHz)	Modulation scheme	Maximum bit rate
IEEE 802.11	2.4 - 2.4835	FHSS	2 Mbps
IEEE 802.11	2.4 - 2.4835	DSSS	2 Mbps
IEEE 802.11b	2.4 - 2.4835	DSSS	11 Mbps

Table 4. Maximum data rates for WLANs

5. Use of wireless LANs in the nuclear environment

The WLAN technology based on the IEEE 802.11 standard has a very promising future for its use in nuclear power plants due to features like mobility, reliability, security, scalability and compatibility with other communication networks technologies in order to provide new services such as voice over IP (VoIP) and IP video (Shankar, 2003). A key issue of these technologies is that no wires are needed to implement new services. Currently, WLAN technology is been installing and evaluating in nuclear power plants, due to it provides enhanced features compared to traditional wireless communications technologies such as conventional mobile radio in two key aspects: higher operation frequencies and lower output power which translates in very high data rates and very low electromagnetic interference. However, wireless technology may exhibit greater vulnerability to the nuclear power plant EMI/RFI environment than existing instrumentation and control systems. The typical environment in a nuclear power plant includes many sources of electromagnetic interference (EMI), radio-frequency interference (RFI), and power surges, such as hand-held two-way radios, arc welders, switching of large inductive loads, high fault currents, and high-energy fast transients associated with switching at the generator or transmission voltage levels. Hence, operational and functional issues related to safety in the nuclear power plant environment are required to address the possibility of troubles and malfunctions in instrumentation and control systems caused by EMI/RFI and power surges.

The wireless communications technologies that can be implemented in nuclear power plants include paging systems, mobile radio, cellular systems and wireless local area networks (WLANs). Recently, there has been a significant increase in the use of wireless technology in nuclear power plants. For example, wireless technology has been installed and evaluated in nuclear power plants by utility companies throughout the U.S., Canada, Mexico, Europe, and worldwide (Bahavnani, 2001; Telrad Connegy, 2001; Wireless Magazine, 1995). Wireless technology has many applications in the industry, including cellular phone systems, paging systems, two-way radio communication systems, dose management in nuclear power plants, remote monitoring and tracking systems, and operator logs (EPRI, 2004a). In addition, EPRI has reported the following case studies: Exelon´s Peach Bottom Nuclear

Power Plant, Texas Utilities Comanche Peak Nuclear Station, and in the San Onofre Nuclear Generating Station de la Southern California Edison (EPRI, 2004b). Other application cases of wireless technology in nuclear power plants in the U.S. include: Nine Mile Point Nuclear Station, Robinson Nuclear Plant (CP&L) and in the Seabrook Nuclear Power Plant (Shankar, 2003; SpectraLink, 2004; Kjesbu, S. and Brunsvik T., 2000).

6. Mobile IP telephony system for a nuclear power plant

Mobile IP telephony is the new generation of communications networks that makes possible the convergence of voice and data over wireless local area networks. It refers to the transmission and reception of voice conversations over wireless IP data networks. This technology combines data networks with mobile technologies to support voice and data applications over a common integrated network. A key advantage of IP telephony is that it allows the transmission of voice signals from wireless telephones over IEEE 802.11 wireless LANs connected to an IP network backbone, being either a public network (Internet) or a private network (Intranet). IP telephony is designed to replace the legacy TDM (time division multiplexing) technologies and networks with an IP-based data network. Digitized voice will be carried in IP data packets over a LAN and/or WAN network. A major aspect involved in a voice conversation using mobile IP telephony is the conversion of analog or digital voice signals from conventional phones to IP packets for further transmission either to a fixed or mobile phone, over an IP network.
This section presents the design considerations for the implementation of a mobile IP telephony system for voice communications applications in Laguna Verde nuclear power plant (CNLV), Federal Commission of Electricity (CFE), Mexico based on national and international standards as well as, the technical requirements that commercially available equipment must meet. In addition, the technical requirements of a mobile IP management system are presented.

6.1 Components of the mobile IP telephony system

One of the main requirements of a mobile IP telephony system is that the proposed system must meet the design technical requirements for its exclusive operation in a nuclear power plant in Mexico. Regarding standardization, it must meet national and international standards and regulatory guides applicable to nuclear power plants. Also, this system will operate upon the existing CNLV´s data backbone which is based on Gigabit Ethernet switching technology. The model proposed in this work for the mobile IP telephony system is showed in figure 3 and it is composed of the following components:

- Voice gateway
- Priority server
- Wireless Access points (AP)
- Wireless telephones

The voice gateway provides the conversion of analog or digital voice signals from conventional phones to IP packets for further transmission to any access point connected to the CNLV´s IP data network backbone. The priority server has as a main function to assign priorities according to the type of traffic to be sent over the network, in order to give the highest priority to voice packets guaranteeing in this way high quality voice conversations. The wireless access points have as a main function to interconnect the portable wireless

mobile telephones to the CNLV´s backbone. The functionality of the proposed system shall be so good that allows any user located in a wireless coverage area using a conventional analog or digital telephone connected to a private branch exchange (PBX) equipment to make calls and communicate with any other user located in some other area carrying a wireless telephone assigned to an access point or having nearby a conventional wired telephone.

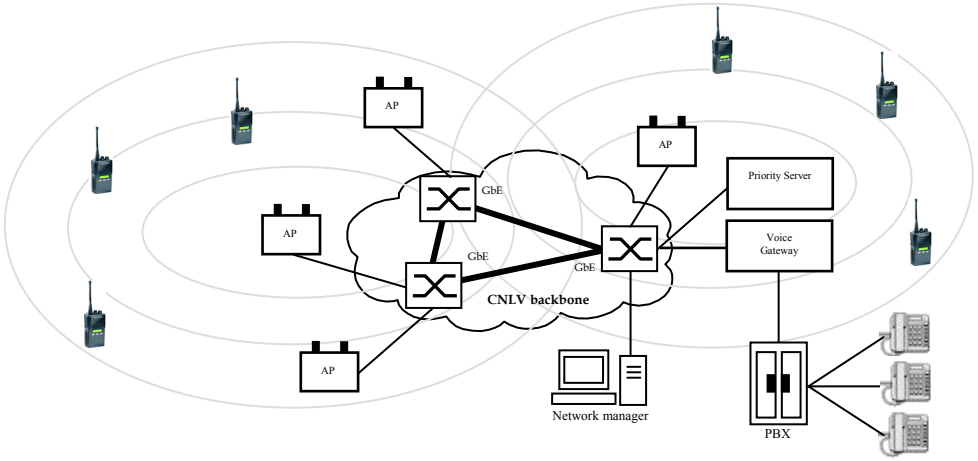


Fig. 3. Mobile IP telephony system for the CNLV nuclear plant, Mexico.

In addition, a network management system for the mobile IP telephony system is proposed. It is composed of a server and the network management software which will allow visualizing all network components such as access points, wireless telephones, voice gateways, priority server and airspace by means of a human machine interface (HMI). Next, the technical requirements that every component of the proposed mobile IP telephony system shall meet, are presented.

6.2 Technical requirements of the mobile IP telephony system

The proposed mobile IP telephony system shall support voice over IP communication applications within the CNLV controlled areas. Regarding standardization, it must meet national and international standards and regulatory guides applicable to nuclear power plants. The components of the wireless system shall be compatible with IEEE 802.11b standard (IEEE, 1999b), use direct sequence spread spectrum (DSS) a modulation technique and support a data rate of 11 Mbps. The systems must operate in the ISM (Industrial, Scientific and Medical) band from 2.4 to 2.4835 GHz, according to the NOM-121-SCT-94 standard (CCNNT, 2001) from the Mexican Normalization in Telecommunications Consultative Committee. The maximum output power radiated form the wireless components shall be 100 mW. The system shall be composed of industrial type equipment for use in nuclear power plants and must meet the acceptable electromagnetic interference (EMI) and radiofrequency interference (RFI) ranges for electronic equipment operating in frequencies above of I GHz, established in the regulatory guide 1.180 (NRC, 2003), issued by the United States Nuclear Regulatory Commission (NRC). The system must meet military

levels of electromagnetic emissions, electromagnetic compatibility and electromagnetic interference susceptibility for electronic equipment operating in the frequency range from 1 GHz to 10 GHz, according to MIL-STD-461E (DoD, 1999) military standard. Also, the system must meet the applicable electromagnetic interference and radiofrequency interference standards from the International Special Committee on Radio Interference (CISPR), the International Electrotechnical Commission (IEC) as well as, from the Institute of Electrical and Electronics Engineers (IEEE).
The system must meet the technical requirements of electromagnetic compatibility (EMC) and radiofrequency interference contained in the IEC-61000 (IEC, 2002) set of standards. It must meet national and international wireless communication standards such as NOM-121-SCT-94 (CCNNT, 2001), FCC Part 15.247 (FCC, 2004), IEEE 802.3 (IEEE, 2002) and IEEE 802.11b (IEEE, 1999b). Regarding the system components, they must be designed to operate in risky and severe environments such as: high humidity levels, temperature, radiation, noise, vibration, shock, as well as in controlled levels in a nuclear power plant. Regarding security, the system shall provide military like high security levels by using data encryption techniques as integral part of the voice over IP communication system. Also, the system must use robust electronic equipment and provide an availability at least of 99.999%, considering a redundant and fault tolerant communications architecture, as well as an uninterruptible power supply (UPS) with the aim of avoiding the system collapse.
The system shall be designed with the necessary security and authentication levels in order to deny the access to private information of validated users or their privileges by means of interception and modifications of their data messages. In addition, it must provide high confidentiality in order to guarantee very high efficiency in the activation and deactivation of voice conversations, data messages, control messages or authentication passwords. With regard to availability, then mobile IP telephony system must guarantee the continuous voice communications between mobile users using wireless telephones and the access points installed in and out the CNLV power plant. Regarding wireless telephones, the system must handle 64 spread spectrum wireless telephones per unit, being 128 in total to serve the 2 units available to the CNLV.

6.2.1 Voice bandwidth

In the traditional voice world a single T1 leased line is used to carry 24 digital channels of 64 Kbps, toll quality telephone calls from the public switched telephone network (PSTN). In its turn, those with a private point-to-point T1 connection can compress the voice to less than 64 Kbps bandwidth for more efficiency, but the quality may be sacrificed. The three most common modulation schemes for encoding voice are:

- 64 Kbps (PCM)/1.544 Mbps = 24 simultaneous calls on a T1
- 32 Kbps (ADPCM)/1.544 Mbps = 48 simultaneous calls on a T1
- 8 Kbps (CELP)/1.544 Mbps = 120 simultaneous calls on a T1

Bandwidth use and voice compression play an important role in provisioning the LAN/WAN.

6.2.2 Voice over IP bandwidth

The concept of combining traditional voice on data networks is simple because voice traffic uses a lot less bandwidth than traditional LAN-based computer networks. A single toll-quality phone call over the public network uses 64 Kbps in each direction – that's only

0.0625% of a 100 Mbps full duplex link. On a 100 Mbps Ethernet network, each voice call takes up to 85.6 Kbps (64 Kbps + IP header + Ethernet header) in each direction supporting up to 1,160 calls over a full duplex link. On a Gigabit backbone, up to 11,600 simultaneous calls can be handled.

6.2.3 Voice gateway

The voice gateway has as main function to provide the conversion of analog or digital voice signals from conventional phones to IP packets for further transmission to any access point connected to the CNLV´s IP data network backbone. For the implementation of the mobile IP telephony system, it is required to integrate the analog or digital voice circuits from the private branch exchange (PBX) Harris equipment model 20-20L installed in the CNLV to a voice over IP system based on an Ethernet wireless LAN, and compatible with the IEEE 802.11b standard (IEEE, 1999b). This function must be carried out by the proposed voice gateway. Next, the most relevant technical requirements the voice gateway shall meet are presented.

Regarding the functional characteristics, the proposed voice gateway must provide the interface between an enterprise PBX and a wireless LAN network by supporting digital and/or analog line interfaces, and shall provide the voice encoding and packetization required to transmit telephone conversations over an IP network. It must provide RJ-21 physical interfaces for its connection to the PBX, as well as a RJ-45 interface for connection to an Ethernet LAN. It must provide the necessary control signaling to the host PBX allowing wireless telephone users to access all the features and capabilities of the PBX. The setting up of voice connections and the control of voice calls within the voice gateway shall support a wide variety of packet-oriented protocol standards such as ITU-H.323 (ITU, 2009), RFC3261 (IETF, 2002), ITU-H.248.1 (ITU, 2005) or a subset of these protocols. Also, the voice gateway shall support H.323-compliant clients, allowing third-party wireless devices to make and receive calls through the host PBX. In addition, the voice gateway must provide coding and decoding of voice packets according to VoIP compression standards such as ITU-G.711 (ITU, 1988), ITU-G723.1 (ITU, 2006)or ITU-G.729 A/B (ITU, 2007) with design characteristics that enhance the quality of the voice signals to be transmitted providing silence suppression and echo cancellation functions according to the ITU-G.168 standard (ITU, 2009).

With regard to security and availability, the voice gateway to be used in a nuclear power environment shall meet military requirements of electromagnetic interference emissions, compatibility and susceptibility specified in MIL-STD-461E (DoD, 1999) and IEC-61000 (IEC, 2002) standards. In its turn, it must meet security technical requirements applicable to electronic communication equipment stated in the IEC-60950 (IEC, 2005) standard or other equivalent standards like UL-60950 y EN-60950.

6.2.4 Priority server

The proposed priority server has as a main function to recognize and to assign priorities to voice and data packets according to the type of traffic to be sent over the network, in order to give the highest priority to voice packets to reduce the voice packets time delay and causing the minimum impact on data throughput, guaranteeing in this way high quality voice conversations. Next, the most relevant technical requirements the priority server shall meet are presented.

The proposed priority server shall use a packet filtering mechanism to guarantee a high quality of service (QoS) to voice packets with the possibility of a transparent migration of the system to the IEEE 802.11e standard (IEEE, 2005) when it has been ratified. It must has the capacity of supporting up to 120 simultaneous voice calls, provide an Ethernet 100Base-T port with a RJ-45 connector, support full-duplex transmission compatible with the IEEE 802.11b standard (IEEE, 1999b) for operation in Wi-Fi WLANs and also, it must be capable of supporting native VoIP protocols and interfaces provided by the proposed voice gateway. With regard to the WLAN monitoring, the priority server shall provide a RS-232 interface for configuration and diagnostics purposes, WLAN connection indicators, system activity, collisions events and error state indicators. It shall support local mode configuration by using the RS-232 and USB interfaces or support remote mode configuration by using the Telnet application protocol over a LAN or WLAN.
Regarding the operational conditions, the priority server shall support an operation temperature range from 0° C to 40° C, relative humidity between 0% and 95% non-condensed, and a supply voltage of 110/220 VAC at an operation frequency of 60 Hz.

6.2.5 Wireless access points

The proposed wireless access points have as a main function to interconnect the portable wireless mobile telephones to the CNLV nuclear power plant´s data backbone. According to the normative defined by the Mexican Normalization in Telecommunications Consultative Committee (CCNNT, 2001) as well as, for the CNLV nuclear power plant environment operating conditions, the access points providing the wireless communication links to the mobile telephones must operate in the ISM frequency band, from 2.4 to 2.4835 GHz an must use a maximum output radiated power of 100 mW. Next, the most relevant technical requirements the wireless access points shall meet, are presented.
For the implementation of the wireless VoIP application, the proposed wireless access points shall are required to be compatible with the IEEE 802.11b (IEEE, 1999b) standard, use direct sequence spread spectrum modulation technique and support a data rate of 11 Mbps. With regard to security, the wireless access points shall provide very high security mechanisms to voice and data packets during transmissions of voice conversations by supporting at least the WEP (Wired Equivalent Privacy) encryption technique with 128 bit keys, and the possibility of easily migrate to the WPA (Wi-Fi Protected Access) encryption scheme, as well as to support the security mechanisms included in the IEEE 802.11i standard (IEEE, 2004).
Other security functions are desirable such as continuous airspace monitoring, detection and blocking of access points that are not part of the CNLV´s data backbone, etc. The wireless access points shall be compatible with quality of service mechanism supported by the proposed priority server and the voice gateway, with the possibility of migrating in a transparent manner to the IEEE 802.11e (IEEE, 2005) when it has been ratified. The physical distribution and the number of the wireless access points installed in the CNLV nuclear power plant must guarantee a wireless coverage area of 100 % in each controlled area so the use of dual external antennas shall allow a much better distribution of electromagnetic signals radiation pattern so that no black spots exist.
In order to improve the reliability and availability of the mobile IP telephony system, the wireless access points shall provide two 10/100 Mbps Ethernet ports one for physical connection and the other one for a backup network segment. The power supply for the wireless access points shall be provided by means of a UTP network cable supporting Power

over Ethernet (PoE) mechanism, according to the IEEE 802.3af standard (IEEE, 2003). In addition, it is recommended that each wireless access point shall provide an independent 110/220 VAC voltage input.
The legislation that the wireless access points must meet, includes the regulation emitted by the Federal Communications Commission, FCC Part 15.247 (FCC, 2004) for digitally modulated intentional radiators devices, and the security and electromagnetic interference requirements (DoD, 1999), (IEC, 2002), (IEC, 2005), in order to respect the acceptable electromagnetic interference and radiofrequency ranges for electronic communication equipment operating at frequencies above 1 GHz according to the Nuclear Regulatory Guide 1.180 (NRC, 2003), emitted by the Nuclear Regulatory Commission.

6.2.6 Wireless telephones

The proposed wireless telephones will be used by personnel working in the external areas of the CNLV nuclear power plant, conducting fieldwork so that they have to be robust designed for using in industrial and nuclear power plants, in particular. Next, the most relevant technical requirements the wireless telephones shall meet, are presented.
The wireless telephones shall be compliant to the IEEE 802.11b (IEEE, 1999b), H.323 (ITU, 2009), G.711 (ITU, 1988), G.729 (ITU, 2007) standards as well as to VoIP protocols emitted by international standards bodies. Besides, they must support the capability of sending and receiving short text messages via open application interface. The wireless telephones shall support both static and dynamic (DHCP) IP addressing configuration and must operate in the ISM frequency band, from 2.4 to 2.4835 GHz, according to the NOM-121-SCT-94 standard (CCNNT, 2001), issued by the Mexican Normalization in Telecommunications Consultative Committee. They shall be compliant to the IEEE 802.11b (Wi-Fi) standard (IEEE, 1999b), use direct sequence spread spectrum (DSSS) modulation technique and support data rates of 11, 5.5, 2 and 1 Mbps, which must be automatically selected according to the communication channel conditions and voice quality of service.
With regard to radiated power, the wireless telephones shall produce a maximum transmission power below 100 mW (20 dBm), which must be automatically adjusted in order to have always the same radiated power level. They shall provide very high security mechanisms to voice and data packets during transmissions of voice conversations by supporting at least the WEP (Wired Equivalent Privacy) encryption technique with 128 bit keys, and the possibility of easily migrate to the WPA (Wi-Fi Protected Access) encryption scheme, as well as to support the security mechanisms included in the IEEE 802.11i standard (IEEE, 2004).
In addition, the proposed wireless telephones shall provide an LCD backlit dot matrix display with icons and line-status indicators with the aim of visualizing the entire display in darkness conditions. They shall support the instant communication feature known as push-to-talk (PTT) by using IP multicast addresses. This requires that multicasting be enabled on the subnet used for the wireless telephones, priority server, and voice gateway. They shall provide an integrated TFTP client in order to allow remote software updates via the TFTP (Trivial File Transfer Protocol) application. Also, wireless telephones must be lightweight with a weight less than 200 grams.

6.3 Mobile IP management system

In addition to the mobile IP telephony system, a network management system is proposed. It consists of the network management server and the network management software. Next,

the most relevant technical requirements the network management system shall meet, are presented.
The proposed network management server shall provide the following minimum capacities: 1.8 GHz processor (Pentium IV), 256 MB RDRAM, internal 40 GB hard disk, a CD-ROM unit, a 20" color monitor, and a 10/100 Mbps Ethernet network card. For its part, the network management software must be capable of visualizing all components of the mobile IP telephony system such as access points, wireless telephones, voice gateway, and priority server) as well as the airspace. With regard to capacity, the network management software shall provide management functions like configuration, performance monitoring, fault detection, network statistics, and security, among others. Regarding functionality, it shall support functions such as discovering, configuring and monitoring all access points connected to de CNLV data backbone, allowing the configuration all wireless devices specified in the design of the mobile IP telephony system with just one click.
In addition, the network management system shall provide management tools such as monitoring and measurement of the wireless network performance (delay, throughput, etc.), used and available bandwidth, wireless network use, among others parameters. It shall provide wireless network statistics such as transmitted and received signal level, number of transmitted and received IP packets, frequency deviations, and changes in data rate for each access point.
With regard to security, the network management system must be a centralized-type system, and be capable of providing the mobile IP telephony system with a high level of security by means of monitoring both the physical network devices and the wireless pace used by the system. Also, it shall detect most of wireless network cyber attacks including massive attacks, intrusions, impersonation, sniffers, denial of service (DoS), etc., and finally the network management system must has the ability to perform remote software upgrades to wireless telephones from the network management´s central station.

6.4 Implementation of the mobile IP telephony system at CNLV

In this section, an example of use of the proposed mobile IP telephony system for voice communications applications in Laguna Verde nuclear power plant (CNLV) is presented. Once the design considerations for the implementation of a mobile IP telephony for voice communications applications were carried out, the Federal Commission of Electricity (CFE), Mexico began the system acquisition phase with an international bidding in order to have a winner. Then, the components of the mobile IP telephony system such as: access points, voice gateway, priority server, and wireless telephone, etc., were supplied and installed in the selected controlled areas of the CNLV nuclear power plant. After this, the implementation phase began. The acquired mobile IP telephony system was installed at CNLV´s telecommunications room, and now it is operating upon the existing CNLV´s data backbone which is based on Gigabit Ethernet switching technology. The system provides communication applications such as telephony and voice over IP.
Another example of use of wireless LAN technologies in the nuclear power plant environment from the previous project is that, CFE has initiated a new implementation phase consisting of the introduction of wireless IP video technology with the aim of having a true integrated data, voice and video system using the same CNLV´s network infrastructure. The proposed IP video system will be used for remote video monitoring and video surveillance within the CNLV nuclear power plant taking advantage of the IEEE 802.11b/g standard-based wireless

network technology already installed. The main components of the system are: wireless IP video cameras, massive storage unit (terabyte network attached storage), and a video monitoring and surveillance station. The proposed IP video system which will be integrated to the existing wireless network is shown in figure 4.

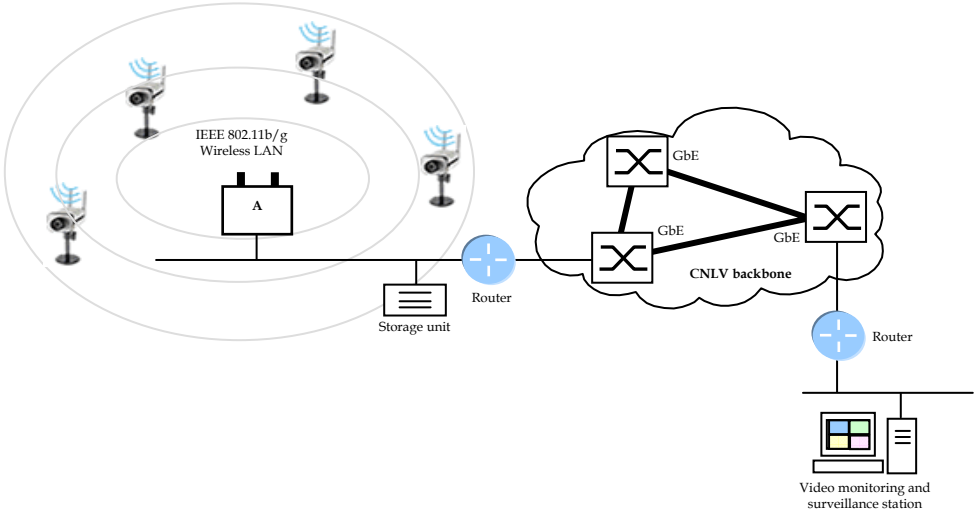


Fig. 4. Proposed wireless IP video system for the CNLV nuclear plant, Mexico.

7. Conclusions

In this chapter, the design considerations for the implementation of a mobile IP telephony system for voice communications applications in Laguna Verde nuclear power plant (CNLV), Federal Commission of Electricity (CFE), Mexico based on national and international standards were presented. Also, this work gave an analysis of the most relevant wireless technologies currently available that can be implemented in nuclear power plants and also identified nuclear regulatory guidelines, wireless networks standards, electromagnetic and radio-frequency interference standards. With regard to the use of wireless LANs in the nuclear environment, there is clear evidence that the electromagnetic interference and radio-frequency interference conditions can adversely affect the performance of safety-related instrumentation and control equipment. EMC is an element of addressing that requirement. Operational and functional issues related to safety in the nuclear power plant environment are required to address the possibility of troubles and malfunctions in instrumentation and control systems caused by electromagnetic emissions (EMI/RFI) from wireless technology. On the other hand, WLAN technology based on the IEEE 802.11 standards, has a very promising future for its use in nuclear power plants, due to its features like mobility, reliability, security, scalability and compatibility with other technologies. Currently, WLAN technology is been installing and evaluating in nuclear power plants worldwide, due to it provides enhanced features compared to traditional wireless technologies such as conventional mobile radio in two key aspects: higher operation frequencies and lower output power, which translates in very high data rates and

very low electromagnetic interference. With regard to system design, a mobile IP telephony system based on wireless local area networks which will operate upon the existing CNLV´s data backbone, has being proposed. In addition, the technical requirements that each commercially available component system must meet for its correct operation regarding the compliance with national and international standards, recommendations, regulatory guides, reliability and availability metrics, and security mechanisms, were established. Within the most important aspects identified in this work, are that the mobile IP telephony system must meet the design technical requirements for its exclusive operation in a nuclear power plant in Mexico, as well as to compliant to existing national and international standards applicable to nuclear power plants. Finally, the technical requirements of a network management system consisting of a network management server and network management software for the mobile IP telephony system, have been specified.

8. References

Shankar, R. (2003). Guidelines for Wireless Technology in Nuclear Power Plants, *11th International Conference on Nuclear Engineering, ICONE11*, pp. 1-9, Tokio, Japan.

IEEE (1999a). IEEE Standard 802.11, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*.

IEEE (1999b). IEEE Standard 802.11b, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Higher Speed in the Physical Layer Extension in the 2.4 GHz Band*.

Martínez, E. (2002), Estándares de WLAN, *Revista Red*, No. 139, pp. 12-16, Mexico.

IEEE (1999c). IEEE Standard 802.11a, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. High Speed Physical Layer in the 5 GHz Band*.

IEEE (2003). IEEE Standard 802.11g, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Further Higher Data Rate Extension in the 2.4 GHz Band*.

IEEE (2005). IEEE Standard 802.11e, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. MAC enhancements for Quality of Service*.

IEEE (2004). IEEE Standard 802.11i, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. MAC enhancements for enhanced security*.

NUREG (2003). Final report NUREG/CR-6782, *Comparison of U.S. Military and International Electromagnetic Compatibility Guidance,* USNRC, pp 34-36.

NRC (2003). RG 1.180, *Guidelines for Evaluating Electromagnetic and Radiofrequency Interference in Safety-Related Instrumentation and Control Systems*, U.S. Nuclear Regulatory Commission.

IEEE (1996). IEEE 1050 Standard, *Guide for Instrumentation and Control Equipment Grounding in Generating Stations.*

DoD (1999). MIL-STD-461e1 Standard, *Requirements for the control of electromagnetic interference, characteristics of subsystems and equipment,* U.S. Department of Defense.

IEC (2002). IEC 61000 Standard, *Electromagnetic Compatibility (EMC)-Testing and Measurement Techniques*, International Electrotechnical Committee.

EPRI (2003). Electric Power Research Institute (EPRI), *EMI/RFI Issues*, Technical Note, sections 3.3-3.6, pp. 49-50.

EPRI (2002). Electric Power Research Institute (EPRI), EPRI Report TR-03T023027, *Guidelines for Wireless Technology in Nuclear Power Plants,* available from

http://www.epri.com/targethigh.asp?program=249866&value=03T023027&objid=284710.

FCC (2004). CFR 47, Part 15, *Radio frequency Devices*, Federal Communications Commission.

CCNNT (2001). NOM-121-SCT1-94, *Telecomunicaciones - Radiocomunicaciones - Sistemas de Radiocomunicación que emplean la Técnica de Espectro Disperso*, Comité Consultivo Nacional de Normalización en Telecomunicaciones.

Meel, J. (1999). Report, Spread Spectrum (SS) Introduction, De Nayer Instituut, Belgium, pp. 1-33.

DoE (2002). U.S. Department of Energy, *Industrial Wireless Technology for the 21st Century*, white paper, DoE.

Pearce, J. (2001). *FCC Considerations for Spread Spectrum Systems*, available from http://www.sss-mag.com/fccss.html.

Bahavnani, A. (2001). *An Analysis of Implementing Wireless Technology to further enhanced Nuclear Power Plant Cost efficiency, Safety and Increased Employee Output*, Pressure Vessel and Piping Design and Analysis, Vol. 430, pp. 369-372, ASME 2001.

Telrad Connegy (2001). Telrad Connegy web page, available from http://www.telradusa.com/pr_chernobyl.htm

Wireless Magazine (1995). *Wireless Improves Safety at Hungary Nuclear Power Plant*, Wireless Magazine, Vol. 4, No. 6, Nov/Dec, 1995.

EPRI (2004a). EPRI *Wireless Technology newsletter No. 1009624*, July 2004.

EPRI (2004b). EPRI Journal on-line, http://www.epri.com/journal/details.asp?doctype=products&id=533&flag=archi ve.

SpectraLink (2004). SpectraLink web page, available from http://www.spectralink.com/solutions/case.html

Kjesbu, S. and Brunsvik, T. (2000). *Radiowave propagation in Industrial Environments*, 26th Annual Conference of the IEEE Electronics Society, IECON 2000, pp. 2425-2430, Nagoya Japan.

IEEE (2002). IEEE 802.3 Standard, *Local and Metropolitan Area Networks - Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*, Institute of Electrical and Electronic Engineers.

ITU (2009). H.323 Recommendation, *Packet-Based Multimedia Communications Systems*, International Telecommunications Union.

IETF (2002). RFC 3261, Session Initiation Protocol (SIP), Internet Engineering Task Force.

ITU (2005). H.248.1 Recommendation, Gateway Control Protocol, International Telecommunications Union.

ITU (1988). G.711 Recommendation, Pulse Code Modulation of Voice Frequencies, International Telecommunications Union.

ITU (2006). G.723.1 Recommendation, *Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbps*, International Telecommunications Union.

ITU (2007). G.729 Recommendation, *Coding of Speech at 8 kbps using Conjugate-Structure Algebraic Code Excited Linear-Prediction*, International Telecommunications Union.

ITU (2009). G.168, *Digital Network Echo Cancellers*, International Telecommunications Union.

IEC (2005). IEC 60950 Standard, *Information Technology Equipment–Safety*, International Electrotechnical Commission.

IEEE (2003). IEEE 802.3af Standard, *Power over Ethernet*, Institute of Electrical and Electronic Engineers.

5

Smart Synergistic Security Sensory Network for Harsh Environments: Net4S

Igor Peshko
Department of Mechanical and Industrial Engineering, University of Toronto
Department of Physics and Computer Science, Wilfrid Laurier University
Canada

1. Introduction

This chapter discusses the basic requirements for the design and algorithms of operation of a multi-parametric, synergistic sensory network – Smart Synergistic Security Sensory Network or Net4S – specially adapted for operation at nuclear power plants or other potentially dangerous sites. This network contains sensors of different types and is capable of analyzing the dynamics of environmental processes and predicting the most probable events. The discussion includes analysis of: 1) the technical aspects of operability of the sensors, optical and electrical telecommunication channels, and computers in the presence of ionizing radiation; 2) the influence of environmental parameters on the sensors' accuracy and network operability; and 3) the development of simulators capable of advising safe solutions based on the analysis of the data acquired by the Net4S. Such a real-time operating network should monitor: (1) environmental and atmospheric conditions – chemical, biological, radiological, explosive, and weather hazards; (2) climate/man-induced catastrophes; (3) contamination of water, soil, food chains, and public health care delivery; and (4) large public/industrial/government/military areas. Military personnel, police officers, firefighters, miners, rescue teams, and nuclear power plant personnel may use the mobile terminals (man-operated vehicles or unmanned robots) as separate multi-sensor units for local and remote monitoring.

Among different types of sensors, only optical laser sensors can respond immediately and remotely. Such sensors can simultaneously monitor several gases, vapours, and ions with the help of single tunable laser; however, the use of several lasers operating at different, well separated wavelengths, dramatically improves accuracy and reliability, and increases the number of monitored substances. The Net4S, monitoring a number of parameters inside and outside a Nuclear Power Plant (NPP), can serve as the security, safety, and controlling system of the NPP.

Besides the technical issues, the chapter also discusses the social aspects of the Nuclear Power Plants' design, construction, and exploitation. Some power consumption-free technologies that significantly improve the reliability of the Nuclear Power Plant are discussed.

In principle, open access publishing is a purely commercial project. After submitting a paper to classical journals, the author should wait for a relatively long time and should fight with the reviewers – the "narrow specialists" are the author's competitors and usually state that

everything is known, that the subject of publication is not interesting, and that the author is of low qualification. The "wide specialists" do not understand what the paper is about and criticize in general – the current tendency in science and technology are out of the subject that the author discusses, that any laser now can be bought off the shelf, and so on. Because of this paradox, a lot of the papers that were later nominated for prestigious awards were initially rejected. In some sense, open access publishing is free from these disadvantages. However, since the publisher should generate maximum profits from this activity, the high requirements for the quality of publications are difficult to be completed.

The next argument is then why so critically select the papers if, anyway, no one can or wants to estimate the real value of new papers. At the same time, open access publications have one very serious advantage. Government experts mostly write in reports what their chiefs expect to hear from them, post-graduate students write to please their supervisors, and Professors write proposals on subjects that the funding agencies declare in calls, and so on. These are because, directly or indirectly, all these categories are payable from the "top publishers". So, once a funding agency declares a solicitation for the investigation of ozone hole, a lot of researchers demonstrate how dangerous the hole is. As soon as the funding ends, nobody remembers what the ozone hole is.

A somewhat different situation is present with open access publishing: the *author pays* for the publication, so he/she is almost free not to lie. However, other public requirements, such as generating more publications before a thesis defense, getting a Professorship position, or being awarded by a Government or private agency push people to publish something. Thus, they invest money in future benefits. No one is absolutely honest and those who believe that they are, very often have limited knowledge of the subject they discuss and analyze. The ways to develop a really safe and effective Nuclear Power plant are very twisted and long. The NPP is very big, complex, expensive to be built and proven in different variants. Drosophila flight is much more perfect in design and implementation since the generation time is several tens of hours, not tens of years as it is for NPPs. Until now, the problem of design and safe exploitation of a NPP is very challenging and uncertain.

The author of this chapter is a specialist in laser physics and optical sensors, not in atomic physics or its applications. However, Dr. I. Peshko was working in Kyiv, Ukraine at the moment of Chernobyl's "peaceful explosion" and watched the reaction and behavior of regular people, academics, government organizations, and researchers. These observations can be very useful for analytical specialists who develop general principles of design, exploitation, and control of the NPPs. In such a "twilling zone" as the NPP, the probabilistic estimation of a single independent person may sometimes be more valuable than official reports and opinions of specialists. The bottom line is that official reports are typically prepared by specialists and officials to protect themselves and to hide their past mistakes, not to protect the future of millions of people. Every time I think about Chernobyl's events, I remember my mother who spent all her life as a housekeeper in a small town in Northern Ukraine and understood nothing about atomic energy. One day, when a radio broadcast informed us about the government's decision to build Chernobyl's Nuclear Power Plant, my mother said, "My feelings are very bad. How is it possible to construct a nuclear station in a place that is a source of water for tens of millions of people?" As I laughed, I replied, "The Chief of the Atomic energy program promised to install his bed on the top of the reactor to demonstrate how safe a reactor is." Unfortunately, time has shown how wrong the best specialist was and how right a regular housekeeper was.

2. Synergistic sensory network

2.1 Threat classification

Nuclear Power Plants are strategically important objects that may be affected by internal and external threats. Consequently, a NPP is considered a potential source of danger to its surroundings and, in turn, environmental elements – natural, artificial, and human factors – are potential dangers for a NPP.

Five types of possible threats potentially affecting the NPP are:

1. Natural catastrophes;
2. Technological (internal and external) problems resulting in emergencies;
3. Terrorist threats;
4. Personnel and security staff sabotage;
5. Scientific uncertainty and scams.

The first and second threats in the list above were widely discussed and documented during the initial stages of the development of nuclear technologies. The third threat became extremely evident after the 9/11 attacks and until now, is a very popular topic of discussion at different public and government levels. The fourth threat may be linked with both internal and external country sources and may have criminal and political backgrounds. Finally, the fifth threat, to our knowledge, is discussed for the first time in this book. It is not an issue for detailed discussion here but this is a very serious problem of modern and future life. The falsification of scientific results; demonstrations of non-existing products or unachieved parameters on the Internet; publication of preliminary, "fast" materials in numerous journals; and awarding grants on the basis of relationships rather than merit result in unpredictable events with critical technologies.

I would like to present one example from my personal experiences. A very famous Canadian Professor, whom I was working with, proposed a thin diffractive grating filled with a biological material as a biosensor. The more specific substance the grating accumulates, the stronger the diffraction is. This works in some range of small changes in grating strength. However, the Bessel function that describes the diffraction process of the thin grating has multiple zero points (solutions); in other words, for several *different amounts* of measured substance, the output signal will be *the same*. I gently mentioned that this kind of technology cannot be used for sensor applications and two weeks later, was fired for some formal reasons. If a tenured Professor of a famous University does not know the properties of the Bessel functions, this is very bad. However, if the Professor knows this and hides it just to receive a grant for the "development" of critical technology, this is much worse.

In attempts to forecast the future, the principle question is: if we know that we don't know, how do we develop a probabilistic solution of the problem with minimal material losses? How can we estimate and forecast of "unpredictable" events? First of all, we need to collect maximal real-time flows of information. To control the situation inside and outside of a NPP, the Sensory Network should monitor several zones: a) core (reactor) area; b) plant building and surrounding territory; c) 30-km radius zone (the Chernobyl tragedy showed that the strongest radioactive poisoning happened within a 30-km zone); d) in North America: Mexico - USA - Canada region (depending on the specific plant location). Thus, a NPP is a duplex element of the global security network. It needs to accept information from near and far environmental areas, and information regarding what is going on inside the NPP should be retrievable from any control station in the country.

The safety zone classification depends on the reactor construction, type of emergency, population density, and the locations of other industrial plants. In the case of the recent Fucushima reactors catastrophe in Japan, the officials specified 5-km and 20-km evacuation zones.

2.2 Principles of 4SNet

The development of a Global Monitoring Security Network is the main task on route to several scientific, technological, business, military, and political directions of modern life. Such a real-time operating network should monitor: (1) environmental and atmospheric conditions: chemical, biological, radiological, explosive, and weather hazards; (2) climate/man-induced catastrophes; (3) contamination of water, soil, food chains, and public health care delivery; (4) large public/industrial/government/military areas. Such a system is expected to consist of mobile robotic and stationary platforms, equipped with a set of portable environmental sensors that are connected to the monitoring centers. Each sensor should be a self-registering, self-reporting, plug-and-play unit that uses unified electrical and/or optical connectors and operates with the IP communication protocol. Military personnel, police officers, firefighters, miners, rescue teams, and nuclear power plant personnel may use the mobile terminals (man-operated vehicles or unmanned robots) as separate multi-sensor units for local and remote monitoring. Some of the objects being monitored require special attention, such as nuclear and chemical plants, offshore oil platforms, mines, military ammunition production facilities, and so on. The Net4S components must operate at varying pressures and temperatures; at indoor and outdoor conditions; be immune to mechanical, thermal, electro-magnetic and radiological noise; and be able to operate in case of electrical blackouts.

In different areas of the reactor and surrounding territories, different types of sensors can be installed. This makes it possible to map temperature, ionizing radiation of different types, gas molecules and ion concentrations, vapors, and presence of dust particles. The overlapping of all these maps and reconstruction of their dynamics can predict what will happen in the close future. During several initial cycles of reactor operation in a "manual regime", the dynamics of all parameters should be recorded and analyzed. During the next routine operation, the total network should permanently measure the data, map them, and compare with previously averaged data. If even small changes of parameters are accumulated along time, this is a sign for alarm. It does not matter which parameter is out of the norm. A negligible event may initiate a catastrophe: a cup of coffee left by a personnel on the operational panel may flip over and cause damage to the electronics located under the desk. Of course, everyone can tell me that nobody is permitted to drink coffee on the command desk, and I absolutely agree, but I definitely know that real life is much richer with possibilities than any designer or programmer can imagine.

During the design stage, any chains of possible undesirable events should be simulated and analyzed. Let us continue the hypothetical "flipped coffee" example. Because of the short circuit in the desk electronics, several high power circuits in the power commutation station are simultaneously activated. This results in a fire and uncontrollable activation of the fuel reloading system that, in turn, results in the quick heating and destruction of the reactor. This example is naïve, very simplified, and may never be realized in practice due to specific reactor construction details and algorithms of operation; however, it helps to understand that to design a nuclear reactor, psychologists and specialists in the traditions of different cultures should be involved, not just specialists in nuclear physics. Previous background and

experience are very important as well. In case of a sudden earthquake, people who experience it for the first time will chaotically look around; those who have survived a strong earthquake may be in panic, but will run away as fast as possible. In both cases the reactor may be out of personnel control. So, it is better if the territory around the plant is supplied with sensors that can measure the amplitude of impact, activate the reactor shut down system, and sound alarms for the personnel. An even better solution is one where the Global Security Network can directly and automatically inform the NPP that a tsunami is approaching.

2.3 Reliability of an inhomogeneous network

In order to improve reliability, sensor redundancy (using multiple instances of a sensor) can be implemented; however, adequateness (ensuring the measured signal pertains only to specific parameters) is still not guaranteed. In real life, it is practically impossible to isolate a single process and be certain that the measurement is related to just one variable. A readable sensor signal may appear as a result of one "strong" interaction with an object, or several indirect interactions that affect the sensor in the same way as the "strong" one. Thus, the problem of reliability is apparent in these measurements, especially if we need to measure in unexpected, unpredictable, and unfriendly conditions. As a simple example: some house fire alarm sensors are typically activated every time someone takes a shower; both water and fire are interpreted as the same entity by the sensor. These sensors were tested for fire emergency events and definitely work well in corresponding conditions; however, nobody thought to test them in high humidity conditions, an absolutely "opposite" range of application. The result of this is that after several false alarms people typically turn off a fire alarm sensor. Thus, the adequateness of measurements is questioned every time.

A sensory network, where the sensors operate in different physical domains, should be used. This creates an inhomogeneous network with a variety of sensors capable to perform joint analyses and mapping of different datasets.

References (Peshko, 2007; Matharoo, 2010) discuss a concept of an "inhomogeneous network". This network combines a set of different types of sensors to measure different parameters (sub-networks), and different types of sensors that measure the same parameter but based on different physical phenomena. For example, temperature can be measured by a bi-metallic thermometer (mechanical thermo-deformation), by a thermocouple thermometer (a junction between two different metals that produces a voltage related to a temperature difference), and can be calculated from the gas optical absorption spectra (spectral line broadening is proportional to the temperature). Evidently, in this hierarchy, the simplest implementation (and one that does not require any power supply) is the bi-metallic thermometer. It may not give information very precisely, but it does "survive" in harsh conditions. In an inhomogeneous network, the sensors synergistically collect and analyze information that individual sensors cannot. This information may be used as a rough measurement for evaluation of more sophisticated multi-parametric processes. If one knows the local temperature of a gas even with relatively low accuracy, the gas concentration remote measurement based on spectroscopy principles may be many times more accurate than if the temperature is unknown.

The sensor network should be analyzed and tested very carefully for the possibility of very rare but theoretically possible scenarios: due to strong irradiation, signals may saturate the transmittance of the processing system that may be interpreted as no signal or a very weak signal.

The required ability to interface with different sensors poses a challenge in maintaining a high level of overall system reliability. Using duplicate sensors for the same task decreases the probability of failure. If different sensors are used, each type of sensor needs to be rigorously tested to identify its most appropriate ranges and conditions of operation. Once this data is available for all the different types of sensors, an algorithm will be deployed to choose the sensor that has the likelihood of providing the most accurate reading at those environmental conditions. This provides a base platform for synergistic reliability. The best way is if the same set of parameters, such as level of radiation, temperature in some specific places, humidity, and presence of some gases or ions, can be measured locally and remotely. A difference in data, being acquired by local and remote sensory networks, means that "something is wrong".

A typical situation in science and technology is one when different groups of scientists and engineers developing devices working in the same area of research or technology fight with each other, proving which technology is better, cheaper, more accurate, and so on. For such sites as a NPP, the “single best choice” is unacceptable as nobody can predict for sure which technology will survive longer and would be more accurate in some unexpected conditions.

The data acquired at a NPP should be accessible (monitored) at plant command station but the NPP's personnel should not have access and ability to modify these data. They should be transferred to the external command and processing center. Even in cases when the data seems incorrect or “stupid”, they should be transferred and analyzed together with data from surrounding areas. A meteorite can be registered by seismic, gaseous, and temperature sensors 5 km away from a NPP and this can be interpreted by the NPP’s security network and personnel as a nuclear bomb explosion. In any case, the reactor cannot be stopped immediately, so each minute is crucial when preparing for critical events.

2.4 Synergistic cross-data

In an inhomogeneous, multi-level security network, each sensor, first of all, is responsible for measuring some specific parameters; at the same time, it supplies other sensors with some additional information that serves for more accurate measurements, more precise description of the investigated multi-parametric phenomena, and for the development of some conclusions about the characteristics of monitored events. Typically, the smart sensory network uses a set of sensors that control some secondary phenomena but still help in evaluation of the main process. For example, the level of ionizing radiation around a reactor in a power plant can be monitored with a set of scintillators; however, the concentration of the ionized air over the reactor can be measured well remotely and the radiation level can be estimated. Of course, this is not a direct measurement and it strongly depends on the reactor construction and principles of operation. Though, in case of an emergency, such estimations can be done from hundreds or even thousands of meters away from danger zones. Being preliminarily calibrated, this technique can provide quite accurate measurements.

To introduce the concept of an “inhomogeneous network of synergistic sensors”, consider a simple example. If your home thermometer, barometer, and humidity meter show values of 28°C, 750 torr, and 70% respectively, considering these devices individually, one can conclude that the weather is beautiful. Now, a synergistic complex, which is actually a set of different sensors – humidity, temperature, pressure, oxygen, methane, carbon oxide/dioxide, etc. – tells you that during the last two hours, the pressure fell from 770 to 750 torr, the humidity increased from 45% to 70%, the temperature increased by 3°C, and

the concentration of methane in your kitchen increased from 0.0002% to 0.2%. The dynamics of pressure, humidity, and temperature readings tell you that a hurricane is approaching, while the methane reading tells you that there is a gas leak in the house. Each separate reading does not say something terrible, but the history of parameter changes may predict that the roof of your house (that you were going to repair), may be destroyed by a hurricane, and because of the methane explosion, your house will be on the news.
A very important feature of the synergistic sensory complex is its ability to predict events; thus, the complex can alert you that the current, "beautiful" environmental data is just the beginning of a critical event. As another example, all gasoline stations are supposed to be equipped with fire alarm sensors; however, no one has considered implementing detectors for the presence of explosive materials or checking the quality of the electrical ground of fuel tanks and electronic equipment at the station. Potential sources of sparks, burned cigarettes, or explosive materials should be monitored *before* the fire starts and is then detected. Thus, the fire alarm sensory network should be "inhomogeneous" – it must contain different types of sensors capable of synergistically analyzing different scenarios.
A combination of several sensors can provide an estimation of an environmental event or emergency. For example, in case of a fire, CO, CO_2, H_2O vapour, and other specific gases (C_xH_y, NO_x) are emitted. However, the temperature and relative concentrations of these gases are different in the case of burning gasoline, wood, or plastic. A smart, multi-gas, multi-functional sensor would be able to tell the difference between a well-done BBQ on the stove versus a stove on fire. The difference is in the corresponding gas concentrations and character of light. A flame is chaotically modulated whereas a lamp over the stove irradiates light with constant intensity.
By referencing the measured concentrations with a database and analyzing the deviations in environmental conditions, the sensory platform can immediately generate the most plausible reason for the emergency. Analysis of space-time event maps and weather conditions will help to remotely identify the event and predict its dynamics.

3. Natural inhomogeneous network

It is often said that nature is the best creator, and that after many years of evolution, we are all products of good design. Our bodies are complex systems comprising of sensors, a central processing unit, and actuation devices. The human sensory-network is an example of a "well-designed" system. Every day, we use our senses of smell, touch, taste, hearing, balance, and vision, and although different sensors located throughout the human body register these sensations, the information gathered is sent through the same neurons to the brain, where it is processed and interpreted. After the data is processed and a decision is made, the "CPU center" activates a movable platform – the body. The decision made is based on information extracted from sensors specializing in different domains, i.e. analysis of electromagnetic fields, mechanical vibrations, chemical reactions, etc. The design of new technology is often driven by efforts to mimic designs found in nature. The problem is how to develop a "smart" sensory network for a power plant and environment monitoring that operates in a similar fashion to its biological counterpart, yet is capable of performing tasks not possible by natural sensory-organs, in an effort to increase public and private security.
As an additional example, imagine that you say to your significant other that you love them. You then receive feedback signals from different information channels – verbal responses, facial expressions, body movement, breathing patterns, etc. Each separate channel may

generate a false signal or no signal, e.g. they may close their eyes, but is it because they are happy or afraid to say "no"?
If any sensor/channel of information fails, the total human ability drops down; however, because of synergistic inhomogeneity, a human still operates, i.e. visually impaired people.
Another very interesting capability of the human sensory network is that if one channel fails, the other ones increase sensitivity to compensate for the lost data set. This is why visually impaired people often have an "absolute musical" hearing and can easily recognize similar sources of sound belonging to different objects, i.e. the footsteps of different people. How to teach or train the 4SNet for these capabilities is not currently clear.

4. How and what to do?

From an initial glance, the market is full of different types of sensors; however, there are still some gaping holes. For example, there are many methane sensors on the market, but thousands of miners around the world still die each year due to methane asphyxia or explosions. Similar arguments can be made for carbon monoxide sensors. NASA still announces a competition for the development of O_2, CO, and CO_2 sensors for extra-terrestrial missions; military and recreational divers still lack compact, reliable, and long-lasting sensors for the control of breathing gases; soldiers still die from roadside bombs; and airport security systems still do not detect explosives well. Current tendencies in advanced technologies pertain to the development of simple, cheap hardware and sophisticated software. Each sensor measures something; the deficiency, however, is in the interpretation of the data, shifting the problem from the real to the virtual world – complicated software might be more unpredictable and unstable than complicated hardware. However, it is much cheaper to correct software and to reload processors than to repair or upgrade millions of sensors.
To summarize, we then pose the following question: What are the basic requirements for a "universal", portable alarm sensor capable of operating on a movable robotic platform or in a life-supporting system? Such a sensor should demonstrate:

1. Immediate response;
2. Reliability: several processes are used to measure one parameter;
3. Multi-functionality: one process is used to measure several parameters;
4. Operability in hard environmental conditions;
5. Cheap, effective, simple hardware;
6. Sophisticated, "smart" software;
7. Low power consumption;
8. Self-calibration ability;
9. Synergistic data processing;
10. No additional external devices: pumps, calibrator, power supplies;
11. Immune to thermal, radiation, and mechanical noise;
12. Compatible with other sensors, sensory networks, and scientific instruments.

5. Nuclear power plant operational conditions

A nuclear power plant is a very specific object where the requirements for the Net4S are especially high. There are some technical problems in the sense of network exploitation. The optical elements (fibers, lasers, optics) can be colored under ionizing radiation. The main

components of electronics (semiconductor materials) are affected by such radiation as well. The penetrating radiation can affect the computer and electronics operation without even physically destroying these elements, resulting in the generation of false signals through the system. So, the optical sensors have some troubles, operating in this area.

In space, nuclear power, and other scientific applications, optical glass may be exposed to high-energy radiation like gamma-, electron, proton, and neutron radiation. With the accumulation of higher doses, this radiation changes the transmittance of optical glass especially near the UV-visible edge of the spectrum. The investigations of resistance of glasses versus ionizing radiation were intensively provided in 50's; these investigations were connected with research on nuclear bomb action on optical devices and other techniques.

Generally speaking, a long history of space exploration and NPP exploitation has accumulated enough knowledge on safe operation of opto-electronic devices at regular reactor conditions. However, for emergency cases, the sensory network should be protected so as to survive in catastrophes similar to the one in Chernobyl. First of all, a circuit of well-protected sensors should be installed on the perimeter of the NPP to supply the "outside" world with information in case the internal system is down. As this chapter is oriented for a wide range of readers, let us consider very shortly the problems in design and construction of internal opto-electronic sensors.

Firstly, any glass components (fibers, objectives, prisms, filters, etc.) located in the reactor and surrounding zones can be affected by ionizing radiation. Ionization caused by photon and particle radiation, changes the transmittance of optical glasses (Friebele, 1974; Schott, 2007; Sigel, 1974; Smith, 1964). An absorbed radiation dose of 10 Gy (10J energy of absorbed ionizing radiation by 1 kg of matter) gamma radiation leads to recognizable loss in transmittance over the complete visible spectral range. The decrease of transmittance is most significant at the UV-edge of the spectrum. Most glasses become unusable for optical applications if the radiation is increased to 100 Gy. The intensity of the color change does not only depend on the type of radiation dose but also on the energy of the ionizing radiation and the radiation dose rate.

Optical glasses can be stabilized against transmittance loss caused by ionizing radiation by adding cerium to the composition. The extent of stabilization depends on the glass type. In general, the higher the cerium content, the more the glass is stabilized against higher total doses but the more the intrinsic transmittance is reduced. In addition, the impact to the color change by addition of cerium depends on the glass matrix.

Most of the modern technological and telecom lasers work within the 1-2 microns wavelength range. So, the ionizing irradiation affects the transparency of glasses mostly in the wavelength range where the typical lasers do not work.

It should be mentioned that most of the currently operating NPPs have been designed and built 20-40 years ago. During this time, a lot of new radiation-protected technologies have been developed. One techno-cluster that absorbs a lot of new, specially developed technologies is the Large Hadron Collider (which started to work in 2010). These technologies are extreme radiation-resisting plastics, micro-cables, and radiation detectors. These technologies were designed to survive the radiation levels that are equivalent to a 100-megaton nuclear bomb explosion. Now is definitely the time to use them on old and new NPPs.

Generally speaking, all semiconductor devices are very sensitive to ionizing radiation. The attempts to use robots on the Chernobyl NPP failed very fast. The fact that

semiconductor devices irradiated by nuclear bomb ionizing and radio pulses stop operating tens of kilometers around a bomb explosion is well known. However, old electronic bulb devices still survive despite being very close to the epicenter (if not destroyed mechanically). So, two variants are possible: 1) all robot controllers and other electronics units should be located in a protected cabin with a cable connected to the robot engines, or 2) the electronics should be designed with old-fashioned components that are very insensitive to ionizing radiation.

6. Located on-robot

A sensor system for a reconnaissance mobile robot must monitor many environmental parameters; however, in miniature systems, we cannot simply combine several different sensors because of weight, size, and power consumption limitations. Therefore, all available processes and information gathered from sensor-environment interactions should be used for monitoring these different parameters.

The current tendency in the development of technologies for dangerous sites is the application of mobile robots. Such systems are under intensive development in Japan, USA, Canada, China, and the EU. Robots, as "environmental" guards, have some advantages and disadvantages. From one side, having limited "intellectual" abilities, a robot cannot find probabilistic solutions for unexpected problems. On the other hand, a robot has no human characteristics such as panic, fatigue, or narcotic/alcohol dependency that can suppress normal human abilities. The most useful and current application for a robot is as a carrier of sensors with preprocessing of the data. Regular reconnaissance robots may be applied without limitation within a NPP zone. However, to increase the emergency protection of the robots within the core (reactor) and secondary (building and territory) zones, the robots should be designed with high radiation and temperature protection.

A mobile robot with multi-gas sensors and a multifunctional spectrometer on-board is capable of identifying more than a hundred gases, liquids, and solids, locally and remotely. Such a system can be additionally supplied with a non-linear microscope, cameras, rangefinders, a laser-ultrasound scanner, and other techniques for detailed scanning of the environment and atmospheric conditions. This system is under development at several industrial companies and Universities in Canada: 1) Engineering Services, Inc. (Toronto) (ESI, n.d.), University of Toronto (Department of Mechanical and Industrial Engineering) (RAL, n.d.), P&P Optica, Inc. (Waterloo) (P&P Optica, n.d.).

The end-goal is to develop a smart sensory network for environmental monitoring, which is capable of performing tasks not possible by natural sensory-organs, in an effort to increase public and private security (Peshko, 2007; Matharoo, 2010). As the first step in achieving this goal, the design of an integral part of the proposed smart sensor-network: an all-in-one, multi-gas, photonic sensor (for CO, CO_2, CH_4, N_2O, O_2, and H_2O vapor sensing) is provided. The sensory platform also houses independent total-pressure and temperature sensors, infrared, ultraviolet, and γ-ray radiation detectors.

7. Catastrophe simulator: Computer forecasting of processes and events

The problems of continuous reliability and adequateness are apparent in measurements, especially if we need to measure in some unexpected, unpredictable, and unfriendly conditions. Sometimes, occasional combinations of the sensor signals may be interpreted as

an alarm signal, and, sometimes, at really dangerous situations, the alarm system "is sleeping" because an unpredictable interference of the environmental parameters may mask the real event. Thus, modeling the environmental processes together with the reactor's operational processes is necessary. This includes mapping the internal temperature and pressure parameters (dependent on the external ones), radioactive background during the reload process of technological elements and normal standard operation, and other repair/maintenance operations. A simulator will help check a number of situations that may or may not have happened in real life over thousands of years. Additionally, analyzing space-time event-maps and weather conditions will help to remotely identify the event and predict its dynamics. The end-goal is to develop a smart sensor-network for monitoring a nuclear power plant and its surrounding areas to estimate the most probable means that are necessary to predict and prevent catastrophic events.

A full size simulator should include: 1) Modeling of meteorological conditions (in case of an emergency, the area located along the wind path should be alarmed first); 2) Temperature 3D map: the heating/cooling plant model (sun/wind action, reactor operation, air-conditioning operation) should be taken into account – is it an internal source of unexpected heating or are external current factors resulting in local internal heating; what are the amplitudes of possible construction stresses in case of catastrophes); 3) Map of the over-ground and underground rivers, big water reservoirs; 4) Possible action of earthquakes, hurricanes, and other natural phenomena; 5) Modeling of the security system and problems with its operation in case of a catastrophe, cyber or direct terrorist attacks, errors, and emergencies.

8. Reactor zone security monitoring

The general reason for an alarm in any type of the security system is, "something is wrong". This concept is not connected with any specific technology. It is based on pre-calibrated standard scenarios and logical chains of events that typically happen if "everything is right". For example, let us consider the monitoring of personnel motion inside some protected zone:

1. Someone inserts a card key into the (corridor) door (does not matter who as the key may be stolen);
2. The cameras monitoring the door space confirm a moving object (it does not matter who (what) is imaged on monitors, as the security system may be hacked and some recording transferred to the monitors);
3. The motion sensors confirm that something is moving along the corridor;
4. The sound analyzers confirm that the sound spectrum of steps belongs to a person who did open a door (codes of the key), the person is alone, and moves along the way he/she is authorized to walk.

Non-confirmation at any stage of the described chain results in the activation of an alarm. In this case, the most important thing is not the right signal at each stage that may be falsified or not mentioned by security personnel, but the right sequence of actions with some specific signs at each stage.

If no motion is detected by the cameras (comparing pixel information variations, not by motion sensors!) for 20-30 seconds in the security room, it means that the security guards are neutralized or sleeping; an alarm should be activated automatically. This algorithm can be applied in any protected zone: banks, treasures, military sites, and so on.

It is very important that the same logic and the same sensors can be used for NPP safety control.

9. After 9/11

After the events of 9/11, governments are paying more attention to the protection of NPPs. USA's Congressional Research Service published open documents that describe the main requirements for the newly designed plants and propose the means for protection of old operating units. These documents are focused on analyses of NPP vulnerability to terrorist attacks (Holt, 2007) and general problems of NPP security and vulnerability (Holt, 2010). I cite here some key paragraphs from these documents because of their high importance.

"Nuclear plant security measures are designed to protect three primary areas of vulnerability: controls on the nuclear chain reaction, cooling systems that prevent hot nuclear fuel from melting even after the chain reaction has stopped, and storage facilities for highly radioactive spent nuclear fuel. U.S. plants are designed and built to prevent dispersal of radioactivity, in the event of an accident, by surrounding the reactor in a steel-reinforced concrete containment structure.

The Nuclear Regulatory Commission (NRC) approved its final rule amending the design basis threat (DBT) (10 C.F.R. Part 73.1) on January 29, 2007, effective April 18, 2007. Although specific details of the revised DBT were not released to the public, in general the final rule

- clarifies that physical protection systems are required to protect against diversion and theft of fissile material;
- expands the assumed capabilities of adversaries to operate as one or more teams and attack from multiple entry points;
- assumes that adversaries are willing to kill or be killed and are knowledgeable about specific target selection;
- expands the scope of vehicles that licensees must defend against to include water vehicles and land vehicles beyond four-wheel-drive type;
- revises the threat posed by an insider to be more flexible in scope; and
- adds a new mode of attack from adversaries coordinating a vehicle bomb assault with another external assault.

In October 2006, NRC proposed to amend the security regulations and add new security requirements that would codify the series of orders issued after 9/11 and respond to requirements in the Energy Policy Act of 2005. The new security regulations were approved by the NRC Commissioners on December 17, 2008, and published March 27, 2009:

- Safety and Security Interface. Explicit requirements are established for nuclear plants to ensure that necessary security measures do not compromise plant safety.
- Mixed-Oxide Fuel. Enhanced physical security requirements are established to prevent theft or diversion of plutonium-bearing mixed-oxide (MOX) fuel.
- Cyber Security. Nuclear plants must submit security plans that describe how digital computer and communications systems and safety-related networks are protected from cyber attacks.

- Aircraft Attack Mitigative Strategies and Response. As discussed in the earlier section on vulnerability to aircraft crashes, nuclear plants must prepare strategies for responding to warnings of an aircraft attack and for mitigating the effects of large explosions and fires.
- Plant Access Authorization. Nuclear plants must implement more rigorous programs for authorizing access, including enhanced psychological assessments and behavioral observation.
- Security Personnel Training and Qualification. Modifications to security personnel requirements include additional physical fitness standards, increased minimum qualification scores for mandatory personnel tests, and requirements for on-the-job training.
- Physical Security Enhancements. New requirements are intended to ensure the availability of backup security command centers, uninterruptible power supplies to detection systems, enhanced video capability, and protection from waterborne vehicles."

From my point of view, these documents do not pay enough attention to the tendencies of modern weapons. It is much harder to protect a NPP from small, truck-launched weapons than from a big rocket sent from a plane or ship hundreds of kilometers away from the NNP.

It is interesting to note that the problems that took place at Fucushima's reactors (after earthquake) were listed in the NRC documents listed above. So, these documents and the NPP live their own independent life: corporate interests stand higher than the security of the entire country.

Technological monitoring of a power plant includes the control of radiation level (all types of ionizing radiation), temperature, humidity, and some ions and gases that may appear as a result of normal technological process or abnormal situations, for example, CO_2 or CO, C_xH_y, and NO_x in case of a fire. However, the gas monitoring system should also monitor explosive vapours, nerve/blister agents, and other substances that can be used in terrorist actions or during preparing for such actions.

10. Some simple ideas

The key issue is electrical consumption – feeding the security network and coolers for the reactors and burned fuel. The special attention zone is reserve electrical generators and pumps for cooling systems. The recent catastrophe (March 2011) in Japan definitely demonstrated that a reserve generator should be mounted on the damping pyramid with height two times higher than any tsunami or other floating debris potentially affecting the generator. It would be nice to introduce a technique where if the reserve generator is not checked once a year, the NPP should automatically slow down to a safe power minimum, and no NPP personnel, government official, or president of the NPP operating company can turn off the "shut down" option. The best way is to install a generator out of the NPP territory with several power lines going to the NPP by different ways. The cooling loops should be duplicated and triplicated (as much as engineers would decide). It is strongly recommended to have a lot of small pumps instead of fewer high power pumps.

The best option is to build a reservoir of alarm cooling liquid capable of autonomously operating the coolers until the NPP slows down to a safe level.
Every day, on my way to work, I see big tanks of water along the road in each municipality. A relatively low-power pump delivers water to the tank 25-m high and after that, the water runs to consumers without any pumping. So why this extremely simple technology, which was actually developed during the times of ancient Rome, is not used as an emergency reserve cooler that can work until the risk crew reconstructs a source of electricity to support the main pumps' operation?
Analysis of the recent cyber attacks around the world shows that from time to time, higher and higher protected entities, like banks, governments, and big corporations that put in extra efforts to protect their sites and databases, are successfully hacked. It is time to develop special interfaces that have no electrical (wire/wireless) contact between the inside-outside zones of the protected segments of the network.

11. Conclusions

This chapter discusses the principles of development of a Smart Synergistic Security Sensory Network for Harsh Environments: Net4S. It includes an analysis of:

- the technical aspects of operability of the sensors, optical, and electrical telecommunication channels, and computers in the presence of ionizing radiation;
- the influence of environmental parameters on the sensors' accuracy and network operability;
- the development of simulators capable of advising safe solutions based on the analysis of the data acquired by the Net4S; and
- social aspects of the Nuclear Power Plant design, construction, and exploitation.

In total, such a real-time operating network should monitor:

- environmental and atmospheric conditions: chemical, biological, radiological, explosive, and weather hazards;
- climate/man-induced catastrophes;
- contamination of water, soil, food chains, and public health care delivery; and
- large public/industrial/government/military areas.

The end terminals of the system consist of mobile robotic and stationary platforms, equipped with a set of portable environmental sensors that are connected to the monitoring centers. Each sensor should be a self-registering, self-reporting, plug-and-play unit that uses unified electrical and/or optical connectors and operates with the IP communication protocol.
To control the situation inside and outside a NPP, the Sensory Network should monitor several zones:

- core (reactor) area;
- plant building and surrounding territory;
- 10-30 km radius zone; and
- entire country and neighboring territories.

A concept of an "inhomogeneous network" is also introduced. This network combines a set of different types of sensors to measure different parameters (sub-networks), and different types of sensors that measure the same parameter but based on different physical phenomena. The Net4S aims to solve several problems simultaneously:

- the detection and estimation of critical events by a synergistic sensory-network,
- higher reliability of multi-substance sensors based on different operational principles; and
- prediction of critical events based on a history of monitored parameters.

The reactor area should be monitored by a network of local sensors and by the network of remote sensors, in case the core zone is in a state of emergency.

Among different types of sensors, only optical laser sensors can respond immediately and remotely. Such sensors can simultaneously monitor several gases, vapours, and ions with the help of one laser; however, the use of several lasers operating at different wavelengths, dramatically improves accuracy and reliability, and increases the number of monitored substances. A synergistic sensory network can monitor the background optical losses (scattering), environmental pressure, temperature, and humidity.

The Net4S, monitoring a number of parameters inside and outside a Nuclear Power Plant, can serve as the security, safety, and controlling system of the NPP.

The most critical parts of the cooling systems should be self-operable: the water should be delivered from the highly located tank by free running without any pumps.

In total, the security system should identify natural events (hurricane, earthquake, abnormally high or low temperatures and pressures), unauthorized access to the NPP (terrorist attack, hacker's attack) and wrong personnel actions.

12. References

Friebele, E.; Ginther, R.; Sigel Jr. G. (1974). Radiation protection of fiber optic materials: Effects of oxidation and reduction. Applied Physics Letters, Vol.24, No.9 1974 p.412 - 414.

ESI: Engineering Services, Inc. (n.d.). 01.03.2011, Available from www.est.com

RAL: Robotics & Automation Lab, University of Toronto, Department of Mechanical and Industrial Engineering (n.d.). 01.03.2011, Available from www.mie.utoronto.ca/labs/ral

Holt, M; Andrews A. (2007). Nuclear Power Plants: Vulnerability to Terrorist Attack. 01.03.2011, Available from http://www.fas.org/sgp/crs/terror/RS21131.pdf

Holt, M; Andrews, A. (2010). Nuclear Power Plant Security and Vulnerabilities, 01.03.2011, Available from http://www.fas.org/sgp/crs/homesec/RL34331.pdf

Matharoo, I.; Peshko, I.; and Goldenberg, A. (2010). Synergistically-reliable multi-gas photonic sensors for security networks *Proceedings of the Canadian Society for Mechanical Engineering Forum 2010.* Victoria, British Columbia, Canada, 7-9 June, 2010.

P&P Optica, (n.d.) .01.03.2011, Available from www.ppo.ca

Peshko, I. (2007). New-generation security network with synergistic IP-sensors *Proceedings of IEEE, Optics East: Advanced Environmental, Chemical, and Biological Sensing Technologies V* 6755 ed T Vo-Dinh, R A Lieberman and G Gauglitz. Boston, Massachusetts, USA, 9-12 Sep 2007.

SCHOTT Optical Glass Pocket Catalogue (2007). 01.03.2011, Available from http://www.schott.com/advanced_optics/english/download/tie-42_radiation_resistant_glasses.pdf

Sigel Jr, G.; and D. Evans, B. (1974). Effects of ionizing radiation on transmission of optical fibers. *Applied Physics Letters,* Vol. 24, No. 9, (1 May 1974), p.410-412.

Smith, H.; Cohen, A.. (1964). Color Centers in X-Irradiated Soda-Silica Glasses. *Journal of The American Ceramic Society* Vol. 47, No. 11, p.564-570.

6

An Approach to Autonomous Control for Space Nuclear Power Systems

Richard Wood and Belle Upadhyaya
Oak Ridge National Laboratory & The University of Tennessee, United States of America

1. Introduction

Under Project Prometheus, the National Aeronautics and Space Administration (NASA) investigated deep space missions that would utilize space nuclear power systems (SNPSs) to provide energy for propulsion and spacecraft power. The initial study involved the Jupiter Icy Moons Orbiter (JIMO), which was proposed to conduct in-depth studies of three Jovian moons. Current radioisotope thermoelectric generator (RTG) and solar power systems cannot meet expected mission power demands, which include propulsion, scientific instrument packages, and communications. Historically, RTGs have provided long-lived, highly reliable, low-power-level systems. Solar power systems can provide much greater levels of power, but power density levels decrease dramatically at ~1.5 astronomical units (AU) and beyond. Alternatively, an SNPS can supply high-sustained power for space applications that is both reliable and mass efficient.

Terrestrial nuclear reactors employ varying degrees of human control and decision-making for operations and benefit from periodic human interaction for maintenance. In contrast, the control system of an SNPS must be able to provide continuous operation for the mission duration with limited immediate human interaction and no opportunity for hardware maintenance or sensor calibration. In effect, the SNPS control system must be able to independently operate the power plant while maintaining power production even when subject to off-normal events and component failure. This capability is critical because it will not be possible to rely upon continuous, immediate human interaction for control due to communications delays and periods of planetary occlusion. In addition, uncertainties, rare events, and component degradation combine with the aforementioned inaccessibility and unattended operation to pose unique challenges that an SNPS control system must accommodate. Autonomous control is needed to address these challenges and optimize the reactor control design.

1.1 State of the technology

To support JIMO development, Oak Ridge National Laboratory (ORNL) and the University of Tennessee (UT) conducted an investigation of autonomous control. Overviews of autonomous control characteristics, capabilities, and applications were found that establish the existing experience and current technology readiness (Antsaklis & Passino, 1992; Astrom, 1989; Chaudhuri et al., 1996; Passino, 1995; Zeigler & Chi, 1992; Basher & Neal, 2003). The desirable characteristics of autonomous control include intelligence, robustness, optimization, flexibility, adaptability, and reliability.

Control systems with varying levels of autonomy have been employed in robotic, transportation, spacecraft, and manufacturing applications. However, autonomous control has not been implemented for an operating terrestrial nuclear power plant, and there has not been any experience beyond automating simple control loops for space reactors. Current automated control technologies for nuclear power plants are reasonably mature, and basic control for an SNPS is clearly feasible under optimum circumstances. Autonomous control is primarily intended to account for non-optimum circumstances when degradation, failure, and other off-normal events challenge the performance of the reactor, and near-term human intervention is not possible. There are clear gaps in the development and demonstration of autonomous control capabilities for the specific domain of nuclear power operations.

1.2 Advanced control in nuclear power applications

In the nuclear power industry, single-input, single-output classical control has been the primary means of automating individual control loops. The use of multivariate control, such as three element controllers for steam generators, has been employed in some cases. In a few cases, efforts were made to coordinate the action of individual control loops, based on an overall control goal, and extend the range of automated control.

The application of most advanced techniques for nuclear power control has primarily been the domain of universities and national laboratories. Some of the techniques employed in controls research for both power and research reactors include adaptive robust control for the Experimental Breeder Reactor II (EBR-II), fuzzy logic control for power transition, H-infinity control and genetic-algorithm based control for steam generators, neural network control for power distribution in a reactor core, and supervisory control for the multi-modular advanced liquid-metal reactor (ALMR). Proceedings of past American Nuclear Society (ANS) International Topical Meetings on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies provide a useful compendium of findings from such research activities (ANS, 1993, 1996, 2000, 2004, 2006, 2009, 2010). In recent research for the U.S. Department of Energy (DOE), UT demonstrated aspects of near-autonomy for a representative SNPS design through the development of a model-predictive controller using a simulation of the SP-100 reactor system (Upadhyaya et al., 2007; Na & Upadhyaya, 2007). The approach demonstrated the fault-tolerance and reconfiguration features of the control strategy.

1.3 Autonomy in space exploration

NASA has pursued autonomy for spacecraft and surface exploration vehicles (e.g., rovers) to reduce mission costs, increase efficiency for communications between ground control and the vehicle, and enable independent operation of the vehicle during times of communications blackout. For rovers, functional autonomy addresses navigation, target identification, and science package manipulation. For spacecraft, functional autonomy has focused on automated guidance, navigation, and control.

Autonomy for rovers has progressed during the last decade with prominent examples from efforts to explore the surface of Mars. The Mars Pathfinder rover, Sojourner, explored the Martian terrain beginning in July 1997 (Mishkin et al., 1998). The Sojourner had very limited autonomy to enable navigation and provide for resource management and contingency

response. Because it only provided supervised autonomy, repetitive ground monitoring was required. In January 2004, Spirit and Opportunity, the twin Mars Exploration Rovers (MERs), began a surface exploration mission that has continued into 2011. These rovers employ expanded autonomy over what was feasible for Sojourner and provide model-based recovery, resource management, and autonomous planning capabilities in addition to autonomous obstacle detection and navigation. The integration software architecture used to facilitate MER autonomy is the "Coupled Layer Architecture for Robotic Autonomy" or CLARAty (Volpe et al., 2001). CLARAty provides a dual-layer architecture consisting of a decision layer for artificial intelligence (AI) software and a functional layer for controls implementations. Implicit granularity in each layer allows for a functional hierarchy with nested capabilities.

Spacecraft autonomy has been demonstrated with the Deep Space 1 mission. Deep Space 1 was launched in October 1998 as a test platform to validate high-risk advanced technologies in space (Rayman et al., 1999). In support of autonomous navigation of the spacecraft, a principal experiment involved demonstration of the remote agent (RA) AI system for on-board planning and execution of spacecraft activities.

2. Autonomous control functional definition

2.1 The nature of autonomy

There is a distinction between automated control and autonomous control. Consideration of the Greek root words illustrates the difference. *Automatos* means self-acting, while *autonomos* means independent. Similarly, automated control involves self-action, while autonomous control involves independent action. Autonomous control implies an embedded intelligence. Although automation includes at least a limited inherent authority within the control system, automated control often consists of straightforward automatic execution of repetitive basic actions. It is clear that autonomous control encompasses automated control.

Automated control provides control actions that result from a fixed set of algorithms with typically limited global state determination. As a result, automated control is often implemented as rigidly defined individual control loops rather than as fully integrated process/plant control. Although automated control requires no real-time operator action for normal operational events, most significant decision-making is left to the human rather than being incorporated as part of the control system. In contrast, autonomous control integrates control, diagnostic, and decision capabilities. A flexible functional architecture provides the capability to adapt to evolving conditions and operational constraints and even support self-maintenance over the control system lifetime. While automated control is common in numerous applications, autonomous control is more difficult to achieve, and the experience base is very limited.

Autonomy extends the scope of primary control functions. Such capabilities can consist of automated control during all operating modes, process performance optimization (for example, self-tuning), continuous monitoring, and diagnosis of performance indicators as well as trends for operational and safety-related parameters, diagnosis of component health, flexible control to address both anticipated and unanticipated events and to provide protection of life-limited components (such as batteries and actuators), adaptation to changing or degrading conditions, and validation and maintenance of control system performance.

Key characteristics of autonomy include intelligence, robustness, optimization, flexibility, and adaptability. Intelligence facilitates minimal or no reliance on human intervention and can accommodate an integrated, whole system approach to control. It implies embedded decision-making and management/planning authority. Intelligence in control provides for anticipatory action based on system knowledge and event prediction. To support control and decision, real-time diagnostic/prognostic capabilities are important for state identification and health/condition monitoring. Additionally, self-validation is an aspect of intelligence that addresses data, command, and system performance assessment and response.
In addition to providing an environmentally rugged implementation, robustness is addressed by accounting for design uncertainties and unmodeled dynamics. Fault management is an important consideration in achieving robustness. Fault management involves techniques such as fault avoidance, fault removal, fault tolerance, and fault forecasting. Robustness can also involve self-maintenance or self-healing. This capability is promoted through means such as captured design knowledge and self-correcting features, prognostics to identify incipient failure, and fault detection and isolation.
Optimization implies rapid response to demands, minimal deviation from target conditions, and efficient actuator actions. Optimized control can be facilitated by self-tuning and other forms of adaptation. Flexibility and adaptability are enabled by diverse measurements, multiple communication options, and alternate control solutions. Functional reconfigurability facilitates the effective use of these systems options, while an inherent redesign capability permits adaptation to unanticipated conditions.
The characteristics discussed above represent the possibilities of autonomy, but they do not constitute a necessary set. Therefore, autonomous control can be viewed as providing a spectrum of capabilities with automated control representing the lowest extreme or baseline of the continuum. The incorporation of increasing intelligence and fault tolerance moves the control capabilities further along the spectrum. The higher degrees of autonomy are characterized by greater fault management, more embedded planning and goal setting, and even self-healing. The realization of full autonomy involves learning, evolving, and strategizing independent of human interaction or supervision.

2.2 Space nuclear power mission challenges

The space reactor control and protection paradigm is different from conventional terrestrial reactor applications. Terrestrial reactors have relied upon immediate interaction from readily available human resources. This includes varying degrees of direct human control and decision-making for operations and periodic human intervention for maintenance and refurbishment. In contrast, the SNPS control system must be able to provide continuous, remote, often-unattended operation for a mission lasting up to a decade or more. Because of communication delays and blackouts, immediate human interaction for continuous operational supervision and event management is not feasible. This isolation drives the need for a high level of autonomy. Because of launch considerations, size and mass constraints significantly limit the options for redundancy and diverse systems. This limitation drives the need for functional and environmental robustness. Because of distance from the Earth for deep space or planetary surface power applications, maintenance and refurbishment are improbable if not impossible. This inaccessibility drives the need for long-life dependability. Also, because of the critical and time-sensitive nature of some spacecraft maneuvers, space reactor power must be available on demand. This operational imperative, coupled with the likelihood that an SNPS restart capability will be unavailable, requires that reactor scrams as

a response to plant events must be minimized if not completely avoided. Thus, unlike the design criteria for terrestrial nuclear power plants, mission assurance must be emphasized over reactor protection for space reactors.

Automated control can provide the necessary automation of normal operational control to permit ground control personnel to assume a supervisory role for an SNPS rather than taking on direct, active control responsibilities. This control capability for full-power range maneuverability, including startup, has been developed and demonstrated for terrestrial reactors (Winks et al., 1992). However, considering the challenges of a deep space mission, the SNPS control system must do more than provide automated control of normal operational activities. In addition to automation, the control system must provide a level of autonomy that can detect, diagnose, and adapt to evolving conditions (e.g., failures or degradation) as well as rapidly respond to anticipated events without requiring a reactor scram. As a result, autonomous control is clearly necessary to ensure the successful application of an SNPS for deep space missions. This conclusion is based on the understanding that autonomous capabilities permit the SNPS control system to satisfy essential control objectives under significant uncertainties, disturbances, and degradation without requiring any human intervention. In a sense, the role of the autonomous control system is to act as an extension of and occasional proxy for the ground-based human operators to ensure reliable, continuous operation of an SNPS over an extended lifetime under adverse conditions.

2.3 Space Nuclear Power System control

Autonomous control functions of an SNPS can be defined based on the expected operational modes, which include startup, normal power operation, reactor protection, contingent operation, and end-of-mission shutdown. As a minimum requirement of autonomy, the SNPS control system must be able to switch between normal operational modes automatically (i.e., automated control). Additionally, reactor protective action must be available if the desired operational conditions cannot be achieved.

The phases of power operation include power ascension, steady state power and load following, and power reduction. Under normal conditions, power operation can be relatively simple, with inherent feedback effects serving to maintain stability and provide the means for load following in response to minor fluctuations. Thermal load transients (e.g., turbine failure) can be treated as off-normal events. Other off-normal events include design uncertainties, load/power interruptions, control element jamming, control motor burnout, control linkage failure, unintended control element motion, actuator signal interruption or interference, heat rejection system degradation or damage, control processor fault, rare-event software error, sensor failure, sensor signal interruption or interference, sensor drift, signal conditioning electronics drift, sensor noise increase, and communication failures or retransmissions. The most likely protective action would consist of a rapid power runback. Contingent operation occurs when SNPS operation may be restricted either because of environmental limitations, such as an abnormal thermal environment, or because of power system limitations, such as component failures.

The response to off-normal events is where autonomy becomes especially relevant. The autonomous response includes a reflexive element and a deliberative element. The first element addresses reactor protection. Unlike terrestrial reactors, in which the primary defense against potentially adverse conditions resulting from off-normal events is to scram the reactor, it is quite likely that an SNPS must not shut down until the end of the

mission because a restart capability may not be feasible. Thus, reactor protection is provided through diversity and defense-in-depth to anticipate potential challenges to power operation. A limitation system is one means of protecting the reactor. This is accomplished by defining acceptable operational regimes and overriding control actions that would drive the reactor conditions to violate the limitation boundaries. In effect, the limitation system acts as a bounding system whose primary purpose is to provide a check against operations outside of analyzed conditions. The principal response of the limitation system would be to run back the reactor power to assume a safe low-power condition when necessary. Because of the operational imperative that power must remain available during critical spacecraft actions for deep space missions, the SNPS control system must provide the capability to switch out the protection element (or at least expand the operational boundaries it maintains) on demand from the spacecraft or mission control.

The second element of the response to off-normal events addresses mission assurance. The deliberative nature (i.e., determination and decision) of this element contributes the most relevant attribute of autonomous control that distinguishes it from conventional automation. In the operational control context, the autonomous control functionality involves detection and immediate response to degraded or failure conditions. Fault management is a crucial part of this element of autonomous control; it provides for detection, diagnosis, and adaptation (or reconfiguration) given changing SNPS conditions. An additional aspect of the deliberative element is the monitoring, diagnosis, and validation of control system and SNPS performance. Through this capability, the SNPS control system is able to identify incipient events (transients or failures) for anticipatory rather than reactionary action, determine measures to protect life-limited or vulnerable components, and ensure continued dependable operation of the power plant.

As noted, the SNPS autonomous control functionality revolves around automated control for normal operational modes. In essence, the primary function of the control system is command generation to achieve the desired operational state. Additional functionality to support confirmation of control system performance includes features such as command verification, control coordination with interconnected systems, and strategy enforcement. Mechanisms for implementing these features can involve multiple diverse algorithms for comparison with the principal controller command, inclusion of feedforward action or some representation of unmodeled dynamics (e.g., exogenous variables) in control algorithms, event management according to predetermined sequences of events, and adaptation of the control strategy.

Performance management as part of the SNPS autonomous control functionality involves continuously assessing the condition of the control system and the SNPS to identify when predetermined adjustments to the controller should be invoked. The needed assessments include monitoring control system effectiveness, identifying the dynamic state of the SNPS, and determining the condition of key components. Methods that can be employed are state estimation algorithms, process system diagnostics, component condition monitoring, and control parameter adaptation.

Data management and communications are related capabilities with traditional and autonomous functionality intended to support autonomy and system integration. Data acquisition and signal processing methods provide the data needed for control and monitoring, while signal validation adds information about data quality. For communications, the functional elements include device-level data and control signals,

system-level information and commands, and spacecraft-level status and demands. The effective integration of data and information at each level requires a well-defined functional architecture with a capable physical infrastructure that supports reliable, timely information flow.

Desired functionality for fault management includes detection and identification of field device faults, change tracking for system parameters, detection of off-normal transients and identification of anticipated events, and configuration control. Field device monitoring can be accomplished through model-based and/or data-driven algorithms. Parameter tracking can involve empirical models or first principles estimation. Each capability can be used to facilitate an adjustable system dynamic model that can be used for fault prediction or control system performance validation. Configuration control functions are needed to manage transitions among predefined control strategies or algorithms for the autonomous control system. This is essential for effective fault recovery.

To illustrate the autonomous functionality that can be provided for the SNPS control system, two fault management scenarios are considered in which detection and response are described. The first scenario relates to fault adaptation in the case of sensor failure. The indicators from surveillance and diagnostic functions that the SNPS control system can employ include divergence of redundant measurements, conflict between predicted (based on analytical or relational estimation) and measured values, and detection and isolation of a confirmed fault. The prospective response can include substitution of a redundant measurement or utilization of a diverse measurement. An example of the latter would be using neutron flux instead of temperature (i.e., core thermal power) as a power measurement. Switching to an alternate control algorithm may prove necessary for faulted or suspect measurements.

The second scenario relates to fault avoidance in the case of a degrading actuator. The indicators of an incipient failure can be prediction of actuator failure based on prognostic modeling (e.g., fault forecasting) or detection of sluggish response to commands. The prospective response can be to switch to an alternate control strategy to avoid incipient failure by reducing stress on the suspect component. An example would be utilizing manipulation of core heat removal (e.g., coolant density change) instead of direct reactivity insertion (e.g., control element movement) to control reactor power.

2.4 Enabling autonomous control

Autonomous control must be addressed early in the design of an SNPS to determine the degree of autonomy required. Mission requirements, technology readiness, design trade-offs, and resource constraints will affect the autonomous capabilities to be included. The extent to which the key characteristics of autonomy are realized depends on the level of responsibility that is to be entrusted to the autonomous control system and the degree of mission risk that the autonomous control system must mitigate.

Several factors can influence the degree of autonomy selected for an SNPS control system. These factors include the potential for human interaction (which is physically limited but also may be practically limited due to the economics of maintaining a ground-based team), performance goals, complexity of system demands, technological constraints, mission risk considerations, and the balance between simplicity (i.e., reliability) and complexity (i.e., the capacity to detect and adapt). The trade-off between reliability and mission assurance profoundly affects the level of autonomy employed for SNPS control. While having a highly reliable SNPS control system is important, that fact is meaningless if it cannot accommodate

SNPS degradation. In such a case, the result is a highly reliable control system that becomes useless because the plant has changed.
Finally, as previously described, the experience base for autonomous control is not deep. In particular, autonomous control has not been implemented for an operating terrestrial nuclear power plant. The technology gaps indicated by investigation of the state of the technology for reactor control in general and autonomous control in particular suggest research, development, and demonstration (RD&D) activities that need to be accomplished to fully realize the goal of autonomous control for an SNPS. Key elements of the needed RD&D effort involve establishing a suitable functional architecture, developing foundational modules to support autonomy, and demonstrating the integrated application of autonomous capabilities.

3. Functional architecture for autonomous control

3.1 Architectural approaches

As observed from examples of autonomous control for nuclear and space applications, the principal functional architectures that have been employed, in most cases, involve some form of hierarchical framework with varying distributions of intelligence.
A three-level hierarchy is typical for robotic applications (Antsaklis & Passino, 1992; Alami et al., 1998; Gat, 1998). The general concept of the hierarchy is that commands are issued by higher levels to lower levels, and response data flows from lower levels to higher levels in the multi-tiered framework. Intelligence increases with increasing level within the hierarchy. Each of the three interacting tiers has a principal role. Basically the functional layer provides direct control, the executive layer provides sequencing of action, and the planner layer provides deliberative planning.
As previously described, an autonomous control architecture, based on the CLARAty software environment, was developed to support the MER mission. The CLARAty dual-layer architecture provides an upper (decision) layer for AI software and a lower (functional) layer for controls implementations. The development of CLARAty addresses perceived issues with the three-tiered architecture (Volpe et al., 2001). Those issues are the tendency toward a dominant level that depends on the expertise of the developer, the lack of access from the deliberative or planner level to the control or functional level, and the difficulty in representing the internal hierarchy of each level (e.g., nested subsystems, trees of logic, and multiple time lines and planning horizons) using this representation. In one sense, the CLARAty architecture collapses the planner and executive levels, which are characterized by high levels of intelligence, into the decision layer. Essentially, the deliberative and procedural functionalities are merged into an architectural layer that parallels the functional layer and provides a common database to support decision-making. Additionally, a system granularity dimension is maintained to explicitly represent the system hierarchies of the functional layer and the multiple planning horizons of the decision layer.
The functional layer is an object-oriented hierarchy that provides access to the capabilities of the plant/system hardware and serves as the interface for the decision layer to the subject (robot, spacecraft, plant) under control. The interaction between the two layers depends on the relative granularity of each layer at the interface. At lower granularity, the decision layer has almost direct access to the basic capabilities of the plant/system. At higher granularity, the decision layer provides high-level commands that are broken down and executed by the

intelligent control capability of the functional layer. The decision layer provides functionality to break down goals into objectives, establish a sequential task ordering based on the plant/system state and known constraints, and assess the capability of the functional layer to implement those commands. At lower granularity within the decision layer, executive functions such as procedure enforcement are dominant while, at higher granularity, planning functions such as goal determination and strategy development are dominant.

There is an architectural approach for nearly autonomous control systems that have been applied through simulated nuclear power applications (see Fig. 1). As part of research into advanced multi-modular nuclear reactor concepts, such as the International Reactor Innovative and Secure (IRIS) and the ALMR, a supervisory control system architecture was devised (Wood et al., 2004). This approach provides a framework for autonomous control while supporting a high-level interface with operations staff, who can act as plant supervisors. The final authority for decisions and goal setting remains with the human, but the control system assumes expanded responsibilities for normal control action, abnormal event response, and system fault tolerance. The autonomous control framework allows integration of controllers and diagnostics at the subsystem level with command and decision modules at higher levels.

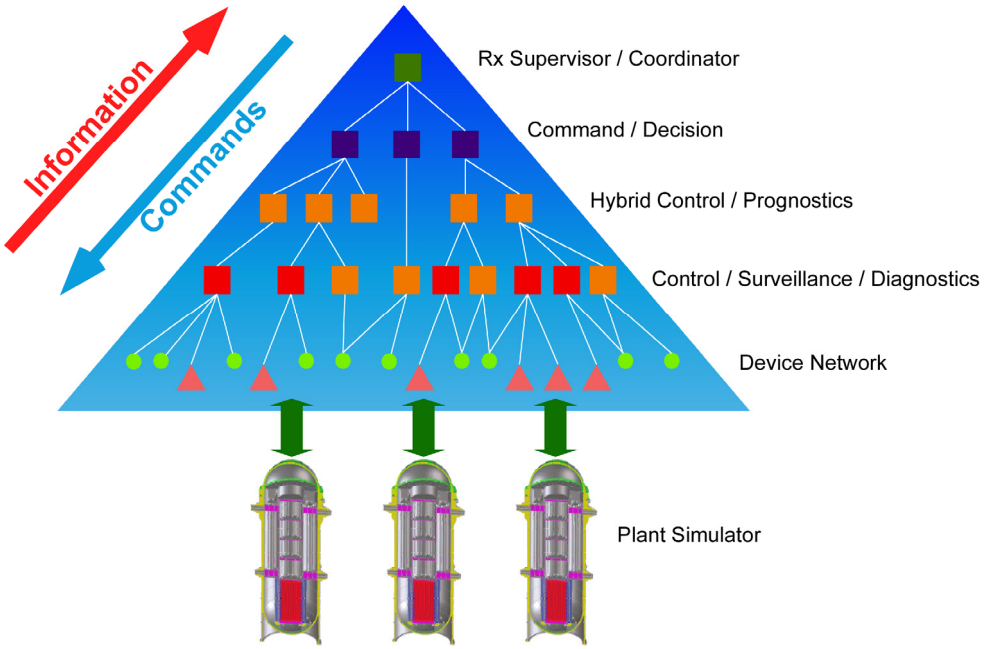


Fig. 1. Supervisory control architecture for multi-modular nuclear power plants

The autonomous control system architecture is hierarchical and recursive. Each node in the hierarchy (except for the terminal nodes at the base) is a supervisory module. The

supervisory control modules at each level within the hierarchy respond to goals and directions set in modules above it and to data and information presented from modules below it. Each module makes decisions appropriate for its level in the hierarchy and passes the decision results and necessary supporting information to the functionally connected modules.

The device network level consists of sensors, actuators, and communications links. The next highest level consists of control, surveillance, and diagnostic modules. The coupling of the control modules with the lower-level nodes is equivalent to an automated control system composed of controllers and field devices. The surveillance and diagnostic modules provide derived data to support condition determination and monitoring for components and process systems. The hybrid control level provides command and signal validation capabilities and supports prognosis of incipient failure or emerging component degradation (i.e., fault identification). The command level provides algorithms to permit reconfiguration or adaptation to accommodate detected or predicted plant conditions (i.e., active fault tolerance). For example, if immediate sensor failure is detected by the diagnostic modules and the corresponding control algorithm gives evidence of deviation based on command validation against pre-established diverse control algorithms, then the command module may direct that an alternate controller, which is not dependent on the affected measurement variable, be selected as principal controller. The actions taken at these lower levels can be constrained to predetermined configuration options implemented as part of the design. In addition, the capability to inhibit or reverse autonomous control actions based on operator commands can be provided. The highest level of the autonomous control architecture provides the link to the operational staff.

3.2 Framework for autonomous control functionality

A variation on the nuclear plant supervisory control architecture and the CLARAty architecture for microrovers seems appropriate for consideration as the framework to support autonomy for an SNPS control system. Figure 2 illustrates the concept. Essentially, the approach of a hierarchical distribution of supervisory control and diagnostic functionality throughout the control system structure is adopted, while the overlaid decision functionality is maintained. It is possible to blend the decision and functional layers for this application domain because the planning regime for nuclear power system operation is much more restricted than for robotic or spacecraft applications. For example, while there are a multitude of paths that a robot may traverse as it navigates to its next site, the states are allowed for an SNPS are much more constrained. Even in the event of transients or faults, the control system will try to drive the plant back to a known safe state. This compression of the dual layers into a truncated three-sided pyramid allows for a deeper integration of control, diagnostics, and decision to provide the necessary capability to respond to rapid events and to adapt to changing or degraded conditions.

The granularity dimension is retained with more complexity shown at the lower hierarchical levels. Additionally, the information and command flow reflects granularity as well. At lower granularity, volumes of data are present. As the granularity increases moving up the hierarchy, the data are processed into system state and diagnostic/prognostic information that are subsequently refined into status and indicator information. On the command side, the transition from the top is demands to commands to control signals with the resolution of the plant/system control growing increasingly more detailed.

As with the supervisory control architecture, the bottom two levels of the hierarchy are the equivalent of an automated control system. The embedded functionality that enables a reliable, fault-tolerant implementation is indicated as a base intelligence. It is expected that there will be some decision capability associated with the control/surveillance/diagnostics level of that baseline system. The higher levels of the hierarchy assume greater degrees of decision capabilities.

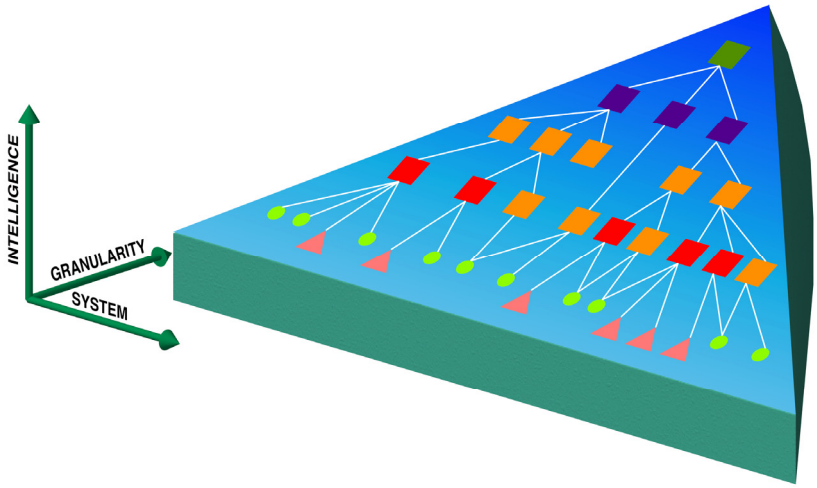


Fig. 2. Hierarchical framework to support SNPS control system autonomy

In addition to managing the communications within the hierarchy, the autonomous control system must coordinate with the spacecraft control system and keep the mission control staff informed. To this end, the reactor supervisor/coordinator node must communicate information about the status of the SNPS and the control system and also receive directives and commands. The information provided by the supervisor node can include SNPS operational status and capability (e.g., constraints due to degradation), control action histories, diagnostic information, self-validation results, control system configuration, and data logs. Additional communication outside of the hierarchy may be required to coordinate control actions with other segments of the spacecraft, such as the power conversion system.

The functionality that is embodied in the hierarchy can be decomposed into several elements. These include data acquisition, actuator activation, validation, arbitration, control, limitation, checking, monitoring, commanding, prediction, communication, fault management, and configuration management. The validation functionality can address signals, commands, and system performance. The arbitration functionality can address redundant inputs or outputs, commands from redundant or diverse controllers, and status indicators from various monitoring and diagnostic modules. The control functionality includes direct plant or system control and supervisory control of the SNPS control system itself. The limitation functionality involves maintaining plant conditions

within an acceptable boundary and inhibiting control system actions. The checking functionality can address computational results, input and output consistency, and plant/system response. The monitoring functionality includes status, response, and condition or health of the control system, components, and plant, and it provides diagnostic and prognostic information. The commanding functionality is directed toward configuration and action of lower level controllers and diagnostic modules. The prediction functionality can address identification of plant/system state, expected response to prospective actions, remaining useful life of components, and incipient operational events or failures. The communication functionality involves control and measurement signals to and from the field devices, information and commands within the control system, and status and demands between the SNPS control system and spacecraft or ground control. The fault management and configuration management functionalities are interrelated and depend on two principal design characteristics. These are the ability of the designer to anticipate a full range of faults and the degree of autonomy enabled by the control system design.

Finally, the distribution of functions throughout the hierarchy must be established based on the degree of autonomy selected, technology readiness, reliability and fault management considerations, software development practices and platform capabilities, and the physical architecture of the SNPS control system hardware. Because an autonomous control system has never been implemented for a nuclear reactor and because several functional capabilities remain underdeveloped (as seen in the overview of the state of the art), there is clearly a critical need for further development and demonstration of a suitable architectural framework.

4. Application of model-based control to Space Nuclear Power Systems

Key functionality that is necessary to establish the basis for autonomous control has been demonstrated through a simulated space reactor application under university research sponsored by DOE. These capabilities related to control elements within the lower layers of the functional hierarchy. Specifically, the research conducted at UT involved development of a highly fault tolerant power controller for the SP-100 space power reactor design (Upadhyaya et al., 2007; Na & Upadhyaya, 2007).

The SP-100 design provides for a fast spectrum, lithium-cooled fuel pin reactor coupled with thermo-electric converters (TE) with the waste heat removed through a heat pipe distribution system and space radiators. The TE generator output is rated at 112 kW, with a nominal reactor thermal power 2000 kW.

A lumped parameter simulation of a representative SNPS was developed based on physics models specific to the SP-100 reactor, which were derived in prior academic work at the University of New Mexico (El-Genk & Seo, 1987). The reactor system modules include a model of reactor control mechanism, a neutron kinetics model, a reactor core heat transfer model, a primary heat exchanger (HX) model, and a TE conversion model. Figure 3 illustrates the elements of the SNPS model. The integrated SP-100 SNPS model was assembled through an iterative algorithm. The model involves both nonlinear ordinary differential equations and partial differential equations. The code development was performed under the MATLAB™/SIMULINK™ environment. The SNPS simulation provided the demonstration platform for the fault tolerant controller development.

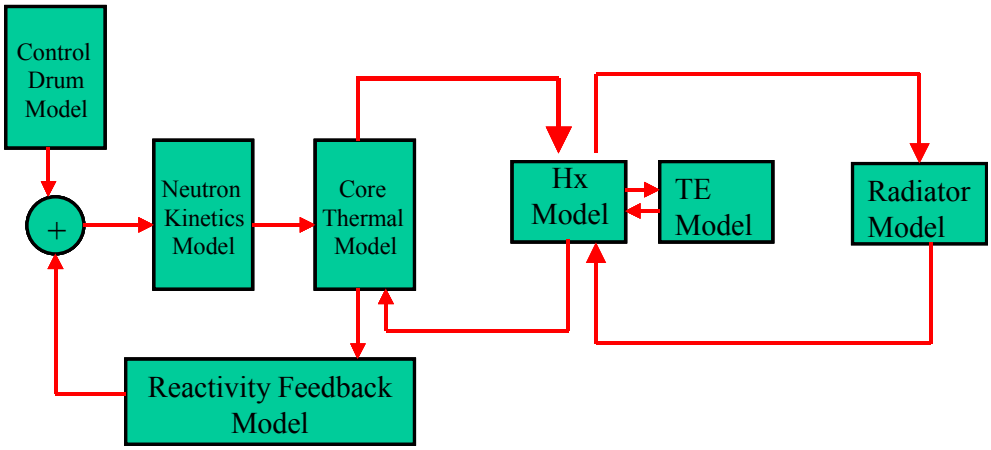


Fig. 3. Schematic of the model development of the SP-100 reactor system

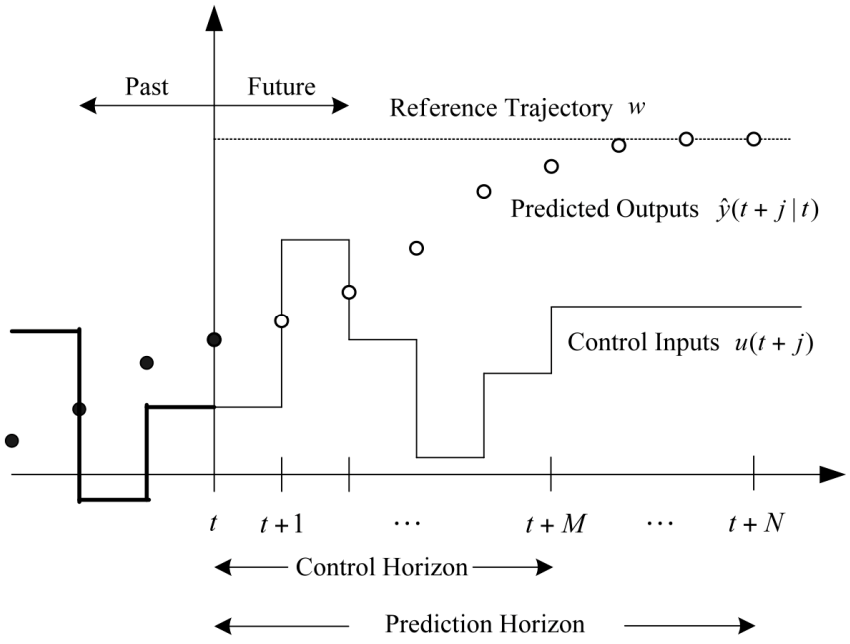


Fig. 4. Basic concept of a model predictive control method

The control approach adopted is a model-predictive controller (MPC) design. The basic concept of the model-predictive control method is illustrated in Fig. 4. The MPC

minimizes a quadratic cost function and takes into consideration any constraints imposed on the control action and the state variables. For a given set of present and future control actions, the future behavior of the state variables are predicted over a prediction horizon *N*, and *M* present and future control moves ($M \le N$) are computed to minimize the quadratic objective function. Out of the *M* control moves that are calculated, only the first control action is implemented. The prediction feature of the controller has an anticipatory effect, and is reflected in the current control action. These calculations are repeated in the next time step by appending the next measurement to the database. The new measurements compensate for the unmeasured disturbances and model inaccuracies, both of which result in the measured system output being different from that predicted by the model. The MPC requires the on-line solution of an optimization problem to compute optimal control inputs over the time horizon. The MPC calculates a sequence of future control signals by minimizing a multi-stage cost function defined over a prediction horizon.

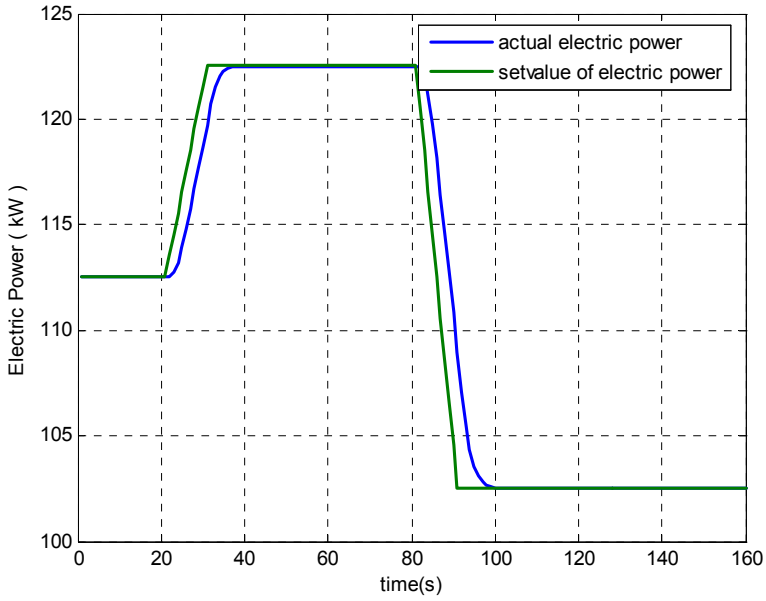
The performance index for deriving an optimal control input is represented by the quadratic objective function given in Eq. (1).

$$J = \frac{1}{2}\sum_{j=1}^{N} Q\left[\hat{y}(t+j\,|\,t) - w(t+j)\right]^2 + \frac{1}{2}\sum_{j=1}^{M} R\left[\Delta u(t+j-1)\right]^2 , \tag{1}$$

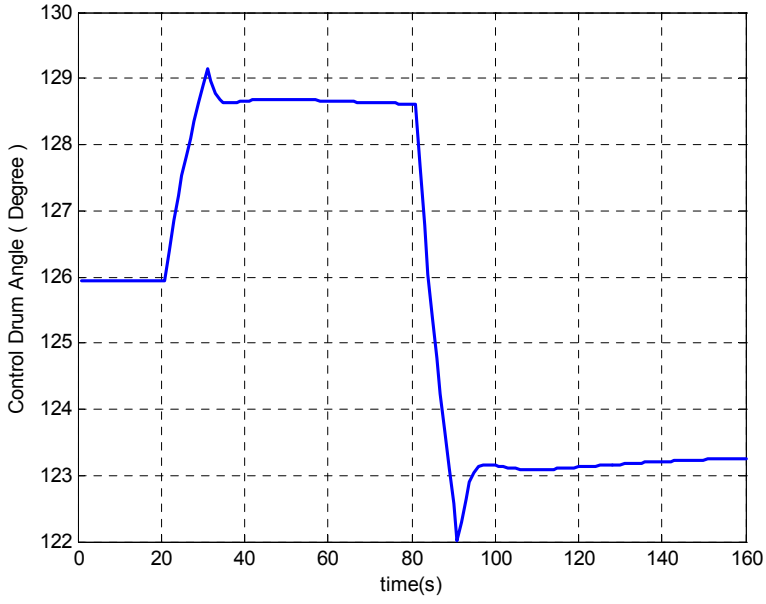
$$\text{subject to constraints} \quad \begin{cases} \Delta u(t+j-1) = 0 \ \text{ for } \ j > M , \\ u_{\min} \le u(t) \le u_{\max} , \\ \left|\Delta u(t)\right| \le \Delta u_{\max} . \end{cases}$$

where Q and R are the weights for the TE generator power (system output) error and the SP-100 control drum angle (reactivity as control input) change between time steps at certain future time intervals, respectively, and w is a set point (desired generator power). The estimate $\hat{y}(t+j\,|\,t)$ is an optimum j-step-ahead prediction of the system output (TE generator power) based on data up to time t; that is, the expected value of the output at time t as a function of the past input and output and the future control sequence are known. N and M are the prediction horizon and the control horizon, respectively. The prediction horizon represents the limiting time for the output to follow the reference sequence. In order to obtain control inputs, the predicted outputs are first calculated as a function of past values of inputs and outputs. The constraint, $\Delta u(t+j-1) = 0$ for $j > M$, indicates that there is no variation in the control signal after a certain time interval M < N, where M is the control horizon. $u_{\min}$ and $u_{\max}$ are the minimum and maximum values of input, respectively, and $\Delta u_{\max}$ is a maximum allowable control perturbation per time step.

The applicability and the effectiveness of the MPC approach were demonstrated through its simulated performance for several operational scenarios, including under degraded or ill-characterized conditions (Upadhyaya et al., 2007). The effectiveness of the MPC controller for tracking the TE power output is illustrated in Figure 6. Figure 6a shows the TE converter set point profile and the actual TE generator power. The corresponding reactivity changes (drum angle variations) are shown in Figure 6b.



(a)



(b)

Fig. 6. (a) Electric power (TE) set point profile and the controller performance. (b) Controller response (i.e., reactivity control) in terms of the drum angle

The MPC approach was shown to provide a fast response and robustness under changing system conditions. Specifically, fault tolerance and reconfigurability features of the control approach were demonstrated in response to sensor faults, drum actuator anomalies, and changes in model parameters (Upadhyaya et al., 2007; Na & Upadhyaya, 2007). Consequently, it is observed that several of the capabilities and characteristics that are necessary to enable autonomous control are provided by the MPC approach.

5. Conclusion

The control system for an SNPS will be subject to unique challenges as compared to terrestrial nuclear reactors, which employ varying degrees of human control and decision-making for operations and benefit from periodic human interaction for maintenance. In contrast, the SNPS control system must be able to provide continuous, remote, often unattended operation for a mission lasting a decade or more with limited immediate human interaction and no opportunity for hardware maintenance. In addition to the inaccessibility and periods of unattended operation, the SNPS control system must accommodate severe environments, system and equipment degradation or failure, design uncertainties, and rare or unanticipated operational events during an extended mission life. As a result, the capability to respond to rapid events and to adapt to changing or degraded conditions without near-term human supervision is required to support mission goals. Autonomous control can satisfy essential control objectives under significant uncertainties, disturbances, and degradation without requiring any human intervention. Therefore, autonomous control is necessary to ensure the successful application of an SNPS for deep space missions.

Key characteristics that are feasible through autonomous control include

- intelligence to confirm system performance and detect degraded or failed conditions,
- optimization to minimize stress on SNPS components and efficiently react to operational events without compromising system integrity,
- robustness to accommodate uncertainties and changing conditions, and
- flexibility and adaptability to accommodate failures through reconfiguration among available control system elements or adjustment of control system strategies, algorithms, or parameters.

Autonomous control must be addressed early in the design of an SNPS to determine the degree of autonomy required. Mission requirements, design trade-offs, and the state of the technology will affect the autonomous capabilities to be included. The extent to which the key characteristics of autonomy are realized depends on the level of responsibility that is to be entrusted to the autonomous control system. Given anticipated mission imperatives to utilize technology with demonstrated (or at least high probability) readiness, it is not practical to strive for the high-end extreme of autonomy. Instead, modest advancement beyond fully automatic control to allow extended fault tolerance for anticipated events or degraded conditions and some predefined reconfigurability is the most realistic goal for an initial application of SNPS autonomous control. A hierarchical functional architecture providing integrated control, diagnostic, and decision capabilities that are distributed throughout the hierarchy can support this approach. The application of the MPC approach to the SP-100 reactor system and demonstration of key fault-tolerant and reconfigurable control features have been accomplished through simulation. The results illustrate the feasibility of incorporating these techniques in future space reactor designs.

Control systems with varying levels of autonomy have been employed in robotic, transportation, spacecraft, and manufacturing applications. However, autonomous control has not been implemented for an operating terrestrial nuclear power plant. Therefore, technology development and demonstration activities are needed to provide the desired technical readiness for implementation of an SNPS autonomous control system. In particular, the capabilities to monitor, trend, detect, diagnose, decide, and self-adjust must be established to enable control system autonomy. Finally, development and demonstration of a suitable architectural framework is also needed.

6. Acknowledgments

Portions of the work reported in this chapter were performed under the sponsorship of NASA's Project Prometheus and directed by DOE/National Nuclear Security Administration (NNSA) Office of Naval Reactors. Other reported work was sponsored by DOE Office of Nuclear Energy. Opinions and conclusions drawn by the authors are not necessarily endorsed by the sponsoring organizations.

7. References

Alami, R., et al. (1998). An Architecture for Autonomy, *International Journal of Robotics Research*, Vol. 17, No. 4, (April 1998), pp. 315–337

American Nuclear Society (1993). *Proceedings of the 1993 ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, ISBN 0-89448-185-1, Oak Ridge, Tennessee, USA, April 1993

American Nuclear Society (1996). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 96)*, Vols. 1 & 2, ISBN 0-89448-610-1, State College, Pennsylvania, USA, May 1996

American Nuclear Society (2000). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2000)*, ISBN 0-89448-644-6, Washington, District of Columbia, USA, November 2000

American Nuclear Society (2004). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2004)*, ISBN 0-89448-688-8, Columbus, Ohio, USA, September 2004

American Nuclear Society (2006). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2006)*, ISBN 0-89448-051-0, Albuquerque, New Mexico, USA, November 2006

American Nuclear Society (2009). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2009)*, ISBN 978-0-89448-067-6, Knoxville, Tennessee, USA, April 2009

American Nuclear Society (2010). *Proceedings of the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC&HMIT 2010)*, ISBN 978-0-89448-084-3, Las Vegas, Nevada, USA, November 2010

Antsaklis, P. & Passino, K. (1992). An Introduction to Intelligent Autonomous Control Systems with High Degrees of Autonomy, In: *An Introduction to Intelligent and Autonomous Control*, P. Antsaklis & K. Passino (Eds.), pp. 1–26, Kluwer Academic Publishers, ISBN 0-7923-9267-1, Boston, USA

Astrom, K. J. (1989). Toward Intelligent Control, *IEEE Control System Magazine*, (April 1989), pp. 60–64

Basher, H. & Neal, J. (2003). *Autonomous Control of Nuclear Power Plants*, ORNL/TM-2003/252, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

Chaudhuri, T. R., et al. (1996). From Conventional to Autonomous Intelligent Methods, *IEEE Control System Magazine*, (October 1996), pp. 78–84

El-Genk, M. S. & Seo, J. T. (1987). SP-100 System Modeling: SNPSAM Update, *Transactions of the 4th Symposium on Space Nuclear Power Systems*, Albuquerque, New Mexico, USA, January 1987, pp. 513-516

Gat, E. (1998). Three-Layer Architectures, In: *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*, D. Kortenkamp et al. (Eds.), pp. 195–210, MIT Press, Cambridge, Massachusetts, USA

Mishkin, A. G., et al. (1998). Experiences with Operation and Autonomy of the Mars Pathfinder Microrover, *Proceedings of the 1998 IEEE Aerospace Conference*, ISBN 0-7803-4311-5, Aspen, Colorado, USA, March 1998, pp. 337–351

Na, M. G. & Upadhyaya, B. R. (2007). Development of a Reconfigurable Control for an SP-100 Space Reactor, *Nuclear Engineering and Technology*, Vol. 39, No. 1, (February 2007), pp. 63-74

Rayman, M. D., et al. (1999). Results from the Deep Space 1 Technology Validation Mission, *Proceedings of the 50th International Astronautical Congress*, American Institute of Aeronautics and Astronautics, *Acta Astronautica,* Vol. 47, pp. 475–488

Passino, K. (1995). Intelligent Control for Autonomous Systems, *IEEE Spectrum*, (June 1995), pp. 55–62

Upadhyaya, B., et al. (2007). *Autonomous Control of Space Reactor Systems*, DE-FG07-04ID14589/UTNE-06, University of Tennessee, Knoxville, Tennessee, USA

Volpe, R., et al. (2001). The CLARAty Architecture for Robotic Autonomy, *Proceedings of the 2001 IEEE Aerospace Conference*, Vol. 1, ISBN 0-7803-6599-2, Big Sky, Montana, USA, March 2001, pp. 121–131

Winks, R. W., et al. (1992). B&W PWR Advanced Control System Algorithm Development, *Proceedings: Advanced Digital Computers, Controls, and Automation Technologies for Power Plants*, EPRI TR-100804, Electric Power Research Institute, Palo Alto, California, USA

Wood, R. T., et al. (2004). Autonomous Control for Generation IV Nuclear Plants, *Proceedings of the 14th Pacific Basin Nuclear Conference*, ISBN 0-89448-679-9, Honolulu, Hawaii, USA, March 2004, pp. 517–522

Zeigler, B. & Chi, S. (1992). Model Based Architecture Concepts for Autonomous Control Systems Design and Simulation, In: *An Introduction to Intelligent and Autonomous Control*, P. Antsaklis & K. Passino (Eds.), pp. 57–78, Kluwer Academic Publishers, ISBN 0-7923-9267-1, Boston, USA

7

Radiation-Hard and Intelligent Optical Fiber Sensors for Nuclear Power Plants

Grigory Y. Buymistriuc
Intel-Systems Instruments, Inc., St-Petersburg, Russia

1. Introduction

Optical fiber sensors (OFS) have a number of intrinsic advantages that make them attractive for nuclear power plant (NPP) applications, including absolute explosion safety, extremely low mass, small size, immunity to electromagnetic interference, high-accuracy, self-calibration, and operation in extremely harsh environments, and it is a well-known fact. Civil nuclear industry essentially encompasses the complete nuclear fuel cycle and therefore the range of possible fiber applications both for communications and sensing is very broad (Berghmans & Decreton, 1994), (Korsah et al., 2006).

In order to expand OFS applications in nuclear engineering it was necessary to overcome a bias that some scientists and engineers used to have at the initial stage of using an optical fiber for communication, about "darkening" of a fiber and sharp growth of optical attenuation under the conditions of ionizing radiation, i.e. availability of convincing proofs of radiation hardness of optical fibers and OFS.

Safety and long-term metrological stability of OFS for NPP assumes:

- Radiation hardness of fiber optic sensors and cables;
- Absence of mechanical resonances of the gauge at frequencies up to 200 Hz;
- Immunity to electromagnetic effects in the range of frequencies 200 kHz and 18 – 20 MHz,
- High reliability of a sensitive element of the OFS ;
- Temperature-insensitive measurements of pressure in the working range of temperatures;
- Self-calibration of the gauge without stopping the process of measurement.

These requirements are satisfied by modern OFS, especially intellectual optical fiber sensors which can self-calibrate, i.e. control themselves at the level of changing their internal (own) parameters depending on the calibrated value (Buymistriuc & Rogov, 2009).

No optical measurement electronics will survive in, or near, an operating nuclear reactor core. Therefore, OFS light emission must be guided to the measurement electronics located in a well-controlled, benign environment. Several different implementations can be employed to accomplish this, each with their own advantages and weaknesses. Recently single material hollow-core optical fibers (referred to as photonic crystal fibers) have become

commercially available. All silica, photonic crystal fibers appear likely to have much larger radiation tolerance than conventional optical fiber technologies.
Monitoring signals from sensors in NPP is not only to diagnose process anomalies but also it is necessary to verify the performance of the sensors and the associated instrumentations. Tests such as calibration verification, response time measurement, cable integrity checking, and noise diagnostics are required in NPP. In-situ test methods that use externally applied active test signals are also used to measure equipment performance or for providing diagnostics and anomaly detection capabilities. Controls and instrumentation were enhanced through incorporation of optical and digital technologies with automated, self-diagnostic features.
The design of the sensitive element of interferometric pressure OFS working with the measured environment of a nuclear reactor without application of pulse tubes is such that its resonant frequency lies in the range of frequencies above 60 kHz, i.e. inadmissible resonances in nuclear reactors at frequencies below 200 Hz are structurally excluded.
Was developed also methods of realization of intelligent OFS on other principles of operation, in particular possibilities of intelligentization of the acoustic emission OFS based on intrinsic optical fiber effect of Doppler, of the strain and temperature OFS based on the fiber Bragg gratings.
Coatings of the sensitive element of interferometric OFS with enhanced adhesion to silica tips and long-term durability was obtained by a molecular layering method or atomic layer deposition. An important advantage of such interferometric pressure OFS is its enhanced reliability determined by a unitary structure of the sensor and extremely high adhesion of molecular coatings to silica optical fibers. Reliability of OFS with such nano-coatings is preserved high under different external effects, including at dose ionizing radiation up 10 MGy.
Safe disposal of spent nuclear fuel (SNF) and high level waste is currently considered a major challenge, a key element to the sustainability of future nuclear power use in most countries. A first priority is obviously ensuring safety during operation under normal and faulty conditions. With this, besides contributing to guarantee operational safety, systems reliably monitoring the repository environment over several decades of years, whenever possible maintenance free and in unattended mode, can become a key element in achieving confidence on repository performance as well as public and regulatory acceptance. Application of fiber optic technologies for monitoring SNF offers distinct advantages compared with conventional systems. Optical fibers not only withstand chemical corrosion and high temperatures much better than conventional systems, but their immunity to electromagnetic interference and their large bandwidths and data rates ensure high reliability and superior performance.
Due to this optical fibers are the preferred alternative for both: sensing and signal transmission in long-term monitoring of NPP and SNF applications.

2. Background

A NPP generally uses about 200 to 800 pressure and differential pressure sensors to measure the process pressure, level, and flow in its primary and secondary systems. For example, fig. 1 shows a typical pressure sensing (pulse) line inside a nuclear reactor containment (Lin K. & Holbert K., 2010).

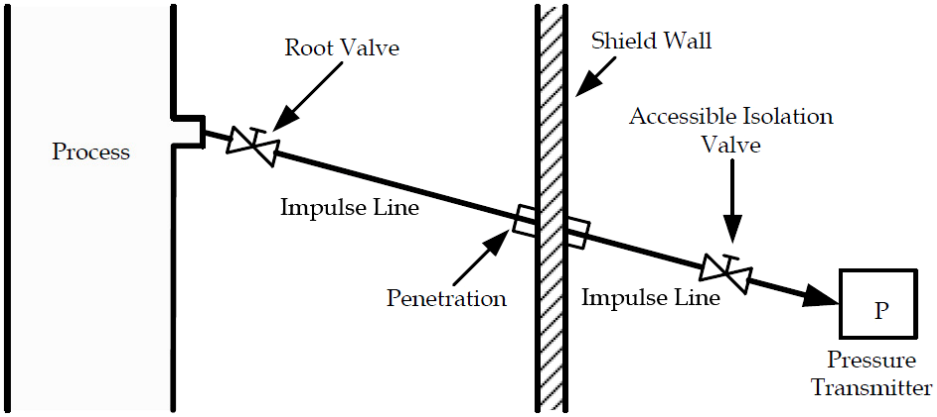


Fig. 1. Typical pressure sensing (instrument) line inside a nuclear reactor containment.

Instrument lines can encounter a number of problems that can influence the accuracy, response time of a pressure sensing system and decrease safety of NPP in consequence of mechanical resonances which appear on frequencies up 200 Hz , for example, fig. 2 shows transfer functions of a pressure sensing system (Lin K. & Holbert K., 2010).

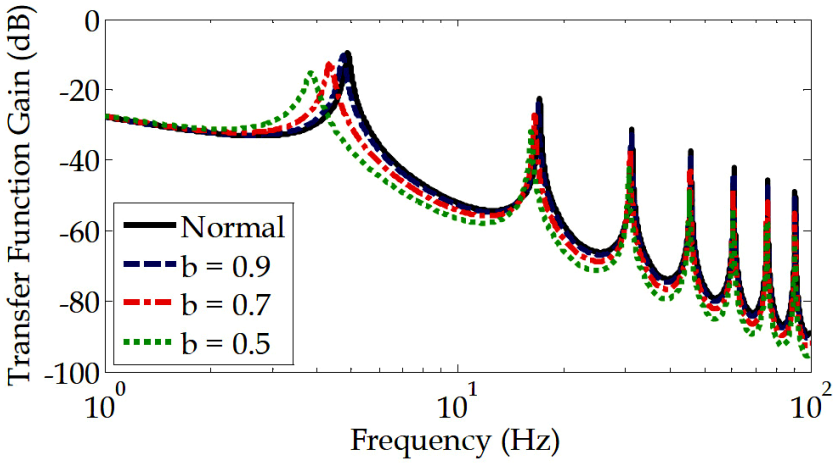


Fig. 2. Transfer functions of a pressure sensing system with resonance frequencies

Exception pulse lines from join of pressure sensors to technological equipment and pipelines in NPP is provided by Technical Regulations of the Russia (TR, 2000). Performance of this requirement became real possible only at use of fiber optic technologies. Advanced concept of construction of water-water nuclear reactors from Russian nuclear research center "Kurchatov Institute" provides use of welded joints of gages with equipment NPP instead of less reliable fitting connections that is possible with

application OFS with the big life time (up to 60 years) and with function of metrological self-calibration (Buymistriuc & Rogov, 2009).
It is important to notice that begun using fiber-optical technologies of communication and measurements in NPPs considerably improves their equipment. Really, typical NPPs used hard wired point-to-point connections from field instrumentation to control systems and panels in the control room. Essentially there is one wire per function or about 30 – 50 thousands wires coming from the field to the cable spreading room and then control room. The use of optical fiber networks, which carry substantially more information and decrease in 9 once weight of connections, instead of copper cabling, can eliminate 400 kilometers of cabling and 12500 cubic meters of cable trays (GE, 2006).
Contemporary optical fiber sensors give a unique possibility to realize the principle of remote measurement (fig. 1) .

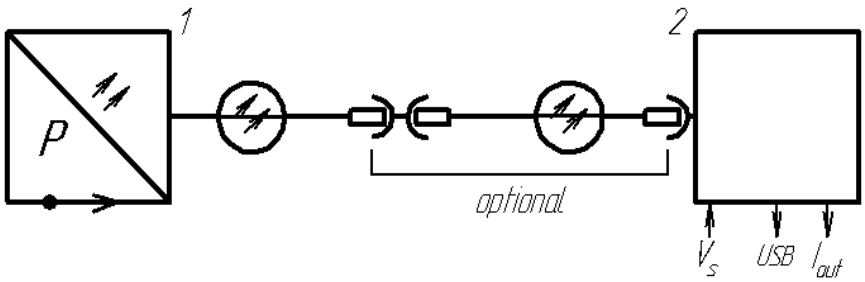


Fig. 3. Concept of remote pressure measurements
1 – OFS; 2 – optoelectronic transceiver

When sensitive element 1 of an OFS placed in harsh environment can be moved away from optoelectronic transceiver 2, which is under comfortable conditions of an equipment room, at the distance up to 3000 meters by means of an optical cable option which replaces undesirable pulse tubes very effectively.
The design of the sensitive element of pressure OFS working with the measured environment of a nuclear reactor without application of pulse tubes is such that its resonant frequency lies in the range of frequencies above 80 kHz, i.e. inadmissible resonances in nuclear reactors at frequencies below 200 Hz are structurally excluded. In fact, the resonant frequency of longitudinal vibration of optical fiber Fabry-Perot interferometer (FFPI) in the form of a quartz glass core is defined as

$$f_1 = \frac{4.91}{L}\sqrt{\frac{E}{\rho}} \tag{1}$$

where E - Young' modulus of elasticity of a glass core, Pa
ρ – glass core density, kg/m^3
E/ρ - own rigidity, in particular for silica glass, 45x10^5 m.
Thus the sensor mechanical resonant frequency is defined by its length L = 0.001 … 0.1 m and lies in the range f_1 = 10,4092 / L [kHz] = 104,092 … 10409,2 kHz.
Frequency resonant characteristic of a typical pressure OFS based on FFPI indicates Fig. 4.

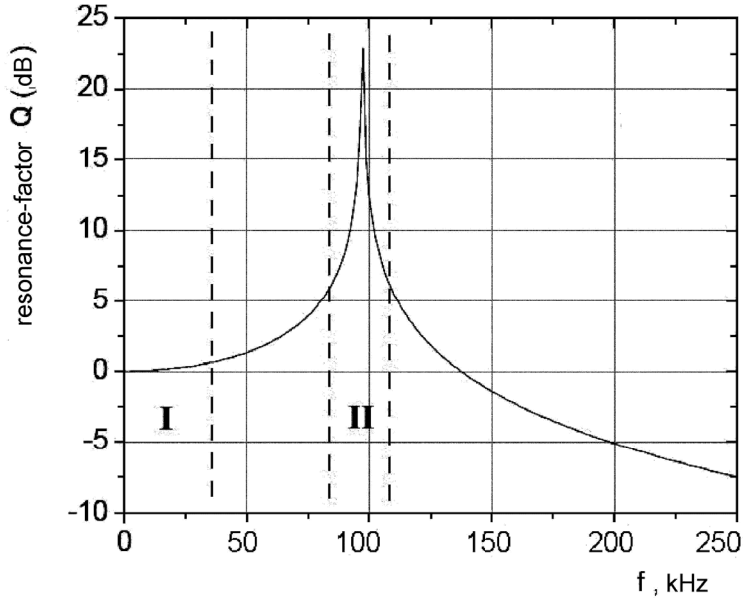


Fig. 4. A resonance frequency response of FFPI-based pressure OFS

OFS of acoustic emission, humidity and others parameters on the basis of coils of a fiber or nano-coatings of an tip of a fiber have resonant frequencies a few tens in MHz.
Use the optical fiber technologies allowing to realize a principle remote measurements changes a principle of construction of measuring systems of NPP and completely to solve a problem of resonances of pulse lines.

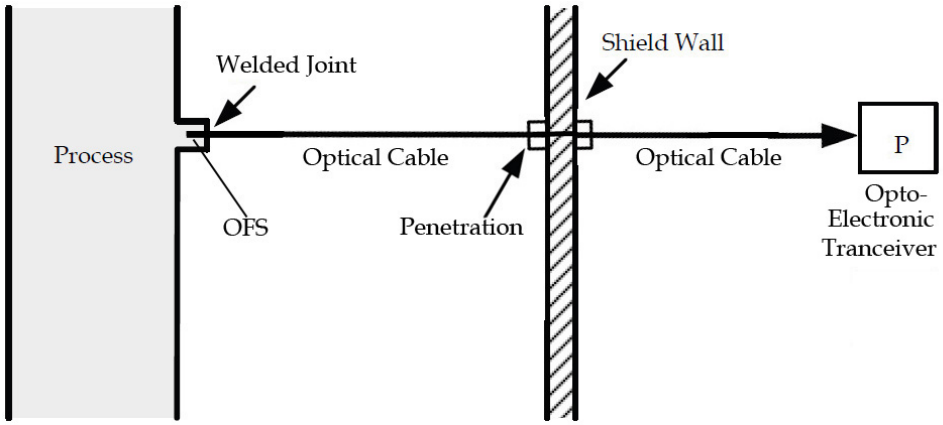


Fig. 5. New advanced structure of the pressure sensing line inside a nuclear reactor containment

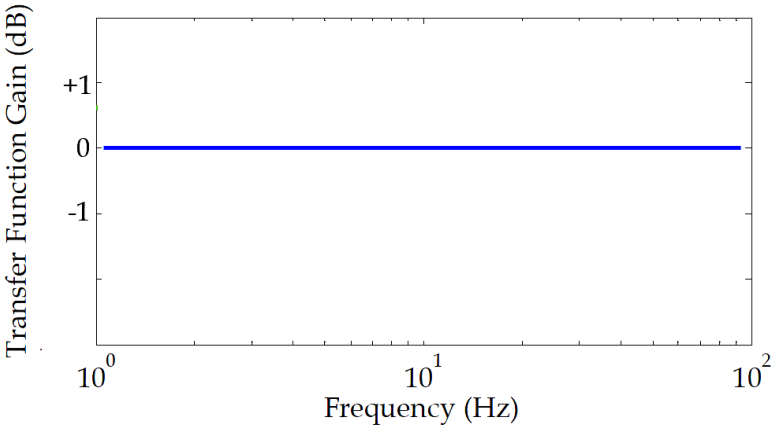


Fig. 6. Transfer gain-frequency function of a pressure optical fiber sensing system

On fig. 5 the new structure of system of measurement of the pressure is shown, one of which results of realization is full elimination resonances of the measuring channel as is shown in fig. 6.

For maintenance of working reliability of offered measuring system, especially in zones NPP with radiation presence, application radiation-hard optical fiber and cables is necessary (TR, 2003).

3. Ionizing radiation hardness of OFS

Use of silica as a material for fiber optic sensors and measuring communication lines is an effective solution both in terms of mechanical properties and radiation hardness of silica fibers which are reached by modern manufacturers, for example, a method of entering and retention of hydrogen in an optical fiber.

Another important factor of applying silica optical fibers under radiation conditions, in particular for OFS based on silica optical fibers, is their low radiation induced losses in the range of wavelengths between1150 nm and 1350 nm as is shown in fig. 7 (Fiedler et. al., 2005).

Radiation hardness of OFS equaled earlier to a general dose of irradiation of about 1.2 MGy with γ-radiation and $2.6 \bullet 10^{16}$ neutrons/cm^2 with a neutron fluence (Berghmans F. & Decréton M., Ed.,1994) but now reaches doses of gamma radiation up to 23 MGy and neutron flux $52 \bullet 10^{16}$ neutrons/cm^2 (Fiedler et. al., 2005).

Photonic crystal fibers (PCFs) were also recently submitted to a number of nuclear environments applications. In hollow core PCFs the light is essentially guided in air, which may significantly decrease the radiation response of such waveguides compared to conventional optical fibers. The structure used by us hole core PCF at 1000X and 10000X imagnifying in a microscope is shown on fig. 8.

The permanent radiation induced attenuation (RIA) levels after radiation of PCF were found to be very low. This was confirmed with hollow core PCF showing at least about 30 to 100 times lower RIA than the best present conventional optical fibers at 1550 nm with theoretical

limit of total dose of gamma radiation over 1 GGy (Henschel H. et al. ,2005). Post-fabrication treatment of the photonic band gap fiber with hydrogen gas has been reported to improve the fiber's resistance to radiation (Tomashuk A.; Kosolapov A. & Semjonov S. (2006).

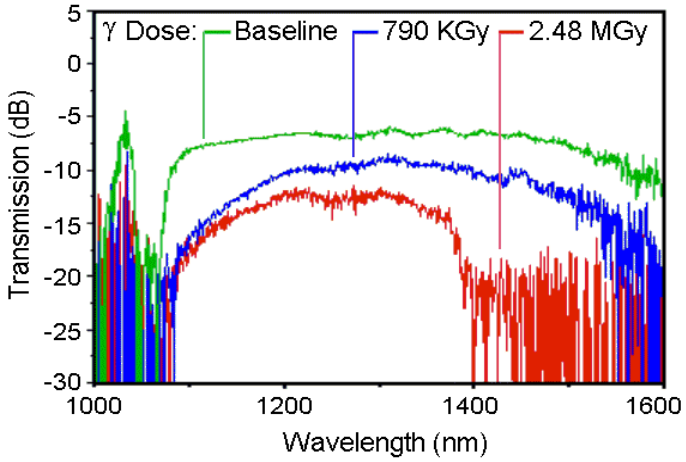


Fig. 7. Spectral transmission for 20% Ge doped silica optical fiber

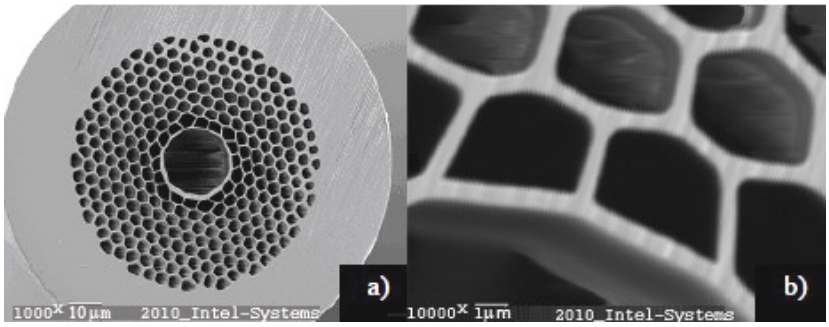


Fig. 8. Structure of the hollow core PCF at 1000^{X} and 10000^{X} magnifying in a microscope.

For extreme dose situations, the light must initially be guided from the measurement location using a hollow-core light guide. All known materials darken unacceptably in the intense radiation field of a nuclear reactor core. However, reflective technologies are available that have been shown to withstand comparable environments.

Conceptually, a hollow-core light guide is simply a mirror that has been formed into a polished titanium tube as shown on fig. 9. This approach was used for measurements temperature and the neutron flux in near-reactor environments (Holcomb D.; Miller D. & Talnagi J. (2005).

The principal innovation of this approach is to combine optical and fiber optical measurement components in a form suitable for deployment in a nuclear reactor core.
The main needs for in-pile concern the assessment of creep and growth of cladding materials, or nuclear fuel rod behaviour, which require elongation measurements or diameter measurements of cylindrical samples. Among others adaptations to the nuclear environment, these OFS will need a radiation resistant fixing.

Fig. 9. Hollow-Core Light Guide Concept.

As the light reaches lower radiation environments, several different optical transmission technologies become possible:
hollow core PCF or conventional radiation-hard optical fibers.
For total doses up to about 10^4 Gy, pure silica core, fluorosilica clad, multimode optical fibers are suitable light guides.

4. Enhanced reliability of OFS in harsh environments

A standard version of FFPI has no face reflecting coverings and works based on natural Fresnel reflections in the amount about 4%. For changing the sensitivity and dynamic range of the pressure OFS , we used TiO_2 reflecting coverings.
The fibers are placed in magnetron sputtering (MS) machine and coated with TiO_2 by vacuum deposition. The reason for using TiO_2 is that it has high refractive index (~2.4, vs. 1.4 for the silica fiber) over visible and infrared spectral ranges and strong bonding on glass based materials. The MS machine is filled with a mixture of 70% argon and 30% oxygen so that the titanium and oxygen atoms ejected toward cleaved fiber end and stick to the fiber until the desired film thickness is reached. However, experiments have shown that adhesion to glass and roughness of the TiO_2 coatings made by this method are not satisfactory. Apparently on fig. 10, mean-quadratic deviation of the surface profile equal about 37.8 nanometers.
A coating with enhanced adhesion and long-term durability was obtained by a enhanced method af atomic leyer deposition - method molecular layering (ML) (Buymistriuc & Rogov, 2009).
Synthesis was carried out by repeated and alternate processing of the surface of the fiber end face by H_2O and $TiCl_4$ steams removing the surplus of not reacted and formed by-products after each stage of processing. Thus, not more than one monomolecular layer with

the thickness of new structural units about 0.3 nanometers are added to the surface in each cycle of ML reactions.

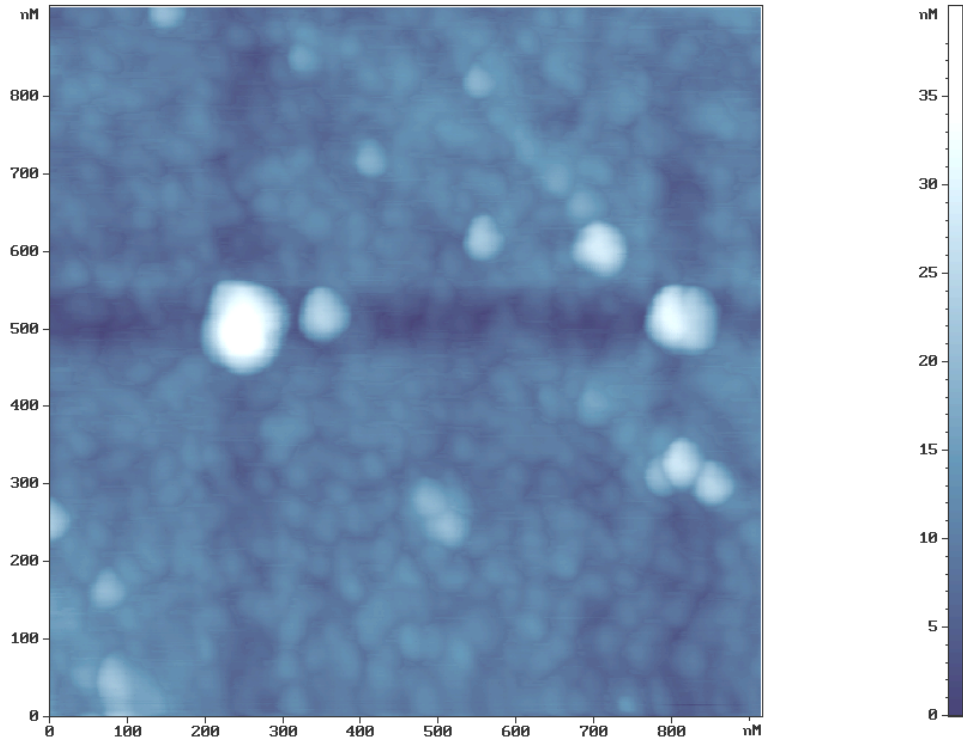


Fig. 10. Microscopic view of optical fiber tip with TiO_2 magnetron sputtering

With processing by $TiCl_4$ steams the reaction on the surface proceeds as follows:

$$(\equiv Si\text{-}OH) + TiCl_4 \rightarrow (\equiv Si\text{-}O\text{-})_2TiCl_2 + HCl \quad (2)$$

With processing by water steams the reaction on the surface proceeds as follows:

$$(\equiv Si\text{-}O\text{-})_2TiCl_2 + 2H_2O \rightarrow (\equiv Si\text{-}O\text{-})_2Ti(OH)_2 + 2HCl \quad (3)$$

At this stage of the ML process we obtain a hydroxylated surface again but now OH-groups are linked not with the atoms of silicon of the initial matrix but with the atoms which are part of the imparted functional groups. The hydroxylated surface is processed

by $TiCl_4$ steams again. At this stage the second titanoxidechloride monolayer is formed as follows:

$$2(\equiv Si\text{-}O\text{-})_2Ti(OH)_2+TiCl_4 \rightarrow [(\equiv Si\text{-}O\text{-})_2Ti(O\text{-})]_2TiCl_2+2HCl \quad (4)$$

Then a reaction product is again subject to processing by water steams. The process was finished when obtaining coatings with thickness from 10 to 180 nanometers with a mean-quadratic deviation of the surface profile about 1.4 nanometers, as shown in fig. 9 which gives the view of the end faces of the optical fiber obtained by means of an atomic-power microscope.

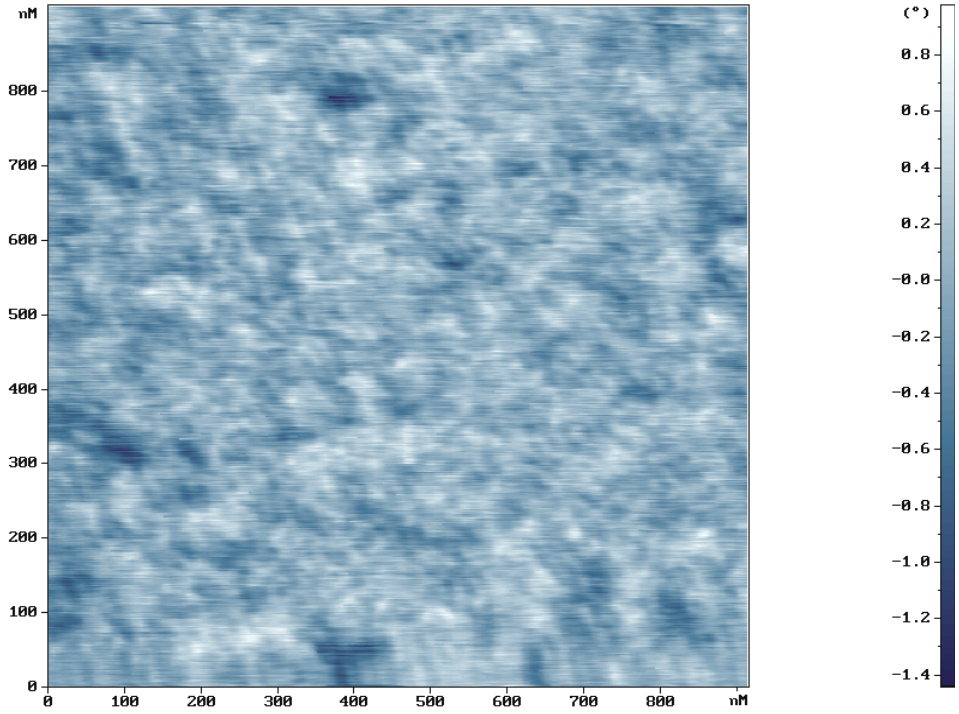


Fig. 11. Microscopic view of fiber tip with TiO_2 molecular layered nano-coating

An important advantage of such pressure OFS is its enhanced reliability determined by a unitary structure of the sensor and extremely high adhesion of molecular coatings to silica optical fibers. OFS fabricated with new ML technology possess the greatest reliability (on distribution of Weibull) than usual MS method as shown on fig. 12.

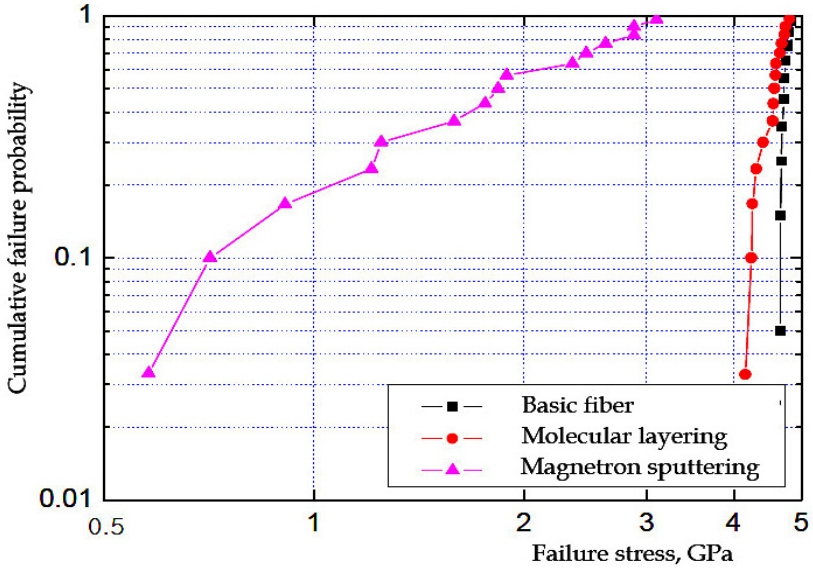


Fig. 12. Weibull distribution plots for basic fiber and fibers coated by using ML and MS technologies

5. Principles and constructions of intelligent OFS

Information redundancy of optical fiber sensors, as well as possibility to their programmable tuning in combination with a minimum structural redundancy allow to develop the so-called intelligent sensors with a function of metrological self-checking. The function of metrological self-checking of optical fiber sensors is provided with their multimodality, i.e. with their similar dependence of an output signal on several variable parameters, f.e., with their dependence on a variable pressure at a constant optical spectrum of an input signal and, accordingly, on a readjusted optical spectrum of the signal at a constant pressure. Construction of intelligent sensors of new generation assumes presence at such sensors of structural (internal) and/or information (external) redundancy (Taymanov R. & Sapozhnikova K., 2008).

Microelectronic sensors of physical quantities generate the unimodal output signal Y depending on change only of one parameter of sensor R:

$$Y = Y_0 + k \cdot \Delta R(x;t) \tag{5}$$

where Y – an output time signal, Y_0 – initial value of an output signal;
ΔR – change of parameter of a sensor, caused in the measured physical quantity x in time t;
k – proportionality factor.

That is, microelectronic sensors do not possess necessary information redundancy. To provide self-checking of such sensors by creation of information redundancy, for example giving on them influences of physical quantity of known value – it is almost impossible while in process controllable equipment in real time. Therefore intelligent microelectronic

sensors are under construction by creation of structural redundancy (embedding of the reference sensor, the additional sensor with parameters close to the basic sensor, etc.) that not always is the optimum decision.

5.1 Self-checking OFS

Application of a fiber optic Fabry-Perot interferometer for measurements of pressure and speed of pressure variation in water reactors of NPPs contributes to improving their safety and long-term metrological stability, which demands for intelligent sensors.

5.1.1 Basic principles

By means of fast tuning of the spectrum of an optical source it is possible to make self-calibration in the course of continuous work of the pressure gauge. Optical cables including connectors, splices, and other components are tested by evaluating the optical losses relationship along the cable.

OFS of physical quantity creates the multimodal output signal depending at least from two parameters of the sensor, for example for OFS based on FFPI output signal I_s depends on change of length of optical resonator G and change of a wavelength of light λ:

$$I_s = I_0\left[1 - Cos\left(\frac{4\pi}{\Delta\lambda_0(t)} \cdot \Delta G(x;t)\right)\right] \tag{6}$$

where I_0 – initial intensity of light coupled into FFPI.

The output signal such OFS according to the equation (2) changes depending on change of length of a cavity of the resonator ΔG, caused by pressure, and depending on change of the central optical wavelength of the coherent sensing channel $\Delta\lambda_0$, provided, for example, by the tuneable spectral optical filter (TSOF) as is shown in fig. 13.

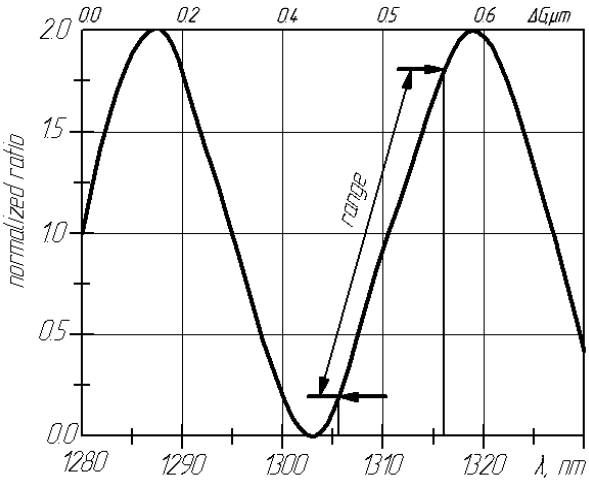


Fig. 13. Response of FFPI output from resonator cavity length and optical source peak wavelength changes

Additional possibility at such intelligent OFS is possibility of stabilization of a quiescent point (Q-point) on its linear site calibration characteristics by fine tuning of a wavelength of light on value $\delta\lambda_0$ sensing channel, compensating a deviation of initial length of a cavity of the optical resonator δG, caused by destabilizing factors during long operation of the sensor. Thus, it is obvious that for realization FMSC interferometric OFS possess necessary information redundancy.Structural redundancy of OFS at realization FMSC is minimal and is reduced to application of the TSOF, as shown on the scheme fig. 14.
On such principle has been realised the intelligent pressure OFS with function metrological self-checking (FMSC) at long operation in extreme conditions (Buymistriuc G. & Rogov A.,2009).
Speed of tuning of a modern TSOF, for example the models "FFP-TF" from "Micron Optics, Inc" or the acousto-optical tuneable filter models "AOTF" from "Fianium, Inc" is rather high also the period of tuning time T on all set spectrum. For example in sequence $T = t_3 - t_1$ on fig. 15, makes value of an order 0,1 … 0,4 microsecond that it is enough for the majority of modes of measurement of pressure, deformation, vibration, temperature, level of etc. controllable industrial equipment.

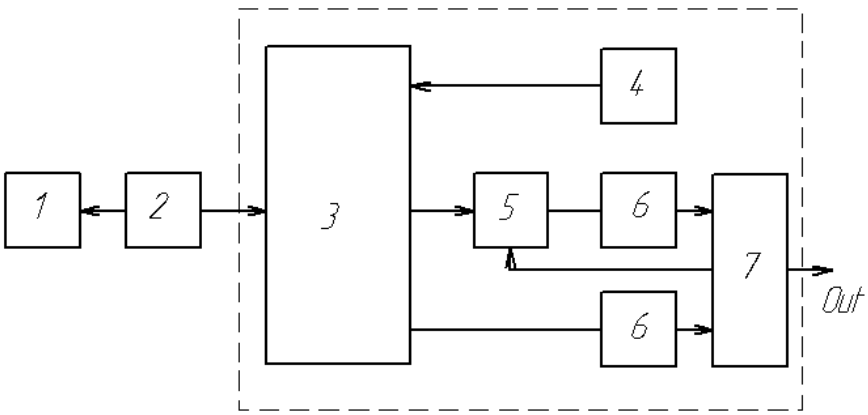


Fig. 14. Intelligent pressure OFS
1 – FFPI; 2 – optical cable; 3 – optical coupler; 4 –light emitting diode; 5 –TSOF;
6 –photodiode; 7 – microcontroller

5.2 Algorithm

Procedure of self-checking of pressure OFS on the basis of FFPI consists of the following consecutive steps.
Step 1. When OFS is manufactured, its calibration characteristic is measured:

$$I_c = f(\Delta G\{P\}) \text{ at } \lambda_0 = const$$

by means of a precision pressure calibrator (for example from DPI-610 from "Druck, Ltd") and stored as initial data in the energy-independent memory of the device.

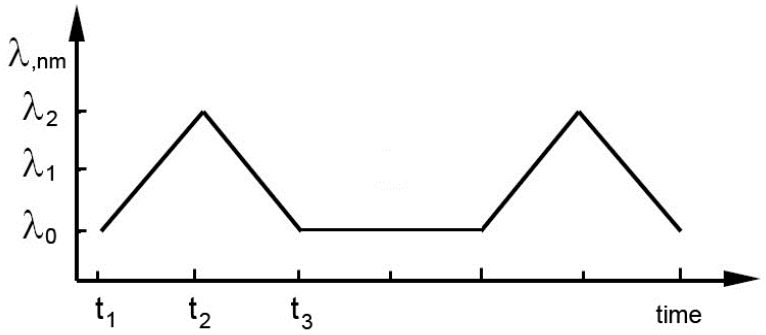


Fig. 15. A spectral peak in time change of tunable optical filter of the OFS

Step 2. While OFS is in service after a certain period of time, which may be equal to an periodic testing interval of the device, the current calibration characteristic of the gauge is measured without stopping the process of pressure measurement :

$$I'_c = f(\lambda) \text{ at } \Delta G\{P\} = const$$

by fast tuning of the central wavelength of the spectral optical filter transmission [$\lambda_0 \rightarrow \lambda_2$] – forward scan, and [$\lambda_2 \rightarrow \lambda_0$] – reverse scan, as shown in fig. 15. The speed of the filter tuning should exceed the rate of the measured pressure change.

Step 3. The graduation characteristic of the pressure OFS is compared with the initial data of calibration and corresponding correction factors are calculated.

Step 4. The measurement of pressure by means of OFS with using correction factors obtained at step 3 is continued.

Thus, interference OFS is internally inherent necessary information and structural in redundancy effectively to carry out FMSC.

5.3 Estimation of error

Error of measurement of the intelligent OFS, caused by instability of the central wavelength of TSOF, i.e. shifted spectra TSOF and light emitting diode and casual displacement of spectral characteristics of the filter us has been investigated by modeling and experimental check.

Distortion s' of an output signal s, is described by following expression:

$$s' = s\left[1 - \frac{\Delta I}{I_o}\right]\left(1 + \frac{\gamma_c}{2} Cos\left[\frac{4\pi}{\lambda - \Delta\lambda} G\right]\right) \tag{7}$$

where γ_c – contrast ("fringe visibility") of interference.

The normalised error of measurement of pressure $\delta s/s$, caused by shift $\Delta\lambda$ the central wavelength of the filter, is described by expression:

$$\frac{\delta s}{s}\Bigg|_{\max} \approx \left\{ abs\left(\frac{1}{I_0}\frac{dI}{d\lambda}\right) + \frac{\gamma_r}{2} \cdot \frac{4\pi G}{\lambda^2} \right\}\Bigg|_{\lambda - \Delta\lambda} \cdot \Delta\lambda \tag{8}$$

A case non-centering (casual displacement) of a wavelengths of the optical filter and a source of optical radiation in the course of long operation.
By means modeling and experiences is established that spectrum mismatching of the TSOF with a light source poorly influences an error of measurements of OFS. Big shifted spectral mismatch, above 3 nm, gives metrological error less then 0.02%.
Besides by tuning TSOF with feedback included on an output signal it is possible to provide stability of a quiescent point (Q-point) interferometric OFS that is to realise the intelligent gauge with self-correction.
Possible application of adaptive OFS is their use as additional measuring transducers of the self-checking channel at construction of intelligent gauges of pressure with elastic elements and gauges of level and other physical quantities for the purpose of substantial growth of their actual interval between testing.

5.4 Intelligent acoustic emission OFS

Was investigated also methods of realization of intelligent OFS on other principles of operation, in particular possibilities of intellectualizing of the acoustic emission OFS based on intrinsic optical fiber effect of Doppler (Li F. et al., 2009).
A Doppler shift of frequency f_D under influence of a sound wave a is defined as

$$f_D = -\frac{n}{\lambda_0(t)} \cdot \frac{dL(a;t)}{dt} \tag{9}$$

where n –index of refraction of a fiber; λ_0 – wavelength of laser radiation; L – length of fiber sensing element; a – acoustic signal; t – time.
Apparently from a equation (9) and fig. 4 that frequency OFS has information redundancy – dependence Doppler shift of frequency both from change of length of fiber L, and from change of frequency of laser radiation f_0, and structural redundancy – the tunable laser diode.

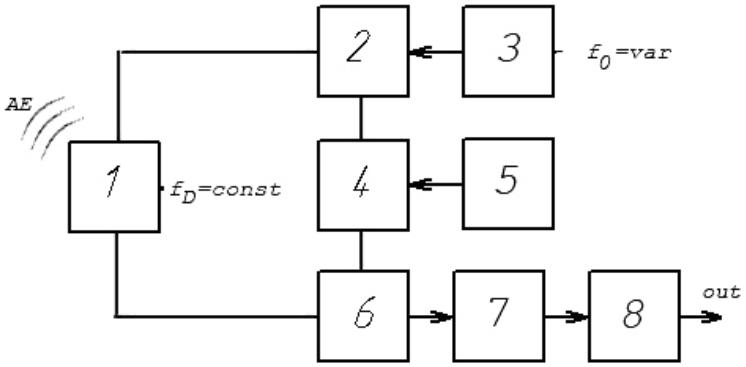


Fig. 16. Intelligent acoustic emission OFS
1-fiber coil sensing element; 2, 6 – fiber optic couplers; 3 – tuneable laser diode;
4 – acousto-optical frequency shifter; 5 – RF generator; 7 – frequency detector;
8- microcontroller

When absent an etalon device for creation, in real-time, on the OFS of acoustic emission of entrance mechanical influence with known characteristics, i.e. when $f_D = const$, self-checking of the amplitude-frequency characteristic of the OFS is made by fast tuning of acousto-optical modulator shift frequency f_M , or by fast tuning of optical frequency of laser diode radiation f_0. Modern tunable acoustic-optical modulators and laser diodes have period of tuning through all spectrum of the frequency characteristic about 1 μs that it is quite enough at frequencies acoustic emission up to 500 kHz.
This graduation amplitude-frequency characteristic of the acoustic emission OFS is compared with the initial data of calibration of OFS and corresponding correction factors are calculated. The measurement of acoustic emission by means of OFS with using obtained correction factors is continued.

5.5 Intelligent fiber Bragg grating OFS

The intelligent strain OFS design uses an induced phase discontinuity (structural redundancy) in the fiber Bragg grating (FBG) that gives information redundancy in optical spectrum which is used for self-monitoring sensor's condition .
The reflectivity at the Bragg wavelength can be estimated using the equation

$$R(\Lambda) = \tanh^2\left[\pi \cdot n\left(\frac{L}{\lambda_B(\Lambda)} \cdot \frac{\Delta n}{n}\right)\left(1 - \frac{1}{V^2}\right)\right] \tag{10}$$

where L – the length of FBG; n – index of refraction;
V - the normalized frequency of the fiber ($V \geq 2.4$);
λ_B – Bragg diffraction wavelength that has multimodal response:

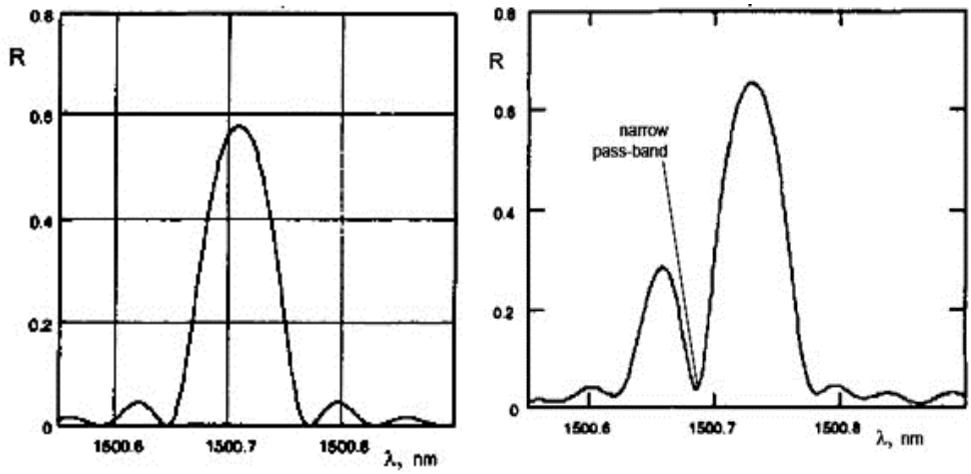
$$\lambda_B(\Lambda;t) = 2 \cdot n \cdot \Lambda(x;t) \ , \ \Lambda = const \ \ \forall x \in L, \ \ \Lambda_0 > \Lambda \ \text{ for } x_0 \in L \tag{11}$$

where Λ – the period of the n modulation of of FBG; x – the fiber axis direction; x_0 – the point of phase discontinuity.
Figures 17a and 17b show the reflection spectrum of FBG-based diffraction OFS with and without a phase discontinuity respectively. The presence of a narrow pass-band in the wavelength spectrum is due forming of phase discontinuity in the FBG structure.
While a strain is applied along axis of the healthy FBG sensor with a phase discontinuity (eq. 7), optical reflection spectrum (eq. 6) shows linear shift as

$$\Delta\lambda_{BS} = \lambda_B\left(1 - \rho_\alpha\right)\Delta\varepsilon \tag{12}$$

where ρ_α - the photoelastic coefficient of the fiber , ε – strain (tension).
When a strain is applied along axis of the damaged FBG sensor with a phase discontinuity (eq. 7), then damage changes the photoelastic coefficient of fiber (strain sensitivity varies) and optical spectrum shows linear shift (eq. 8) and simultaneously changes her shape - the relative position of the main peak of the FBG and narrow pass-band will no longer constant: their disposition are changed that represents a damaged sensor.
It is important to note that the presence of this narrow pass-band in spectra does not add any additional complexity to the signal processing schemes to be used.

Simple tracking the phase-shift and main peak of optical spectrum allows determine technical condition (health) of the FBG strain sensor.



a) b)

Fig. 17. Reflection spectra of strain FBG optical fiber sensor
a)– without discontinuity b) – with discontinuity

6. Applications OFS in nuclear power plants

The analysis of possible ways of influence of electromagnetic disturbances on pressure sensors under the conditions of factual operation on NPP units has shown that the basic influences leading to failures in work of electronic pressure gauges occur on the direct current feed circuit, the current signal output circuit and as a result of impact of electromagnetic fields.

The Russian Federal Agency of Ecological, Technological and Nuclear Supervision yearly registered a few deviations in the work of Russian nuclear stations. The causes included malfunction and false data of the electron pressure sensor.

In addition, after the inspection of electromagnetic conditions on the NPP it was established that the basic components of periodic disturbances at the output of pressure gauges have characteristic frequencies in the range of 18 – 20 MHz and 200 kHz imposed on a supply mains voltage sinusoid, which increases the value of the gauge current signal to a setting value at which emergency protection is activated.

The reason of EMI occurrence at frequencies 18 – 20 MHz and 200 kHz is the electronic circuit of the gauge in the long line operating mode when high-frequency fluctuations fade after several runs from one end of a coaxial cable to another.

This circumstance demands replacement of electronic pressure gauges by noise-proof OFS with fiber optic cable lines.

A general view of intelligent pressure OFS based on FFPI is shown in fig. 18.

a) b)

Fig. 18. General view of intelligent pressure OFS
a)– sensing element (b)- optoelectronic transceiver

The pressure applied to a sensitive element causes blooming deformation of the FFPI glass cover. The cover radius decreases resulting in the growth of length of the L_0 resonator cavity by value ΔG up to hundred nanometers.

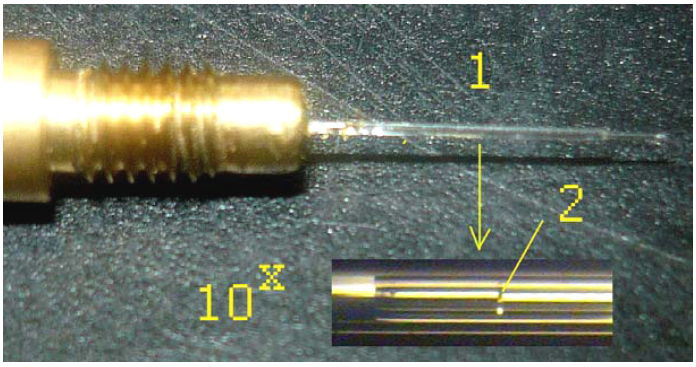


Fig. 19. General view of FFPI-based pressure sensor
1 – sensing element; 2 – FFPI cavity with ΔG length (10^x magnifying)

Analytical expression of the change of the length of the optical resonator ΔG follows from mechanics principles owing to the applied pressure:

$$\Delta G = \frac{L_o}{E}\left[\sigma_z - \mu\left(\sigma_r - \sigma_t\right)\right] \tag{13}$$

where σ_z, σ_r, σ_t - longitudinal, radial and lateral deformation of the sensor.

$$\Delta G = \left[(1 - 2\mu) \cdot k_r \cdot \frac{L_0}{E} \right] \cdot P \tag{14}$$

where P – pressure on the sensitive element;
μ- Poisson factor for quartz glass, 0.17;
E –Young modulus for silica, 73 GPa.
k_r – geometry factor of the sensitive element, 1.16.
FFPI output sensing signal is defined by ΔG value with the constant transmission peak of the spectral optical filter λo:

$$I_s = \left[1 - Cos\left(\frac{4\pi}{\lambda_0} \cdot \Delta G \right) \right] \tag{15}$$

Normalization of the FFPI output signal, i.e. relation formation

$$I_N = \frac{\int_{-\infty}^{+\infty} I_s(\lambda) \cdot \left[1 - Cos\left(\frac{4\pi}{\lambda} \Delta G \right) \right] \cdot d\lambda}{\int_{-\infty}^{+\infty} I_0(\lambda) \cdot \left[1 - Cos\left(\frac{4\pi}{\lambda} \Delta G \right) \right] \cdot d\lambda} \tag{16}$$

allows practically to eliminate the influence of instability of the optical source radiation (for example, ageing or change of the current rating), the influence of the induced losses of optical power in the sensor (for example, ionizing radiation) and the influence of insertion losses in the optical communication cable (for example, abrupt bends of optical cable and connectors). It is experimentally established that, for example, at an average value of relation I_N = 0.27145 and insertion optical losses up to 70% the standard deviation of the relation from the average value does not exceed 0.00063 (0,23 %) as shown in fig. 20.

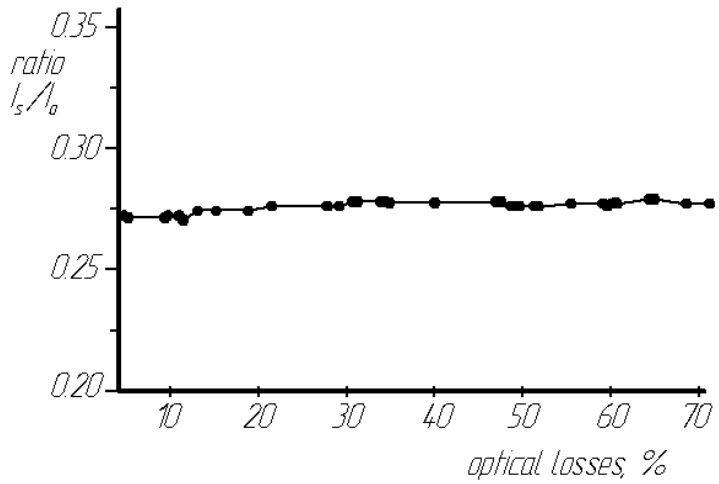


Fig. 20. Normalized output signal of the pressure OFS

The temperature dependence of the resonator cavity length is expressed by the formula:

$$\Delta G = \left(\alpha_o L_o - \alpha_1 L_1 - \alpha_2 L_2\right)\Delta T \,, \tag{17}$$

where ΔT – temperature change;
α_o , α_1, α_2– CTE of capillary cover quartz glasses, input and output fibers, accordingly;
L_o , L_1 , L_2 – lengths of the sensor measuring cavity, input and output optical fibers of the interferometer accordingly.
To reduce a temperature error to a minimum, it is necessary to pick up such values of α_o , α_1, α_2, L_1 that the expression (17) of the formula in brackets almost equals zero:

$$\alpha_o L_o - \alpha_1 L_1 - \alpha_2 L_2 = 0 \tag{18}$$

at $\alpha_o = 6.1 \cdot 10^{-7}$ K^{-1}, $\alpha_1 = 6.5 \cdot 10^{-7}$ K^{-1}, $\alpha_2 = 5.6 \cdot 10^{-7}$ K^{-1}, $L_1 = 5.06$ mm we obtain:

$$\alpha_o L_o - \alpha_1 L_1 - \alpha_2 L_2 = 0{,}0002 \text{ нм} \,.$$

When the temperature ΔT of the measured environment changes by 600°C, the length of the FFPI resonator cavity will change according to the formula (17) by 0.1 nanometers, which is below a permissible threshold. That is, FFPI provides for practical temperature insensitive measurement of pressure in the entire working range of temperatures.
Fig. 7 shows calibration characteristics of pressure OFS with temperature self-compensation for working temperatures 24 °C and 250 °C. Practically, additional temperature errors are absent.

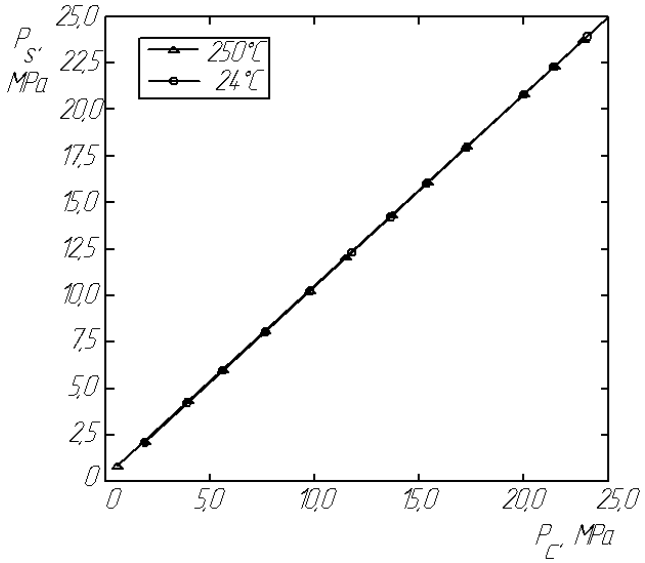


Fig. 21. Calibration characteristics of the temperature insensitive OFS

7. Application of OFS for spent nuclear fuel repository monitoring

Radioactive waste requires, due to its high hazard potential, careful handling from production to final disposal. In all waste management activities the safety objectives defined

by the authorities must be met to guarantee an adequate protection of the public. Waste management starts with the registration of the radioactive waste arising at different locations from different applications in industry, research as well as at nuclear fuel cycle facilities. The waste is then stored, conditioned into an appropriate form for further handling and disposal, intermediately stored whenever necessary over long periods of time, and eventually disposed of (Jobmann M. & Biurrun E.,2003).

Long-term effectiveness, low maintenance, reliable functioning with high accuracy, and resistance to various mechanical and geochemical impacts are major attributes of monitoring systems devised to be operated at least during the operational phase of a repository. In addition, low maintenance and automatic data acquisition without disturbing the normal operation will help reducing operational costs. Due to these reasons Russian "Krasnoyarsk SNF repository" started using of reliable and radiation-hard fiber optic technology as the basis for global monitoring systems at final disposal sites.

Series of parameters important to safety of SNF repository can be monitored by optical sensors. Sensing elements to measure strain, displacement, temperature, and water occurrence together with the multiplexing and data acquisition systems were installed at 1000m depth and the operation temperature is about 40 °C .

The configuration of experimental OFS system is shown on fig. 22.

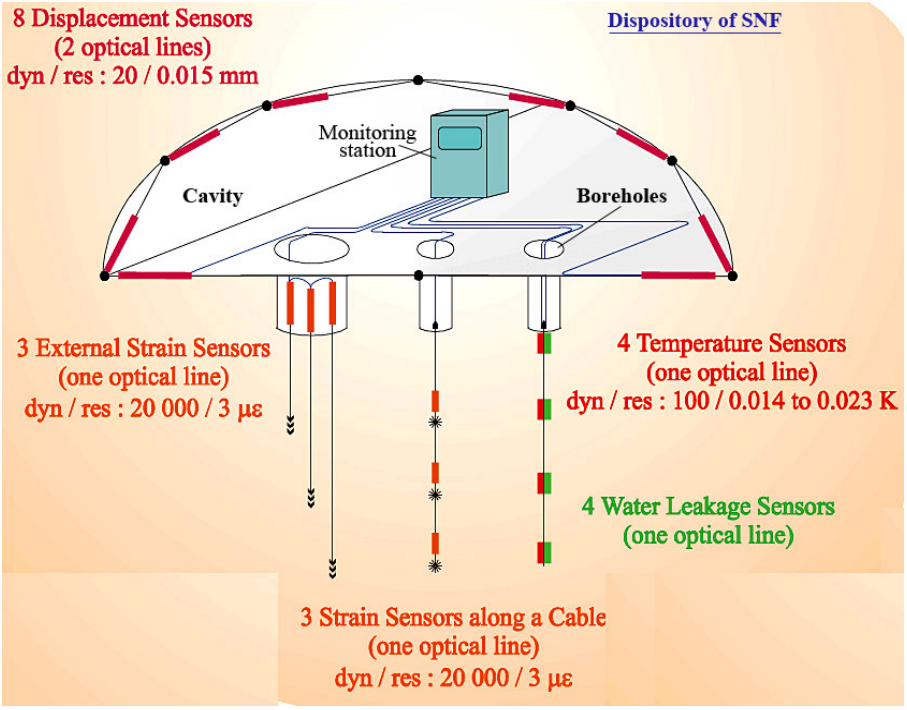


Fig. 22. Configuration of the OFS system in SNF repository

In three boreholes strain, temperature, and water detection sensors are installed, whereas the displacement sensors are fastened around the cross-section of the drift to monitor changes of the cavity geometry.

The complete circuit diagram of the OFS system is shown on fig. 23.

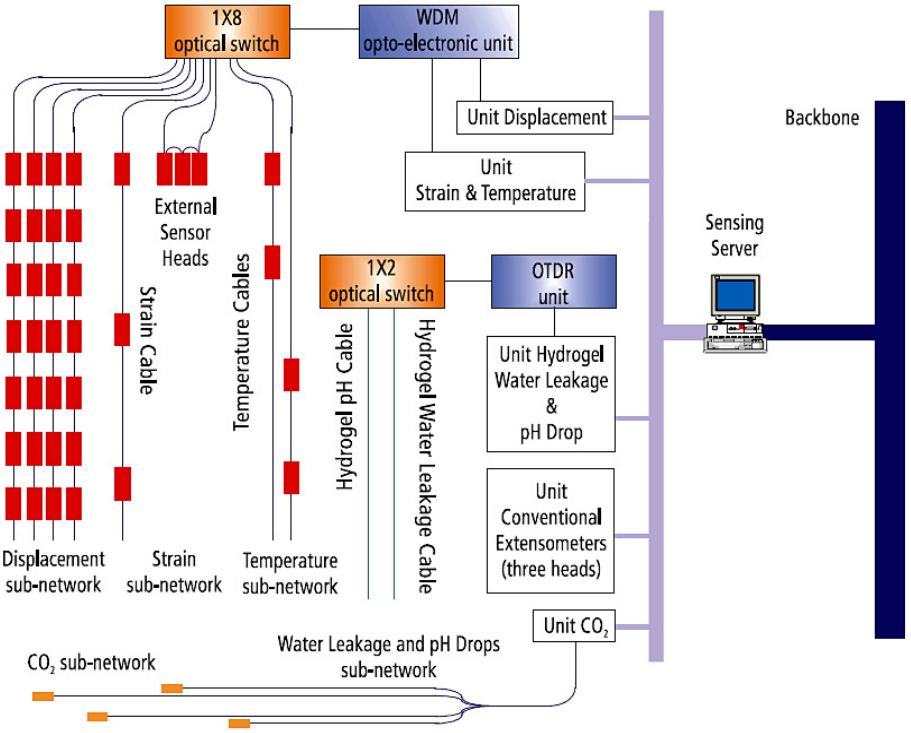


Fig. 23. Circuit diagram of the OFS system in SNF repository

All measured data will be collected by a so called sensing server via the corresponding multiplexing units. The sensing server can be connected to a backbone providing the data in special output files to be downloaded by the user.

The presented fibre optic sensing systems which can be used in an all fibre optic network could be the basis for a high-reliability, low-maintenance, economic monitoring system for operational safety requirements in a final repository as shown in fig. 24.

Monitoring the cavities deformation at representative cross-sections will be the basis for evaluating the operational safety. Together with temperature monitoring as a function of time at different locations, data for validating the thermo-mechanical constitutive laws of the host rock will be available.

Monitoring of harmful gases as methane and carbon dioxide is an important issue in a salt environment because of the ongoing excavations during the operational period. Thus, fibre optic gas sensors at least for measuring methane and carbon dioxide was included in the global monitoring systems for spent nuclear fuel repositories. In an underground repository, the availability of appropriate monitoring tools is a major issue in order to ensure operational safety and to verify that the repository evolves as predicted. The feasibility of measuring safety relevant parameters using sensors and multiplexing systems based on fibre optic technology.

Further developments are necessary to increase accuracy in large sensing networks and to check the long term performance.

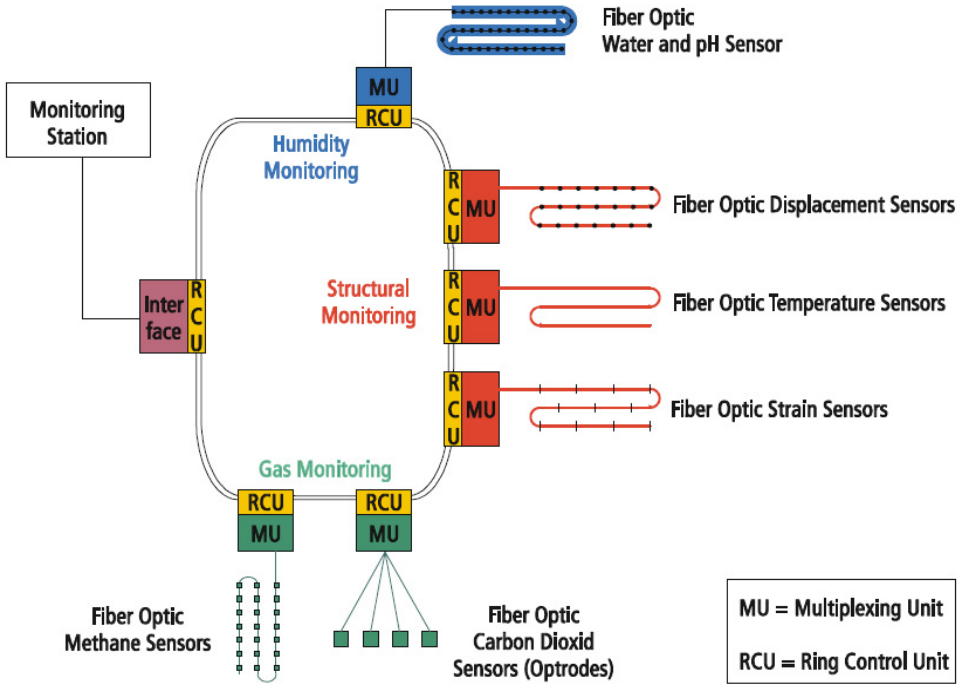


Fig. 24. All-fiber optic sensors network for SNF repository

8. Trends in developments OFS for nuclear energy an industry

In the next decade, nuclear energy is expected to play an important role in the energetic mix. Various national and international programs taking place in order to improve the performance and the safety of existing and future NPPs as well as to assess and develop new reactor concepts. Instrumentation is a key issue to take the best benefit of costly and hard to implement experiments, under high level of radiation.

OFS are contributed to improve instrumentation available thanks to its intrinsic capability of high accuracy associated with the passive remote sensing implementation allowed by fiber optic communication line. It can work under high temperature and high radiation. The small size is appreciated attending the lack of available space in research reactor, while miniaturized sensors will not disturb the temperature and radiation profile on the tested material. The ability of fiber optic sensors to provide smart sensing capabilities, detailed self-diagnostics, and multiple measurements per transducer and distributed OFS for temperature, strain and other parameters profiling are provided. These capabilities, coupled other intrinsic advantages, make fiber optic sensors a promising solution for extremely harsh-environment applications where data integrity is paramount.

The advanced fiber optic sensing technologies that could be used for the in fusion reactors, for example ITER, safety monitoring. The remote monitoring of environmental parameters,

such as temperature, pressure and strain, distributed chemical sensing, could significantly enhance the ITER productivity and provide early warning for hazardous situations.
The development of new intelligent (smart) OFS involves the design of reconfigurable systems capable of working with different input sensors. Reconfigurable systems based on OFS ideally should spend the least possible amount of time in their calibration (Rivera J., et al., 2007).
A traditional NPPs control system has almost no knowledge memory. The neural network, by comparison, learns from experience what settings work best. The system updates the network weighting factors with a learning algorithm. The neural network outputs adjusts the basic parameters (criteria) of technological processes and safety of NPP. This is an excellent artificial intelligence application. Rather than model and solve the entire process, this neural network handles a localized control challenge.
When a complex system NPP is operating safely, the outputs of thousands of sensors or control room instruments form a pattern (or unique set) of readings that represent a safe state of the NPP. When a disturbance occurs, the sensor outputs or instrument readings form a different pattern that represents a different state of the plant. This latter state may be safe or unsafe, depending upon the nature of the disturbance. The fact that the pattern of sensor outputs or instrument readings is different for different conditions is sufficient to provide a basis for identifying the state of the plant at any given time. To implement a diagnostic tool based on this principle, that is useful in the operation of complex systems, requires a real-time, efficient method of pattern recognition. Neural networks offer such a method. Neural networks have demonstrated high performance even when presented with noisy, sparse and incomplete data. Neural networks have the ability to recognize patterns, even when the information comprising these patterns is noisy or incomplete. Unlike most computer programs, neural network implementations in hardware are very fault tolerant; i.e. neural network systems can operate even when some individual nodes in the network are damaged. The reduction in system performance is about proportional to the amount of the network that is damaged.
Beyond traditional methods, the neural network based approach has some valuable characteristics, such as the adaptive learning ability, distributed associability, as well as nonlinear mapping ability. Also, unlike conventional approaches, it does not require the complete and accurate knowledge on the system model. Therefore it is usually more flexible when implemented in practice. Thus, systems of artificial neural networks have high promise for use in environments in which robust, fault-tolerant pattern recognition is necessary in a real-time mode, and in which the incoming data may be distorted or noisy. This makes artificial neural networks ideally suited as a candidate for fault monitoring and diagnosis, control, and risk evaluation in complex systems, such as nuclear power plants (Uhrig R. 1989).
The objective of this task is to develop and apply one or more neural network paradigms for automated sensor validation during both steady-state and transient operations. The use of neural networks for signal estimation has several advantages. It is not necessary to define a functional form relating a set of process variables. The functional form as defined by a neural network system is implicitly nonlinear. Once the network is properly trained, the future prediction can be interpolated in real-time. The state estimation is less sensitive to measurement noise compared to direct model-based techniques. As new information about the system becomes available, the network connection weights can be updated without relearning the entire data set. These and other features of neural networks will be exploited in developing an intelligent system for on-line sensor qualification.

We believe that researchers and instrumentation designers of new generation of NPPs will use novel approaches to conduct real-time multidimensional mapping of key parameters via optical sensor networks, distributed and heterogeneous sensors designed for harsh environments of nuclear power plants and spent nuclear fuel respository. Recent events on Japanese NPP "Fukushima-1" are characteristic that within two weeks the information from gages, as NPP was without power supplies, was inaccessible and electronic gages couldn't transmit the important measuring information for condition monitoring of NPP. Contemporary OFS, as it is known, are radiation-hard and don't need power supplies, and the optoelectronic transceiver can be installated on distance to 80 km from NPP that will allow to supervise NPP during any critical periods and to accept the right decisions on elimination of failures.

9. References

Berghmans F. & Decréton M., Ed. (1994). Optical fibre sensing and systems in nuclear environments, - *Proc. of the SPIE,* vol. 2425. -160 p..

Buymistriuc G., Rogov A. (2009). "Intelligent fiber optic pressure sensor for measurements in extreme conditions". – *1st Int. Conf. "Advan. in Nuclear Instrum., Meas. Methods and Appl."- ANIMMA, Marseille, France, 6-9 June 2009.*

Fiedler R.; Duncan R. & Palmer M.(2005). Recent advancements in harsh environment fiber optic sensors as enabling technology for emerging nuclear power applications. - *Proc. of the IAEA Meeting, Knoxville, Tennessy, 27-28 June 2005.*

GE (2006) Economic Simplified Boiling Water Reactor Plant General Description, *General Electric Company*, p. 12-3.

Henschel H.; Kuhnhenn J. & Weinand U. (2005). High radiation hardness of a hollow core photonic crystal fiber, *Proc. 8th European Conf. RADECS, Cap d'Agde, France, 19-23 September 2005.*

Holcomb D.; Miller D. & Talnagi J. (2005) Hollow core light guide and scintillator based near core temperature and flux probe.- *Proc. of the IAEA Technical Meeting on "Impact of Modern Technol. on Instrum. and Control in NPP, Chatou,France, 13-15 september 2005.*

IEC (2003). TR 62283 Nuclear radiation. Fiber optic guidance.

Jobmann M. & Biurrun E.(2003). Research on fiber optic sensing systems and their applications as spent nuclear fuel final repository tools. –*Symp. on Waste Management, Tucson, Arizona, 23-27 February 2003.*

Korsah K. et al., Ed. (2006). Emerging technologies in instrumentation and controls. Advanced fiber optic sensors". -*Report of the US Nuclear Regulatory Commission, NUREG/CR-6888, ch. 3, pp. 47–52.*

Li F. et al. (2009) Doppler effect-based fiber optic sensor and its application in ultrasonic detection for structure monitoring. - *Optical Fiber Technology, vol. 15.*

Lin K. & Holbert K. (2010) Pressure sensing line diagnostics in nuclear power plants. – "*Nuclear Power*", P. V. Tsvetkov, Ed., Sciyo, Rijeka, Croatia, Chapter 7, pp. 97-122

Liu H.; Miller D. & Talnagi J. (2003) Performance evaluation of optical fiber sensors in nuclear power plant measurements. -*Nuclear Technology, vol. 143; No 2.*

Nannini M.; Farahi F. & Angelichio J. (2000) An intelligent fiber sensor for smart structures. – *J. of Struct. Control,* 1 (7).

Rivera J., et al. (2007) Self-calibration and optimal response in intelligent sensors design based on artificial neural networks.-*Sensors, 7.* ISSN 1424-8220 .

Taymanov R., & Sapozhnikova K. (2008) Automatic metrological diagnostics of sensors, *"Diagnostyka"*, 3(47).

Tomashuk A.; Kosolapov A. & Semjonov S. (2006). Improvement of radiation resistance of multimode silica-core holey fibers", *Proc. of the SPIE* vol. 6193.

TR (2000). OTT 08424262 Instruments and automatic systems for NPP. Common technical requirements (*in Russian*).

Uhrig R. (1989). Use of neural networks in nuclear power plants.- *Proc. of the 7rh Power Plant Dynamics, Control & Testing Symposium, Knoxville, Tennessee, May 15-17, 1989.*

8

Monitoring Radioactivity in the Environment Under Routine and Emergency Conditions

De Cort Marc
European Commission, JRC, Institute for Transuranium Elements, Ispra
Italy

1. Introduction

The main purpose of environmental monitoring is to quantify the levels of radioactivity in the various compartments of the environment disregarding its origin: natural or anthropogenic, under routine or accidental conditions, in view of assessing the health effects on man and his environment. However, because of their historical background, which is connected to the development of nuclear industry, the monitoring programmes established in the European countries focus on artificial radioactivity.

Man-made radioactive matter can get into the biosphere by means of legally permitted discharges from nuclear installations or infrastructures where radioactive material is being used, e.g. hospitals and industry, or as the result of an accident. For each cause, specific sampling and monitoring programmes, as well as systems for internationally exchanging their results, have been implemented in the European Union and are still evolving.

Routine monitoring is done on a continuous basis throughout the country by sampling the main environmental compartments which lead to man; typically these are airborne particulates, surface water, drinking water and food (typically milk and the main constituents of the national diet). The aim of routine monitoring is then also to confirm that levels are within the maximum permitted levels for the whole population (Basic Safety Standards, (EC, 1996)) and to detect eventual trends in concentrations over time. A comprehensive overview of the sampling strategies and principal measurement methods in the countries of the EU will be given, as well as how this information is communicated to the general public.

In case of an accident, monitoring (i.e., sampling, measuring and reporting) is tailored to the nature of the radioactive matter released and to the way in which it is dispersed. In particular during the early phase of an accident with atmospheric release it is essential to be able to delineate the contamination as soon as possible to allow for immediate and appropriate countermeasures. Afterwards, once the radioactivity has deposited, it is important to have detailed information of the deposition pattern; a detailed deposition map at a fairly early stage will serve to steer medium and long term countermeasure strategies (e.g. agricultural, remediation). A summary of the most commonly used techniques, as well as a discussion of the various sampling network types (emergency preparedness, mobile) will be given.

The Chernobyl NPP accident on 26 April 1986 also triggered the European Commission to develop, together with the EU Member States, systems for the rapid exchange of information in case of a nuclear/radiological accident (European Community Urgent Radiological Information Exchange (ECURIE), European Radiological Data Exchange Platform (EURDEP)). Also these systems will be further described.

2. Types of monitoring networks

Depending on the risk, networks have been developed for various purposes. In the first place there is the monitoring of radioactivity releases at nuclear installations, which aims at verifying the authorised discharges. In addition, in most European countries an environmental monitoring programme is operated for the main compartments of the biosphere, i.e., air, water, soil, foodstuffs. The purpose of such an environmental radioactivity monitoring programme is to verify compliance with the basic safety standards for the public.

However, this objective is influenced by the source of radioactivity as well as the environmental compartment(s) affected. Radioactive material mainly comes into the environment by means of discharges into the atmosphere and/or the water. These discharges can happen in a controlled or in an accidental way. Therefore a distinction should be made between routine and emergency situations. In general one can distinguish between the following types of monitoring networks:

- surveillance monitoring networks around nuclear installations to ensure that releases to the atmosphere and acquatic compartment remain below authorized limits, and to verify potential, chronic accumulation of radioactivity in the environment. Results obtained in this way may be used to estimate radiation exposure to critical groups (i.e., members of the public who have been identified as likely to receive the highest doses) (Hurst & Thomas, 2004). In case of accidental release, these networks can also provide information about the off-site contamination close to the installation, usually by means of ambient dose rate or by air concentration measurements;
- national surveillance networks that generally cover the whole territory. These contribute to ensuring compliance with the basic safety standards for the population at large. These networks are operated on a national basis and cover the whole biosphere;
- emergency preparedness networks continuously check levels of mainly ambient dose equivalent rate and airborne radioactivity, in order to detect accidental releases and subsequently monitor the evolution of the radioactive plume. Depending on the country, the monitors belonging to these networks are positioned along national borders and/or distributed over the national territory;
- mobile equipment: depending on the size and the type of accident (release into the air and/or the aquatic phase, types of radionuclides dispersed etc.), additional mobile equipment (terrestrial or airborne) will be needed to obtain more detailed information for more highly contaminated areas.

Whereas the first two types of network mainly have been designed for routine conditions, the latter two types have been designed in view of accidents. Most of the information during the early (or release phase) of an accident will come from the emergency preparedness network and to a lesser extent from the national surveillance networks.

Ideally, an emergency monitoring strategy combines routine monitoring procedures with special requirements set by the emergency e.g. by combining measurement results from fixed monitoring stations (static network) with those from mobile or intervention teams.

During the aftermath of an accident, emergency monitoring is not only important for effective post accident management but also to reassure the general public. Therefore, during a nuclear emergency the measuring and laboratory activities, as well as the general preparedness to perform situation analysis, are enhanced and intensified and special measuring systems (in particular mobile monitoring equipment) are used when appropriate (Lahtinen, 2004).

3. Environmental sampling and measuring techniques

3.1 Exposure pathways

Atmospheric discharges may result in exposure from four pathways, leading to doses to the population:

- external contamination;
- inhalation;
- ingestion;
- external radiation.

To estimate the consequences of the external contamination and inhalation pathways, monitoring by air sampling is performed. For the ingestion pathway this happens by of means of food sampling, e.g., milk, whereas external radiation is determined by direct measurements of external dose or by soil analysis.

Liquid discharges may irradiate man through three pathways:

- ingestion;
- external contamination;
- external radiation.

The monitoring of dose from ingestion in this case is usually carried out through sampling of fish and shellfish. The other two pathways are monitored by sampling of water, aquatic bio-indicators (e.g., seaweed, fish, molluscs) and sediments, and by direct measurements of doses from handling fishing gear or residing on beaches (Aarkrog, 1996).

Internal contamination, as a result of inhalation and/or ingestion, can also be measured directly by whole body counting equipment (see also section 4.4.3.1). In specific radiological situations, like in the case of contamination with pure alfa and beta emitting radioisotopes, monitoring of the internal contamination can be performed by analysis of the blood, urine and/or faeces.

3.2 Air

3.2.1 Introduction

The purpose of monitoring airborne radioactivity in the environment is to check domestic and foreign facilities. Depending on the meteorological conditions, airborne radioactive material can be rapidly transported over long distances in any direction. Man can become contaminated immediately through inhalation or external contamination, or indirectly by deposition and transfer of the radionuclides into the food chain. Therefore, monitoring the air is particularly important in order that contamination be detected as early as possible.

In general, one distinguishes two sampling and measuring techniques for air:

- particulates (alpha/beta or nuclide specific);
- gases (e.g., gaseous iodine, noble gases).

Airborne particulate radioactivity concentration is difficult to measure directly, since the artificial activity concentrations are typically lower than the natural radioactivity concentrations. Therefore accumulation methods are being used. The airborne dust is collected by drawing air through a filter material, which can be made of paper, glass fibre or polypropylene. The sampling devices may be located in diverse environments (in an open field or a courtyard, at ground level or on the roof of a building), which however complicates the intercomparability of measurements and their representativeness. Natural radionuclides include radon and its short-lived decay products (typically 1 - 20 $Bq \cdot m^{-3}$ in outdoor air), ^{7}Be and ^{40}K.

Depending on the response time of the measurement systems, one should make a distinction between on-line and off-line sampling/measuring devices.

3.2.2 On-line measurements

Based on the nuclide category to be measured, generally two measuring methods are considered:

Alpha/beta measurements:

Large area proportional counter tubes are used to measure the accumulated activity. Fixed filter devices only permit sampling periods of maximum one week and require thus considerable operational service. Automated filter changing mechanisms allow automatic operation up to six months and are particularly used in automatic monitoring networks; their drawback, however, is that they require regular maintenance. It is common practice for monitors to have flow rates of up to 25 $m^3 \cdot h^{-1}$, and detect artificial alpha and beta activity concentrations down to 0.1 $Bq \cdot m^{-3}$ in less than one hour in a natural background of several $Bq \cdot m^{-3}$. By increasing the filter speed (in case of a ribbon filter) or by increasing the frequency of the filter exchange, one is able to measure up to 106 $Bq \cdot m^{-3}$ (Frenzel, 1993).

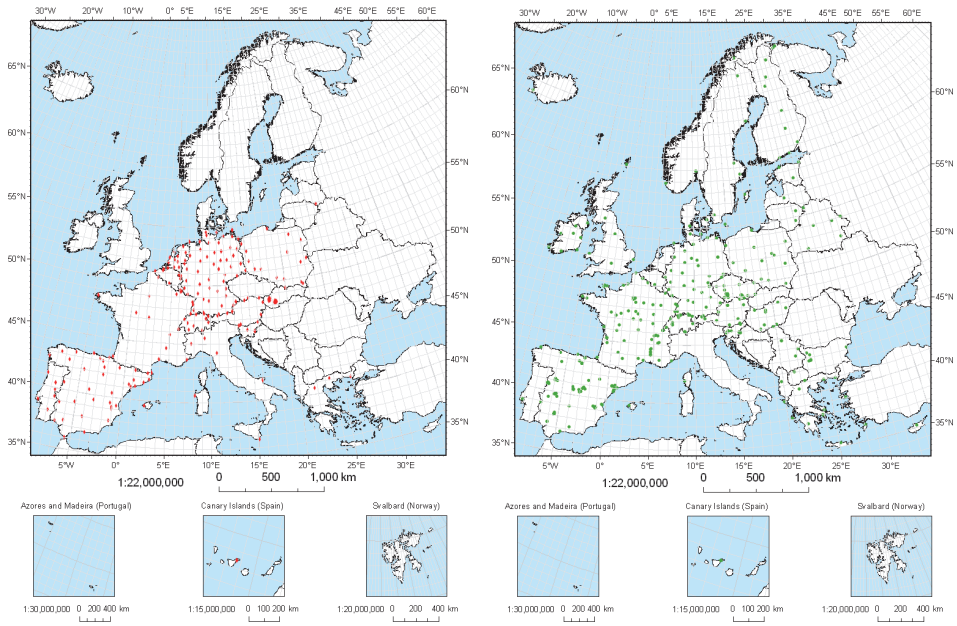


Fig. 1. On-line (left) and off-line (right) aerosol monitoring networks in a number of European countries (Bossew, P. et al, 2008)

Widely used measuring methods are:

- gross alpha: i.e., total alpha minus natural alpha radioactivity. The measurement is made by gas flow proportional counters, with or without anti-coincidence. Lower detection limits range between $1x10^{-5}$ and $4x10^{-2}$ $Bq \cdot m^{-3}$;
- gross beta: i.e., total beta radioactivity with correction for natural radioactivity (mostly influenced by radon daughters). The measuring instruments used are: Geiger-Müller, gas flow proportional counter with different active surfaces and plastic scintillators with ZnS coating for simultaneous coincidence of alpha contamination. Depending on the methods

used, the lower detection limits range between $5x10^{-5}$ and 1 $Bq \cdot m^{-3}$. Distinction between natural and artificial radioactivity can be done very effectively by simultaneous coincidence counting of alpha decay and by assuming that there is no contribution of artificial alpha emitters. In this way concentrations of artificial radioactivity down to 0.1 $Bq \cdot m^{-3}$ can be detected. One may also perform a second, delayed beta-counting after, e.g., 12 h (most short-lived daughters, except ^{212}Po, will then have decayed). This is a valid procedure for routine monitoring, but in emergency situations the alarm level will still be determined by the natural background (on the order of 10 Bq m-3) (Janssens et al, 1991).

Nuclide specific (gamma) measurements:

The filter is measured by solid state detectors:

- semi-conductor detectors (lithium drifted (GeLi) or high purity (HPGe) germanium detector); Measuring systems in inaccessible locations, such as high mountains or remote islands, become independent from the liquid nitrogen by using electrically cooled Ge detectors;
- or scintillation counters (NaI).

Nowadays nuclide-specific identifications are increasingly performed by means of high purity Ge detectors. To ensure optimum early warning, these instruments are designed to allow simultaneous alpha-beta measurement and nuclide-specific measurement. In the automatic mode, modern instruments are capable of reaching low detection limits (50 $mBq \cdot m^{-3}$ for ^{60}Co in 1 h) and can analyse spectra for up to 100 different nuclides. Mostly ^{137}Cs and ^{60}Co and some natural nuclides such as ^{7}Be are measured.

3.2.3 Off-line measurements

Artificial airborne radioactivity in the range of $\mu Bq \cdot m^{-3}$ cannot be detected by automatic instruments in 'real-time', since the natural radioactivity level is too high. Most institutes perform correction from natural radioactivity by waiting 2 to 5 days before measuring the filter, to allow for short-lived radon/thoron daughters to decay.

Typically, high-volume air samplers (HVAS) with air flow rates ranging from 100 to 1,000 $m^3 \cdot h^{-1}$ collect airborne particulates during one week. In case of an emergency, the sampling frequency can be increased to daily. By means of gamma spectroscopy detection levels of a few $\mu Bq \cdot m^{-3}$ or less can be obtained (e.g., ^{137}Cs concentrations on the order of 0.1 $\mu Bq \cdot m^{-3}$ can be measured in a $3x10^5$ m^3 filter sample). Chemical treatment of certain filters afterwards allows the analysis of pure alpha or beta emitters (e.g., ^{239}Pu and ^{90}Sr) (Frenzel, 1993).

3.2.4 Measuring gaseous components

Iodine is selectively accumulated in the thyroid gland and retains special attention because of its potential health hazard. Iodine in air can be bound to aerosols or can be gaseous, each requiring special measuring techniques. When bound to aerosols, iodine is measured with techniques as described previously.

Gaseous iodine is sampled by means of special filters (e.g., silver impregnated activated carbon). The active carbon adsorbs and thus accumulates the gaseous iodine. The drawback of these filters is that, depending on the airborne iodine concentration, they become saturated and have to be replaced. This method may also be used for collecting noble gases (e.g., ^{85}Kr). The active carbon cartridge surrounds the detector or is located directly next to it. The specific peak of 364 keV of ^{131}I is usually measured with scintillation (NaI) detectors. Although HPGe detectors are used for measuring the contribution of different iodine radioisotopes, nowadays ^{131}I concentrations in the order of hundreds of $mBq \cdot m^{-3}$ can be measured in 1 hour.

3.3 Surface water

Surface water includes river, lake or sea water. It is one of the environmental compartments to which radioactive effluents from nuclear installations can be directly discharged. Some of the sampling methods are automatic and continuous and are designed to detect contamination of water purification stations by radioactive effluents from industrial and research laboratories, hospitals having a nuclear medicine, etc... (e.g., the Telehydroray system installed on French rivers, and the Belgian automatic monitoring system on the river Meuse (Debauche, 2004). These automatic devices are designed to perform total gamma counting and to measure concentrations of ^{131}I, ^{99m}Tc and ^{137}Cs with detection limits in the order of 1 $Bq \cdot l^{-1}$.

The time and frequency of sampling is very important for rivers with large differences in seasonal hydrological variations. In all cases, additional information on the river flow rate is very important. Radionuclides can be found in the water phase and associated with suspended particles, becoming eventually incorporated into sediments and living species. Natural radionuclides in surface water include ^{3}H (0.02 - 0.1 $Bq \cdot l^{-1}$), ^{40}K (0.04 - 2 $Bq \cdot l^{-1}$) radium, radon and their short-lived decay products (< 0.4 – 2 $Bq \cdot l^{-1}$).

For routine conditions, river water samples can be taken continuously (or daily) and are then bulked into monthly or quarterly analysis. Alternatively spot samples are taken periodically and analysed individually.

Some laboratories (e.g., in France) filter their surface water prior to measurement. Measurement is then performed on the filtered water and the suspended material separately. More elaborate chemical separations are needed for ^{90}Sr, whereas ^{3}H, which is also produced by nuclear industry, is measured after multiple distillation or electrolytic enrichment of the sample. Usually, residual beta (total beta less ^{40}K activity) contamination is reported (De Cort et al., 2009), although there is a clear tendency in many countries to perform nuclide-specific measurements.

With the exception of tritium in rivers with nuclear industry, usually the levels of radionuclide contamination in surface water are below the detection limit, due to the diluting factor. Hence countries nowadays make more use of biological indicators (aquatic moss, molluscs, vegetation) as these organisms have the capacity to concentrate specific chemical (stable and radioactive) elements. Fish is also frequently sampled, being a better activity integrator in the longer term (Sombré & Lambotte, 2004).

3.4 Soil/sediments/deposits

Airborne particulates are removed from the atmosphere by gravitational settling and turbulent transfer to ground surfaces (dry deposition) or by incorporation in or scavenging by rain droplets (wet deposition). The latter is the predominant process in European countries, and monitoring networks do not generally measure the two components separately.

Depending on the circumstances and the objectives of the measurement, deposition can be sampled and measured in many different ways:

- for routine measurements, the radioactivity is mostly accumulated on artificial collection surfaces, with active surface areas ranging from 0.05 to 10 m^2. The materials used for the collecting areas vary (stainless steel, plastic, polyester, PVC). In most cases these are designed for collecting precipitation. Some detectors distinguish between wet and dry deposition (they open and close appropriately in response to rainfall), but most sample total (wet + dry) deposition. The rainwater may be directly fed to a bottle or to an ion exchange column where the nuclides are fixed. Measuring the activity in the collected rainwater and the amount of rainfall allows calculating the average total

deposition during the sampling period. Between the different European countries sampling periods range from daily, over weekly to monthly (De Cort et al., 2009);

- in case of emergency with large transboundary release to the atmosphere, vast areas can be contaminated. Once the radioactive material has deposited, it is important that, in order to determine appropriate countermeasures in the immediate post-accident management period and to reassure the public, rapid and reliable monitoring of contaminated areas is performed. Airborne gamma ray spectrometry is a fast and effective means to monitor large areas for deposited radioactivity. It is performed by solid-state detectors (usually by means of high volume scintillation detectors (NaI), high purity germanium detectors or a combination of both) mounted on an aeroplane or helicopter that flies at constant speed and low altitude (typically 50-100 m). Along the flight track, time-integrated gamma ray spectra (typically 1 s) are recorded which, along with the information obtained from on-line GPS and radar-altimetry for automatic position and height correction, allow terrestrial radioactivity levels to be quantified and mapped. One of the main sources of uncertainty arises from the uncertainty in the activity depth distribution in the soil. Most commonly nowadays calibration is done by comparison with ground based spectrometry on representative radiation fields. Beyond its obvious benefits in case of a large nuclear accident, airborne gamma ray spectrometry is also applied for geological mapping for mineral exploration, soil mapping for agriculture, pollution studies and lost sources. (Dickson, 2004)
- core sampling is a conventional technique and together with high resolution spectrometry it can be highly accurate and sensitive. It is the only method for radioisotopes which do not emit gamma rays (eg. ^{90}Sr). It serves as an indicator of long-term build-up of radioactivity in the environment and is therefore essential for studies of vertical profiling and measurements by alpha, beta or mass spectrometry. It has the drawback of being time-consuming. Migration in the soil depends on the chemical form of the radionuclide, the soil type, hydrology and agricultural practices. Most artificial radionuclides are found in the upper 30 cm soil layer. Because of the possibility of large micro-scale variations in deposition, it is important to take a sufficient amount of soil samples in order to obtain a reasonable estimate of the deposition of radioisotopes at a given site. Subsequently the samples are thoroughly mixed in order to obtain a representative aliquot which can then be further analysed and measured (Aarkrog, 1996);
- in-situ measurements, where a gamma spectrometer is placed at 1m above the ground, properly shielded by lead to measure in solid angle and thus reducing ambient gamma radiation (Raes, 1989). In-situ gamma-spectrometry measurement of the mean surface radioactivity concentration for a large area (~1 ha) in a relatively short time (generally 15-30 minutes for a deposition of 10 kBq m^{-2} of ^{137}Cs) (Dubois & Bossew, 2003). The uncertainty of the measurement is influenced by local topographic variations (buildings, trees,…), by vegetation cover and by the vertical activity distribution of the radionuclide. Good results have been obtained by means of an advanced analysis of the measured gamma spectrum, referred to as 'Peak-to-valley' method (Gering et al., 1998; Tyler, 2004). In-situ measurements are also common practice in emergency preparedness networks where gamma dose-rate detectors (usually Geiger-Müller probes, proportional counters, ionisation chambers) monitor continuously the ambient gamma dose-rate.
- herbage is also a readily available indicator of deposition and of incorporation into green vegetables. Furthermore herbage forms part of the milk pathway to man (Hurst & Thomas, 2004).

3.5 Food chain

3.5.1 Drinking water

Radioactivity in drinking water is an important indicator of radionuclide transfer from the environment to man. The most important natural radionuclides in water for drinking consumption are ^{3}H (0.02 - 0.4 Bq l^{-1}), ^{40}K (typically 0.2 Bq l^{-1} but widely variable), radium, radon and their short-lived decay products (0.4 - 4 Bq l^{-1}).

The sampling of drinking water varies in the European countries, depending on the national water resources and distribution systems. Drinking water thus may be sampled from ground or surface water supplies, from water distribution networks, mineral waters and table water in bottles. After sampling, the water is mostly evaporated for direct measurement of the residue or is separated on ion-exchange columns. More elaborate chemical separations are needed for ^{90}Sr, whereas ^{3}H is measured by liquid scintillation after purification of the water sample by multiple distillations (De Cort et al., 2009).

The European Commission has issued a directive on water quality, including radioactive aspects. In particular, a limit of 100 Bq l^{-1} of tritium and a total indicative annual dose of 0.1 mSv (natural radioisotopes not included) from water intended for human consumption has been established (EC, 1998). Member States are adapting their national monitoring programme to meet this demand.

3.5.2 Milk

Milk constitutes a principal pathway for exposure to airborne effluents. In addition it is an important foodstuff because it is produced continuously in large quantities, and it is consumed as such or it forms the basis for other foodstuffs (e.g., dairy products). It is essential for children, and the most important fission products such as ^{90}Sr, ^{131}I and ^{137}Cs are secreted in it.

At national level, milk samples are mostly taken at dairies that cover large geographical areas (in order to obtain representative samples), at farms (from raw milk) or at the super market (from bottled milk). To complete the national programme, supplementary milk samples are usually collected at single farms close to nuclear installations. Generally the samples are taken on a monthly basis, but sometimes only over the pasture season. Usually the milk samples are dried before gamma spectrometric analysis. Chemical separation is applied for the determination of ^{90}Sr activity and separation of the water phase with subsequent distillation for the separation of ^{3}H.

In addition, the concentrations of the stable isotopes Ca and K are determined because of the similarity of their metabolic behaviour with Sr and Cs, respectively. Typical values in milk are 1 - 2 g l^{-1} for Ca and 1 - 3.5 g l^{-1} for K (De Cort et al., 2009).

In case of an accidental atmospheric release, only a few days after deposition occurred, radionuclides already reach their maximum activity concentration in the milk (eg 2-4 days for ^{131}I, 4-6 days for ^{137}Cs or ^{90}Sr) (Mercer et al, 2002). Hence immediate monitoring is required. However, in the immediate aftermath of an accidental release, some information may be available on the source and scale of the release, but it is very unlikely that any measurement data on environmental materials will be available.

Predictive models, eventually in combination with data on activity deposited in soil, are essential at this point to provide an initial estimate of the dispersion of the activity released. Estimates of the evolution of activity concentrations in milk are required in the early stages following an accident. These are determined by the time of the year the deposition occurred (greater contamination of milk in summer and autumn when cattle are grazing in the pasture or delayed contamination can occur when contaminated fodder has been harvested).

Within this time, an extended monitoring programme with intensified sampling of milk at affected dairies would have to be started, and subsequently samples sent to laboratories. Results for gamma ray emitting radionuclides (eg ^{131}I or ^{137}Cs) would be available within 1 hour, whereas the determination of pure beta emitters (like ^{90}Sr) would require several days.
An efficient alternative has been developed by the Radiation Protection Division/Health Protection Agency, UK. It consists of a portable specialised NaI detector to measure individual milk samples at bulking depots located in the vicinity of the contaminated area. Information on the radionuclide composition would be required to ensure a proper calibration of the measuring equipment. A minimum detection limit of 100 Bq l^{-1} within 100 s of counting time is achievable (Mercer et al, 2002).
Information on the radionuclide composition of the deposited activity is a priority for a sampling and measurement programme. This enables the radionuclides of primary radiological interest to be identified and the analytical strategy to be determined. Gamma-emitting radionuclides can be determined rapidly without destroying the sample. However, it is also important to determine the contributions from beta emitters like ^{89}Sr and ^{90}Sr. (Mercer et al, 2002)

3.5.3 Foodstuffs

Foodstuffs are measured as separate ingredients (e.g., cereals, meat, fish, vegetables and fruit) or as whole meals (e.g., in canteens of factories or schools). The reason for measuring ingredients is to complete the monitoring programme for migration of radionuclides in the food chain or to check contamination of the public at large through ingestion, whereas sampling the whole meal gives a more direct estimate of the dose received by the population through ingestion. For general surveillance programmes the latter is more representative for the ingestion dose of the population, although ingredient monitoring is generally applied to obtain information on the propagation of radioisotopes in the food chain. Ingredients are measured particularly in case of emergencies to monitor the evolution of radioactive contamination of specific foodstuffs.
In some European countries (e.g., Denmark, United Kingdom and Finland), an important programme for monitoring fish has been established. This is needed to determine radionuclide transfer in the aquatic environment. Usually, fish types most commonly caught for human consumption are sampled. They are sorted by their origin, species and size (Saxen 1990).
Because of differences in the composition of national diets, there is a tendency in the EU to sample complete meals at schools or factory canteens in order to give a representative figure for contamination in a mixed diet (expressed in Bq·d^{-1} per person). Knowledge of the contamination of the different ingredients together with the composition of the national diet can also lead to a representative figure for the radioactivity level in a mixed diet.
The radioactivity levels legally permitted in foodstuff in the EU are laid down in the appropriate EC legislation (EC, 1989; EC, 2000a). Monitoring of foodstuff in the aftermath of a large scale nuclear accident will require major and specific efforts, depending on the type and scale of the atmospheric release. Crop monitoring programmes can be significantly rationalised by using results from airborne gamma spectrometry surveys of the contaminated area, combined with appropriate food chain models.
In the long term it is important that specific food types continue to be monitored, in particular foodstuffs coming from affected areas, even when radioactivity contamination levels in agricultural products have returned to normal. Typical examples are semi-natural foodstuffs that concentrate Cs, such as mushrooms, reindeer (through lichen), wild boar and carnivorous lake fish (EC, 2003).

3.6 Ambient dose equivalent

External radiation is measured as an instantaneous gamma dose-rate or a gamma dose integrated over a certain time period. It is non-nuclide specific and provides information covering large areas. For emergency preparedness purposes it is of specific importance as it can provide 'real-time' information about the progression of the radioactive cloud. Therefore all European countries already have or are establishing automatic monitoring networks for ambient gamma dose rate. Table 1 illustrates how the mean distance between stations ranges from approximately 10 km to 150 km (values for Iceland and Russia excepted because of the non uniform spatial distribution of the stations considered in this study).

Country	Area (1000 km²)	No. stations	Mean distance* (km)
Albania	29	5	76
Austria	84	346	16
Belarus	208	22	97
Belgium	31	192	13
Bulgaria	111	26	65
Cyprus	91	7	36
Czech Republic	79	413	14
Denmark	43	13	58
Estonia	43	10	66
Finland	305	263	34
France	544	157	59
Former Yugoslav Republic of Macedonia (FYROM)	26	1	160
Germany	357	2 067	13
Greece	132	24	74
Hungary	93	77	35
Iceland	103	5	144
Ireland	70	14	71
Italy	301	57	73
Latvia	65	17	62
Lithuania	65	21	56
Luxembourg	2.6	18	12
Malta	316	1	18
Netherlands	34	191	13
Norway	324	28	108
Poland	313	35	95
Portugal	92	17	74
Serbia & Montenegro	103	5	143
Slovak Republic	49	63	28
Slovenia	20	42	22
Spain	505	914	23
Sweden	411	35	108
Switzerland	41	116	19
United Kingdom	244	93	51

* The mean distance is obtained by the square root of the surface of the national territory divided by the amount of monitors.

Table. 1. Number of automatic gamma dose-rate monitoring stations per country (Bossew, P. et al, 2008)

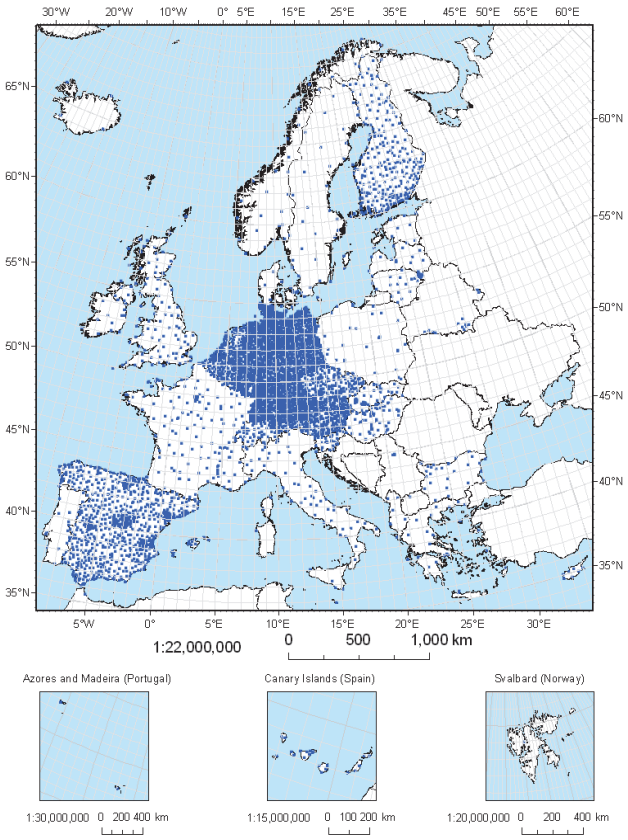


Fig. 2. Gamma dose rate monitoring networks in a number of European countries (Bossew, P. et al, 2008)

The detectors most frequently used are energy-compensated, gas-filled counter tubes, mostly Geiger-Müller (GM), ionization chambers and proportional counters. Occasionally scintillation detectors (NaI crystals) or even germanium spectrometers are used, in particular to determine the nuclide composition of the cloud by gamma spectroscopy. Some countries, e.g. Germany, are exploring to replace a part of their GM tubes by detectors that provide nuclide specific information of the ambient radiation, such as CdZnTe semi-conductors that operate at room temperature.

In order to cover the measuring range from natural radiation up to high dose-rate levels encountered during accidental releases, two counters are usually built into a detector, one for low-range and one for high-range measurements. Modern detectors are equipped with local electronics for data transfer. They may also feature data storage and perform automatic background correction.

A drawback, however, is that local dose rate measurements are very sensitive, so that even minor variations of the natural radioactivity concentration can be detected (e.g., due to radon daughters during precipitation or a pressure drop – see Fig. 3.).

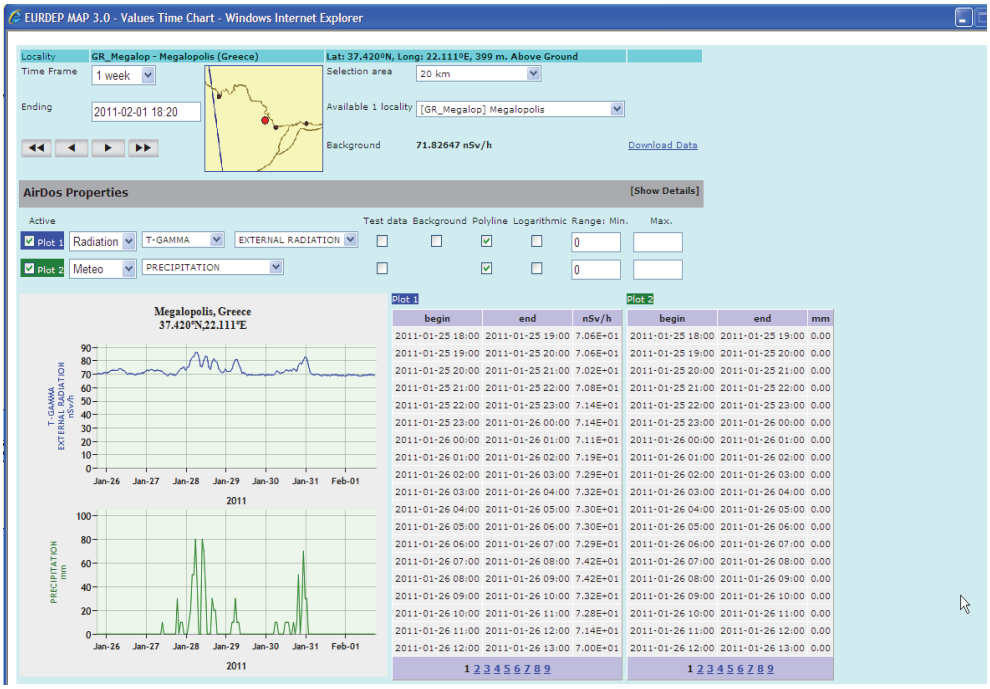


Fig. 3. Example of the effect of precipitation (lower left) on the ambient dose equivalent rate (middle left), due to radon daughter washout (http://eurdep.jrc.ec.europa.eu)

Furthermore, time-integrated dosimeters (film badges and thermo-luminescent dosimeters (TLDs)) are used, mostly on the perimeter fence of nuclear installations in order to obtain a cumulative measurement of the direct gamma radiation of the plant. Results from film dosimeters have to be treated with care, especially for low doses. They may also be affected by temperature and humidity (Hurst & Thomas, 2004).

4. Monitoring networks

4.1 Legislative background – information exchange and access

It is obvious that the use of radioactive and nuclear materials holds an increased risk for health. Therefore monitoring of environmental radioactivity is subject to strict legal obligations. This paragraph describes more into detail which legislation is in place and which measures have to be taken by the EU Member States.

It is normal practice to monitor not only exposure of critical groups but also to review exposure of the population at large. Within the European Community there is an obligation to do so both in terms of Article 35 of the Euratom Treaty and in terms of Article 13 and 14 of the Basic Safety Standards (EC, 1996).

Chapter III of the Euratom Treaty deals with Health and Safety aspects of the development and growth of nuclear industries and in particular with the establishment of uniform safety standards to protect the health of workers and of the general public. Article 35 deals with radioactivity levels in the air, water and soil.

Consequently Member States have set up programmes to monitor radioactivity levels in the air, water and soil over their entire territory. The main responsibility for establishing an environmental monitoring programme lies with the Member States. Under Article 35 the Commission also has access right to the monitoring facilities in order to verify their operation and efficiency.
The authorities shall keep the Commission informed of radioactivity levels to which the population is exposed, by periodically communicating data obtained with the facilities referred to in Article 35. The data are transmitted to the Commission in accordance with Article 36 of the Euratom Treaty. The Commission adopted a Recommendation to Article 36 (2000/473/ Euratom) (EC, 2000b) which specifies in much more detail the sample type and corresponding nuclide measurements that should be reported on, as well as the sampling frequency and the way in which these data have to be submitted to the European Commission. The routine environmental monitoring data are transmitted by the European Union (EU) Member State authorities to the Radioactivity Environmental Monitoring (REM) data base. This database is the basis for preparing the annual reports describing the radioactivity levels in the EU. Both REM database and the annual reports can be consulted by the public (http://rem.jrc.ec.europa.eu).
The vastness of the radioactive contamination following the Chernobyl accident also emphasized the need for improved international collaboration on emergency response. Shortly after the accident a legal framework was realised by establishing common international procedures for notification, data and information exchange and mutual assistance (i.e., the Conventions on Early Notification and Early Assistance by the International Atomic Energy Agency, and Council Decision 87/600 of the European Commission); these procedures were subsequently adopted in the national legislation of the Member States. The official notification system of the EC is the ECURIE system, which is the technical implementation of Council Decision 87/600. An EU Member State that in case of a radiological or nuclear emergency decides to take countermeasures to protect its population against an increase of radioactivity must promptly notify the European Commission. Upon receipt and verification, the EC will immediately forward this information to all Member States, after which the latter are required to inform the EC at appropriate intervals about the measures they take and the radioactivity levels they have measured.
EURDEP is a system by which automatic monitoring results (currently mainly ambient dose equivalent rate, but also air concentration data) are exchanged, irrespective of an emergency situation or not. Countries using EURDEP are exempted of sending the same radioactivity measurement results by ECURIE.
Although it has been designed originally for Europe, there are concrete plans between the EC and IAEA to extend the system to a world-wide coverage. At this moment about 4400 detectors in 35 European countries exchange continuously gamma dose rate measurements. The information can also be accessed by the public (see http://eurdep.jrc.ec.europa.eu)

4.2 Routine monitoring

The routine monitoring programme, which has been developed in parallel to nuclear industry, essentially aims at verifying:

- the discharge authorisation of nuclear installations;
- compliance with the dose limits or constraints laid down for protecting the population.

Therefore one should distinguish between national monitoring programmes and surveillance of nuclear installations.
The national monitoring programme has been designed to verify compliance of the Basic Safety Standards. Hence the monitoring programme should aim at providing information on the overall dose received by the population at large. The monitoring network is therefore to be designed so that the results are representative on a national level. This means that sampling locations must be far enough from nuclear installations in order to avoid direct influence of discharges, that sampling must be made of all the compartments of the environment and that sampling intervals must be performed at time periods that allow verification of dose limits (EC, 1996).
In addition there are environmental monitoring programmes for nuclear facilities in operation. These depend on the type of discharges, whether into the atmosphere and/or to the water. Verification of releases from nuclear installations is best performed by installing equipment in the stack and in the liquid effluent discharge line of the plant. Monitoring stations installed in the immediate vicinity of the plant ensure independent but indirect verification by the authorities. A network located at short distance and in an appropriate number of wind directions and downstream of the discharge point in the river or sea, assures this. In addition, estimation of the actual doses received by the critical groups requires that local foodstuffs be sampled and analysed.
Usually, for routine conditions, radioactivity levels are below the detection levels so that there is a trend to make more use of bio-indicators such as lichen, aquatic mosses and algae to trace accumulated radioactivity (EC, 2000b).

4.3 Emergency preparedness monitoring

4.3.1 The need for emergency preparedness monitoring

Any notification message about an emergency at a nuclear facility, warning that a significant amount of radioactivity has been released or is likely to be released, should come from the plant operator. This information will then be further disseminated through the proper communication channels on regional, national and international level (De Cort et al, 2007). Nevertheless many countries have set up mechanisms for early warning based on nationwide gamma dose-rate monitoring networks. This would allow the emergency situation to be identified independently and irrespective of whether the release was of domestic or foreign origin.
For this purpose it is necessary to be able to identify the source and to circumscribe the extent of the radioactive cloud with reasonable accuracy. Identification and circumscription are possible only if the density of the network is sufficiently high and if data from neighbouring countries can be relied upon in case of a transboundary movement.

4.3.2 Network design criteria

The basic requirement of an emergency preparedness network is its 'early-warning' capability, by which is meant that it must be able to detect a cloud resulting from an accidental release. Next it must be able to delineate the extent of the cloud and to track its progression over a territory.
Although they are mainly designed for the early detection of accidental releases, gamma dose-rate monitoring networks have limited value for circumscribing and monitoring the evolution of such clouds because they cannot readily discriminate between airborne and deposited radioactivity. Continuous aerosol monitoring devices are therefore preferred,

but they have the drawback of having higher maintenance costs. Also out of financial budgetary reasons, a limited amount of gamma spectrometers is usually incorporated in an on-line network to give qualitative information about the composition of the cloud. After the passage of the cloud, the latter in combination with the gamma dose-rate readings, allow for the assessment of the amount of radioactivity deposited to the ground.

Gamma detectors, in view of their relatively low installation and maintenance cost and wider range of sensitivities, best serve the alarm function of the national monitoring networks. Geiger-Müller detectors, proportional counters, ionisation chambers or scintillation crystals may equally be preferred. The existence of so many systems has the drawback of differences in energy response. For the mere alarm function an accurate calibration (and knowledge of the nuclide composition of the cloud) is not a priority, but when absent, it may create problems when data from different networks are brought together.

Irrespective of the topology of the network, the representativeness of the gamma dose-rate measurements depends on many factors, such as the presence of trees (which enhance dry deposition), the presence of surfaces that promote surface runoff of radioactive material reavenged by precipitation (e.g., paved surfaces, roofs), attenuating obstructions (e.g., vicinity of buildings, walls), the surface roughness and the detector position above ground (e.g., terrain or roof). Differences of more than one order of magnitude may occur between different sites which have been contaminated under the same conditions. As is shown in (Sohier, 2002), these differences can be characterised by environmental parameters, on which basis the measured data can be interpreted correctly. In addition this parameterisation could be used to improve harmonisation between different national networks.

When a network is being designed, a number of general factors should be taken into account:

4.3.2.1 Natural background subtraction

Gamma detectors are also very sensitive to variations in background count rates. Apart from regional variations for which corrections can be made, there are also slow and fast temporal variations due to the dependence of radon daughter concentrations in air and soil on meteorological parameters (e.g., an enhanced soil exhalation in case of a pressure front passing by or a cyclonic pressure drop, or washout of atmospheric aerosols in case of rain). Automatic data processing for pattern recognition of such phenomena or correlation is difficult and in general one relies on the judgement of the operator in the central station for data collection.

Nevertheless alarm levels can be set at dose rates on the order of 100 nSv h^{-1} above background, which is certainly low enough.

Continuous aerosol monitors are preferable because they are not affected by ground deposition. However they consist of moving and consumable parts, so maintenance costs are higher. Continuous beta monitoring can be performed, e.g., by putting a beta detector in front of a slowly rotating paper ribbon. As already mentioned in section 3.2.2, there is the additional complication of having to subtract counts from radon daughters.

A distinction between natural and artificial radioactivity can also be achieved by gamma ray spectroscopy. Knowledge of the radionuclide composition of the release is of interest for predicting further transport and deposition of the cloud and for assessing population doses. The composition is not likely to vary quickly in the course of time, so it is sufficient to

perform such measurements only in a few stations. Detection levels achieved by sampling on step-feed filters and on-line counting with a germanium detector are on the order of 0.05 Bq m^{-3} (typically 2 hours measuring time).

4.3.2.2 Sampling/counting time

For continuous monitoring systems sampling and counting times are identical. In view of the alarm function of such systems one might be tempted to reduce the sampling/counting time at the expense of statistical error which in turn would demand an increase in alarm levels. This is particularly true for continuous beta detectors. In the case of gamma detectors it is to a large extent possible to choose a more sensitive instrument.

Counting times should nevertheless be sufficiently short so that the network yields information on the speed, direction and longitudinal elongation of the cloud. On the other hand, counting times too short may cause the central processing unit to be overloaded with redundant data, i.e., the 'pace' of the network (average distance between the stations divided by the sampling time) should be comparable to the wind velocity.

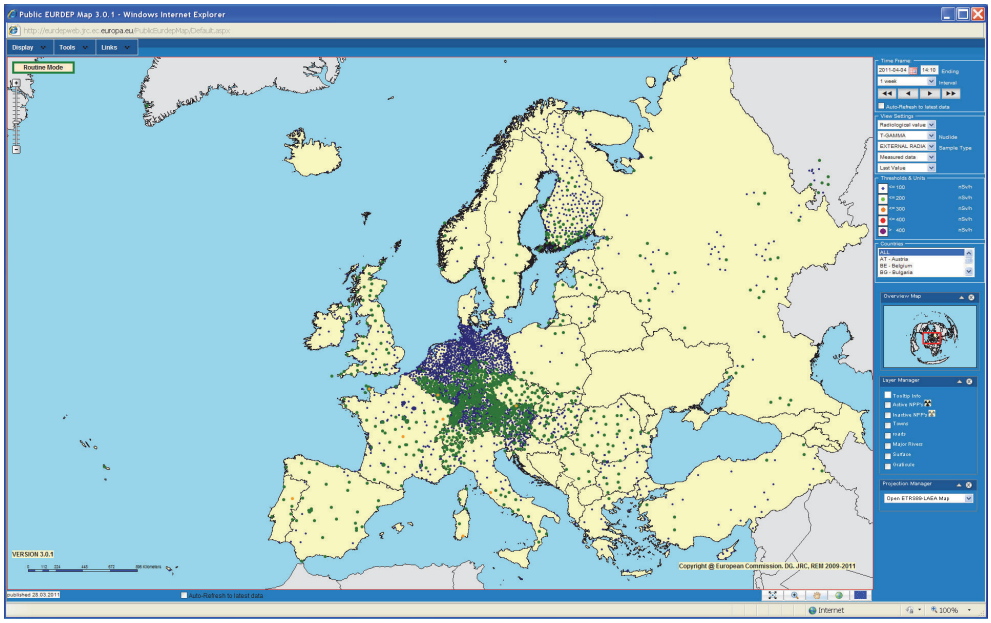


Fig. 4. Location of automatic gamma dose-rate monitoring stations that contribute to the EURDEP system (situation of April 2011), showing the topographic diversity of the national emergency preparedness networks.

4.3.2.3 Spatial homogeneity

The detection capacity of a network also depends on its spatial homogeneity. An elegant analysis of the impact of homogeneous coverage of the territory and of the limited extent of the network (countries' dimensions) can be made in terms of the 'factual' dimension of both

the network and the cloud. Such an analysis demonstrates the need for integrating networks on the largest possible scale and for comparable network densities (Raes et al., 1991; Sohier, 2002). Figure 4 which shows the automatic monitoring stations for gamma dose-rate that contribute to the EURDEP system, demonstrates the lack of spatial homogeneity between the national emergency preparedness networks.

4.4 Monitoring after an accident

4.4.1 General considerations

Even though a thorough analysis of continuous monitoring data may improve the accuracy of predicting further transport and dispersion of the cloud, which may prove critical for timely assessing its possible impact and for preparing adequate countermeasures, additional means are necessary in case of a real accident.

Early decisions such as exhorting the population to seek shelter could be based directly on reading of gamma detectors or aerosol monitors. Evacuation or the distribution of iodine tablets on the other hand might be considered either as a precautionary measure when a major release is expected or otherwise after the plume has passed, but only after the current radiological situation has been assessed. Static networks are inappropriate for this purpose because exposure pathways depend to a large extent on ground deposition. In the absence of rain and in case of a well-circumscribed cloud, calculated values for ground deposition may be available. However the Chernobyl accident has taught us that in case of rainfall one must cope with a patchy deposition pattern, which can be assessed only by portable and preferably mobile equipment.

4.4.2 Monitoring strategies

In case of a possible accidental release of radioactivity, the monitoring programme aims at answering the following questions:

- has an abnormal release occurred;
- is there an action to be taken;
- which remedial measures should be implemented.

As for operational monitoring the design of the programme depends on the nature of the environment which has received the contamination. It is, however, important in any emergency that the results are obtained relatively quickly, which means that rapid methods should be used for making measurements. This usually means a lower sensitivity and thus a greater risk of errors. On the other hand if the contamination is substantial, fast measurements are possible without losing much in sensitivity. Last but not least, at all times one should be aware that equipment can become contaminated, leading to incorrect measurements.

4.4.3 Mobile equipment

4.4.3.1 Terrestrial equipment

Vans have been fitted with all equipment needed for in-situ measurement of dose rates, and for collecting air samples with aerosol filters and charcoal cartridges, and possibly for soil and biological samples, and for their immediate counting and spectrometric analysis. The fitting of such vans, their maintenance, the provision and permanent training of personnel involves considerable costs. It is a matter of choice whether resources are allocated to the extension of static networks or to mobile equipment.

The mobile equipment should also have adequate means of communication and data transmission with the co-ordination centres (GPS navigational instruments and radar altimeter).
The need for environmental radioactivity measurements in case of an accident is not limited to rapid and detailed assessments prior to the implementation of countermeasures. Also in case of a remote accident involving a moderate release, it will be necessary to perform a large number of measurements in different media in order to perform an accurate a posteriori assessment of its radiological impact. Such an assessment fulfils the need for adequately informing the public and may also serve scientific investigations.
These kinds of measurements may be pursued over longer periods of time. The assessment of the impact of short-lived nuclides such as iodine and ruthenium isotopes will be possible only if measurements are performed within a few weeks after the accident.
The direct monitoring of internal contamination of people is usually done by means of whole body counters (WBC). The monitoring infrastructure is usually of a stationary type and installed at nuclear power plants, fuel cycle plants, research centres, hospitals with nuclear medicine, etc... However certain countries, such as France, also dispose of mobile equipment, usually in the form of intervention trailers, for monitoring contaminated persons. This equipment consists of vehicles equipped with 'partially body counters' (e.g., thyroid, thorax) and some even with WBC capacity (including appropriate shielding to reduce the background radiation). Detection limits (e.g. 500 Bq ^{137}Cs for 10 minutes and 100 Bq for 30 minutes counting time) are higher than those for fixed installations; however they allow inspectors to sort people according to their internal contamination. In order to make full use of these resources in case of emergency, the main problem is to have available enough personnel sufficiently trained to guarantee that this infrastructure is full used (Fiedler & Voigt, n.d.).

4.4.3.2 Airborne equipment

During the release phase, given the limitations of ground-based monitoring equipment for three-dimensional circumscription of the cloud, aircraft may prove very useful to carry out an aerial mapping at higher altitudes, in particular above the mixing boundary layer (e.g., 1500 m).
As soon as the release has ceased, airborne gamma ray spectrometry allows for rapid monitoring of vast areas contaminated by radioactivity. The technique has been described in section 3.4. Since the Chernobyl nuclear power plant accident, many European countries have become capable of deploying measurement teams within hours of an accident being notified, in order to measure levels and pattern of deposited radioactive material. In view of combining national efforts in case of a large scale nuclear accident, European teams have been collaborating and exercising on various occasions over the past years. This capacity to act jointly and produce, in real-time, a composite map of an area in Southern Scotland has been demonstrated in the ECCOMAGS project (Sanderson et al, 2004).

5. Discussion and conclusions

Competent Authorities of the European Union Member States have to ensure that the exposure of their population is compliant with the Basic Safety Standards (EC, 1996). To reach this objective European countries have set up environmental monitoring programmes which provide continuously the basic information, i.e., the radioactivity levels in the various compartments of their environment. Subsequently the doses received by the population can be assessed by means of radioecological models. The latter thus play an important role in

designing sound environmental monitoring programmes, including the definition of potentially important pathways and critical groups by determining the most representative sample types and sample locations (Vandecasteele, 2004). Also in the case of radiological or nuclear emergency, and based on the available monitoring and modelling information, Member States are required to submit emergency procedures and practices in accordance with the Basic Safety Standards. At this moment, the latter is being revised to allow for the new ICRP Recommendations (Publication 103) as well as to consolidate all EU radiation protection legislation in a single BSS Directive (Janssens, 2009).

Since the Chernobyl nuclear power plant accident, European countries continue to enhance their capacity and infrastructure to monitor radioactivity in their environment. They also continue to improve their capacity to transfer and handle the monitoring data in real time, combining them with radioecological models and/or decision support systems, in order to convert them into information based on which decisions may be made.

In case of an emergency with radioactive release to the atmosphere, during the release phase most of the information about the environmental contamination will come from the 'static networks', i.e., the emergency preparedness network and the routine monitoring networks, as these operate permanently. Such monitoring data, together with information about the release and appropriate models to forecast the atmospheric dispersion of the cloud, will be used by the authorities to decide on early countermeasures. However, atmospheric dispersion models can give rise to uncertainties in model results, which may lead to different approaches for early countermeasures on the national level. In order to harmonise the information coming from various countries and to work out a reconciled and comprehensive European long range atmospheric dispersion ensemble forecast, the ENSEMBLE system has been developed (http://ensemble.jrc.ec.europa.eu) by the European Commission Joint Research Centre. (Galmarini et al, 2008).

Once the release has stopped and the radioactive material is deposited, it is important to know, as soon as possible, the detailed deposition pattern of the affected areas. Airborne and in-situ gamma spectrometry is currently considered the most efficient way to do this. Combining such results with appropriate food chain models can significantly rationalise milk and crop monitoring programmes.

The vastness of the radioactive contamination following the Chernobyl accident also emphasized the need for improved international collaboration on emergency response. Shortly after the accident both the EC and the IAEA realised a legal framework for notification, data and information exchange and mutual assistance. This need to exchange monitoring data on international level raised questions regarding the representativeness of measurements coming from the various national monitoring networks. The current national monitoring networks are mainly based on historical, political and budgetary constraints. It is most unlikely that these will be converted into a unified operational European monitoring network. Hence the only practical and sustainable approach for the future is to harmonise to the maximum extent possible the data and information produced by these networks as far as they are to be exchanged on an international level. This means that we have to understand fully how monitoring is performed in the various European states. International intercomparison exercises only provide a part of the answer, as they mainly focus on the measurement aspect of monitoring, and usually do not address issues such as sampling and reporting techniques, including representativeness. Notwithstanding this, it is worthwhile to mention that a number of initiatives have already been taken to address this problem, in particular for measurements which are usually performed during the release phase or the

early post-release phase (i.e., the environmental parameterisation of gamma dose-rate detectors and the European collaboration of airborne gamma-ray spectrometry teams).

6. References

Aarkrog, A. (1996). Sampling and monitoring techniques, In: IBC 2nd residential training course on environmental radiation protection, 1996

Bossew, P.; De Cort, M.; Dubois, G; Stöhlker, U.; Tollefsen, T. & Wätjen, U. (2008). AIRDOS Evaluation of existing standards of measurement of ambient dose rate; and of sampling, sample preparation and measurement for estimating radioactivity levels in air. Final Report, project AA N° TREN/NUCL/S12.378241, JRC ref. N° 21894-2004-04 A1CO ISP BE.

Debauche, A. (2004). Continuous radioactivity monitoring systems. From the pre-history of radioprotection to the future of radioecology, *Journal of Environmental Radioactivity,* Vol. 72, pp. 103–108

De Cort, M.; de Vries, G. & Galmarini, S. (2007). International data and information exchange systems in Europe in case of radiological and nuclear emergencies, *International Journal of Emergency Management,* Vol. 4, No. 3, pp. 442-454

De Cort, M., Tollefsen, T., Marsano, A. and Gitzinger, C. (Eds.). (2009). *Environmental Radioactivity in the European Community 2004-2006*, EUR 23950 EN, Office for Official Publication of the European Communities, ISBN 978-92-12984-1, Luxembourg

Dickson, B. L. (2004) Recent advances in aerial gamma-ray surveying, *Journal of Environmental Radioactivity,* Vol. 76, No. 1-2, pp. 225–236

Dubois, G. & Bossew, P. (2003). Spatial analysis of ^{137}Cs in the environment: an overview on the current experience in Mapping radioactivity in the environment: Spatial Interpolation Comparison 1997, G. Dubois, J. Malczewski and M. De Cort (Eds.), EUR 20667 EN, Office for Official Publication of the European Communities, Luxembourg, pp. 21-36

EC (1989). Council Regulation of (Euratom) No. 2218/89 amending regulation (Euratom) No. 3954/87 laying down the maximum permitted levels of radioactive contamination of foodstuffs and feeding stuffs following a nuclear accident or any other case of radiological emergency, In: Official Journal of the European Communities L211 of 27 July 1989

EC (1996). Council Directive (96/29/euratom) of 13 May 1996 laying down basic safety standards for the health protection of the general public and workers against the dangers of ionizing radiation, In: Official Journal of the European Communities L-159 of 29/06/96

EC (1998). Council Directive 98/83/EC on the quality of water intended for human consumption, In: Official Journal of the European Communities L330 of 05/12/1998

EC (2000a). Commission Regulation 616/2000 of 20 March 2000 amending Regulation (EEC) 737/90 on the condition governing imports of agricultural products originating in third countries following the accident at the Chernobyl nuclear power station In: Official Journal of the European Communities L75 of 24/03/2000

EC (2000b). Commission Recommendation (2000/473/Euratom) of 8 June 2000 on the application of Article 36 of the Euratom Treaty concerning the monitoring of the level of radioactivity in the environment for the purpose of assessing the exposure

of the population as a whole, In: Official Journal of the European Communities L191 of 27/07/2000

EC 2003). Commission Regulation 2003/120/EC on the protection and information of the public with regard to exposure resulting from the continued radioactive caesium contamination of certain wild food products a consequence of the accident of Chernobyl nuclear power station, In: Official Journal of the European Communities L-47/54 of 21/02/2003

Fiedler, I. & Voigt, G. (Eds.), (n.d.). Review of infrastructures and preparedness systems in France, Germany and United Kingdom for potential releases of radioactivity into the environment, deliverable 1 of the EC project SAGE, In: http://www.ec-sage.net/D04_01.pdf, 30 April 2011, Available from http://www.ec-sage.net/

Frenzel, E. (1993). Environmental Monitoring: State of the Measuring Technique and Outlook, *Radioprotection*, pp. 255 - 261, February 1993

Galmarini, S.; Bianconi R.; de Vries G. & Bellasio R. (2008). Real-time moniotring data for real-time multi-model validation: coupling ENSEMBLE and EURDEP, *Journal of Environmental Radioactivity,* Vol. 99, pp. 1233-1241

Gering, F.; Hillmann, U.; Jacob, P. & Fehrenbacher, G. (1998). In situ gamma-spectrometry several years after deposition of radiocaesium: II. Peak-to-valley method, *Radiation Environ Biophys*, Vol. 37, pp. 283-291

Hurst, M.J. & Thomas, D.W. (2004). Report on radioactivity discharges and environmental monitoring the nuclear power stations during 1990, HS/NSOB/HP-R/003/91, Nuclear Electric, October 1991

Janssens, A. (2009). Progress with the revision of the EURATOM Basic Safety Standards and consolidation with other Community legislation, Annalen van de Belgische Vereniging voor Stralingsbescherming. Vol. 34, No. 3, pp. 173-187, ISSN 0250-5010

Lahtinen, J. (2004). Radiation monitoring strategy: factors to be considered, *Radiation Protection Dosimetry* (2004), Vol. 109, No. 1-2, pp. 79-82

Mercer, J. A.; Nisbet, A. F. & Wilkins, B. T. (2002). Management options for food production systems affected by a nuclear accident. Task 4: Emergency Monitoring and Processing of Milk, NRPB-W15 (June 2002)

Raes, F. (1989). Monitoring of Environmental Radioactivity in the European Community: inventory of Methods and Development of Data Quality Objectives, EUR 12801, Office for Official Publication of the European Communities, Luxembourg, 1989

Raes, F.; Graziani, G. & Girardi, F. (1991). A simple and fractal analysis of the European on-line network for airborne radioactivity monitoring, *Environmental Monitoring and Assessment,* Vol. 18, pp. 221-234

Sanderson, D.C.W.; Cresswell, A.J.; Scott, E.M. & Lang, J.J. (2004). Demonstrating the European capability for airborne gamma spectrometry: Results from the ECCOMAGS exercise. *Radiation Protection Dosimetry*, Vol. 109, No. 1-2, pp. 119-125, ISSN 0144-8420

Saxen, R. (1990). Radioactivity of surface water and freshwater fish in Finland in 1987, STUK-A77, May 1990

Sohier A. (ed.) (2002). A European manual for 'Off-site Emergency Planning and Response to Nuclear Accidents', SCK-CEN report R-3594, December 2002, p.341

Sombré, L. & Lambotte, J. M. (2004). Overview of the Belgian programme for the surveillance of the territory and the implications of the international recommendations or directives on the monitoring programme, *Journal of Environmental Radioactivity*. Vol. 72, pp. 75–87

Tyler, A. N. (2004). High accuracy in situ radiometric mapping, *Journal of Environmental Radioactivity*, Vol. 72, No. 1-2, pp. 195 – 202

Vandecasteele, C. M. (2004). Environmental monitoring and radioecology: a necessary strategy, *Journal of Environmental Radioactivity,* Vol. 72, No. 1-2, pp. 17 – 23

9

Origin and Detection of Actinides: Where Do We Stand with the Accelerator Mass Spectrometry Technique?

Mario De Cesare
CIRCE - INNOVA, Dipartimento di Scienze Ambientali - II Università di Napoli, Caserta Italia

1. Introduction

The activation products in a reactor's primary coolant loop is a main reason why Nuclear Power Plants (NPPs) use a chain of two or three coolant loops linked by heat exchangers, but isolated against matter exchange. However, during routine operations, damages are likely to occur, and the formation of micro-cracks represents a way of leakage for activation products. Through the same way, also leakage of material directly from the reactor core can take place: fission products and several isotopes of actinides, ^{235}U, ^{238}U, ^{239}Pu and those heavier isotopes produced by single and multiple neutron capture on them. The explosion of a nuclear bomb leads itself to releases of radioactive material, which, in this case, could be widespread in the environment up to a global scale.
Moreover, undeclared nuclear activities and/or illegal use and transport of nuclear fuel can be detected through determination of the isotopic ratios of U and Pu in such samples. As all the activities related to nuclear technology have been, since the second half of the past century, responsible of changes in the concentration of the anthropogenic radionuclides over global scale, it would be not inappropriate referring to a further Global Environmental Change which is recognizable within every environmental matrix: soil, superficial and ground-waters and living organisms. The outcome of the interactions of ionizing particles with living matter is, already, unluckily, well known, that's why a number of studies have been flourishing in order to assess the radiological risk, by analyzing the environmental pathways of radionuclides and their behavior within the human food chain. In parallel, accurate assessment and monitoring of every source of radioactive contamination are required from the point of view of the prevention from this new scenario of radiological exposition (Quinto, 2007). Quantifying the releases and tracing their dispersion in the environment has traditionally been the task of Alpha Spectrometry (AS (O'Donnell et al., 1997; Sanchez et al., 1992) or, more recently, of Inductively Coupled Plasma Mass Spectrometry (ICP-MS) (Ketterer et al., 2003; Ketterer & Szechenyi, 2008; Wyse et al., 2001) or Thermal Ionization (TI-MS) (Beasley et al., 1998; Richter et al., 1999). Although these are mature methodologies, each has its limitations, which are largely surmounted by the relatively new technique of Accelerator Mass Spectrometry (AMS) (Fifield, 2008; Steier et al., 2010). In the chapter the origin of actinides is discussed as well as the potential of actinides to serve as a tracer for geomorphologic processes and as sensitive

fingerprints of releases from nuclear facilities (Quinto et al., 2009; Sakaguchi et al., 2009; Steier et al., 2008).
Moreover the sensitivity of the different actinides measurements method and the peculiarity of the AMS technique with respect to AS and CMS techniques are illustrated. Furthermore the principles and methodology of heavy-element AMS as applied to U and Pu isotopes are described, and the ways in which these have been implemented in various laboratories around the world are discussed with particular attention to the CIRCE (Center for Isotopic Research on Cultural and Environmental Heritage) and ANU (Australian National University) accelerator facilities (Fifield et al., 1996; Terrasi et al., 2007).

2. Origin and detection of actinides

2.1 Sources of contamination

The discovery of the fission properties of ^{235}U and ^{239}Pu and the capability to sustain chain fission reactions, have opened the way to several applications of nuclear energy. Nuclear Power Plants are used to generate thermal, and then electric, power, while Nuclear Weapons use the destructive effects of supercritical chain reactions. Both of these applications of nuclear energy, together with all the activities correlated with production, transport and reprocessing of nuclear fuel, lead to releases of a wide range of radioactive elements in the environment.
The free neutrons produced by fission are partly adsorbed by other nuclei of fissile material, allowing, in the appropriate conditions, the self-sustainability of the chain reaction, and partly adsorbed by other elements in the air and the ground, during fission in nuclear explosions, and by structural and cladding materials and coolant circuits, during the exercise phase of NPP. The production of activation products in a reactor's primary coolant loop is a main reason why NPPs use a chain of two or three coolant loops linked by heat exchangers, but isolated against matter exchange. However, during routine operations, damages are likely to occur, and the formation of micro-cracks represents a way of leakage for activation products. Through the same way, also leakage of material directly from the reactor core can take place: fission products and several isotopes of actinides, ^{235}U, ^{238}U, ^{239}Pu and those heavier isotopes produced by single and multiple neutron capture on them. The input of these radionuclides in the outer coolant circuit results in their release in the environment within airborne and liquid effluents during the routine running of a NPPs. Of course, other ways of leakage would be accidents leading to uncontrolled explosions of the reactor or any kind of exceptional damage at various structural devices. The diffusion area of contamination from routine exercise of NPPs is restricted to a local scale, affecting the surroundings, the river or artificial channel, which provides and drains the reactor cooling water and the following areas connected with the river water circulation (Quinto, 2007). On the some grounds, the above radionuclides can be dispersed in the environment by the explosions of nuclear weapons in the atmosphere or underground.
Schematically it is possible to divide the sources of contamination into **large and small scale contamination**:

1. large scale source of radionuclide contamination includes the explosion of a Nuclear bomb. The contamination could be widespread in the environment up to a global scale. Several atmospheric nuclear tests conducted between 1945 and 1963 by the US and Soviet Union, by UK between 1952 and 1963, by France until 1974 and by the People's Republic of China until 1980, projected their debris into the stratosphere, from where, due to the global wind circulation, they have been deposited around the world within what is called **global fall**

Actinides	$T_{1/2}(y)$	Emitted radiation
^{234}U	$2.5 \cdot 10^5$	α
^{235}U	$7.0 \cdot 10^8$	α
^{236}U	$23.0 \cdot 10^6$	α
^{238}U	$4.5 \cdot 10^9$	α
^{238}Pu	88.0	α
^{239}Pu	$2.4 \cdot 10^4$	α
^{240}Pu	$6.6 \cdot 10^3$	α
^{241}Pu	14	β^-
^{242}Pu	$3.8 \cdot 10^5$	α
^{243}Pu	$5.7 \cdot 10^{-4}$	β^-
^{244}Pu	$82.0 \cdot 10^6$	α

Table 1. Half-lives, $T_{1/2}$ and decay mode of Uranium and Plutonium isotopes in year, y, units.

out. Another large scale source of contamination is due to one of the worst accidents in the history of nuclear energy that occurred on 26 April, 1986, at the Chernobyl Nuclear Power Station near Kiev in Ukraine, affecting mainly Central and Northern Europe, although ^{137}Cs was detectable even in Southern Italy (Roca et al., 1989).

2. small scale includes the operation and decommissioning activities of a NPP which could lead to airborne and liquid releases of radionuclides. At the same level, several steps in the fuel cycle, up to the reprocessing of spent fuel, can release activation and fission products, as well as the fissile material itself.

Obviously, given that the relative concentrations of plutonium and uranium isotopes depend on the nature of the source material and on its subsequent irradiation history, all these sources of contamination do not give the same contributions of contamination. As it will be shown in the following, useful tools to solve among different contributions are the isotopic ratios: $^{236}U/^{238}U$, $^{240}Pu/^{239}Pu$, $^{242}Pu/^{239}Pu$, $^{244}Pu/^{239}Pu$ and $^{238}Pu/^{239+240}Pu$,. Table 1 shows the half lives of the relevant isotopes of U and Pu.

2.2 Different contamination sources

The relative concentrations of plutonium and uranium isotopes depend on the nature of the source material and on its subsequent irradiation history; all these sources of contamination do not give the same contributions of contamination.
Here are shown some example of different contamination sources:

- Being fissile material, ^{239}Pu is the most abundant isotope in weapon-grade plutonium. The average ratio of $^{240}Pu/^{239}Pu$, before detonation is $^{240}Pu/^{239}Pu \leq 0.07$ while after detonation is $^{240}Pu/^{239}Pu \simeq 0.35$ (Diamond et al., 1960), for the US tests. After detonation ^{239}Pu isotope is still the most abundant because the ratio is always less than one.

  ^{239}Pu is produced from ^{238}U via neutron capture where ^{238}U is the most abundant isotope of uranium in nature, $^{238}U \simeq 99.275\%$, $^{235}U \simeq 0.720\%$ and $^{234}U \simeq 0.005\%$. During detonation of nuclear weapons and running of nuclear reactors, ^{239}Pu undergoes neutron capture to generate ^{240}Pu, and also the heavier ^{241}Pu, ^{242}Pu and ^{244}Pu are produced through successive neutron captures.

  The resulting short-lived ^{239}U ($T_{1/2}$ = 23.45 min) decays by β^- into ^{239}Np, which in turn decays by β^- ($T_{1/2}$ = 2.356 days) into ^{239}Pu:

$$^{238}U \xrightarrow{n} {}^{239}U \xrightarrow{\beta^-} {}^{239}Np \xrightarrow{\beta^-} {}^{239}Pu \xrightarrow{n} {}^{240}Pu \xrightarrow{n} {}^{241}Pu \xrightarrow{n} {}^{242}Pu.....$$

In weapon test fallout, the ratio ^{240}Pu/^{239}Pu varies depending on the test parameters in the range of 0.10-0.35. The average for the Northern hemisphere is about 0.18, (Koide et al., 1985).

Significantly different values, in the range 0.035-0.05, are found in Mururoa and Fangataufa atoll sediment, because of the particular nature of French testing, (Chiappini et al., 1999) and (Hrneceka et al., 2005).

- In nuclear reactors, as mentioned before, due to the different composition of fuels, uranium enrichment and burn-up degree, characteristic relative abundances of plutonium isotopes will be obtained: ^{240}Pu/^{239}Pu increases with irradiation time, which, in turn affects ^{238}Pu/$^{239+240}$Pu.

  ^{238}Pu is produced by neutron capture from ^{237}Np, which is itself produced by two successive neutron captures from ^{235}U:

$$^{235}U \xrightarrow{n} {}^{236}U \xrightarrow{n} {}^{237}U \xrightarrow{\beta^-} {}^{237}Np \xrightarrow{n} {}^{238}Np \xrightarrow{\beta^-} {}^{238}Pu$$

  or via the fast-neutron induced ^{238}U(n,2n)^{237}U reaction:

$$^{238}U(n,2n)^{237}U \xrightarrow{\beta^-} {}^{237}Np \xrightarrow{n} {}^{238}Np \xrightarrow{\beta^-} {}^{238}Pu$$

  The ratio ^{238}Pu/$^{239+240}$Pu is useful to resolve between different sources in case they show similar ^{240}Pu/^{239}Pu, e.g., irradiated nuclear fuel in a PWR (Pressurized Water Reactor) with 7-20% of ^{235}U and burn-up 1.4-3.9 GW·d (GWatt·day) reaches ^{240}Pu/^{239}Pu isotopic ratios of 0.13, a value, that could be ascribed also to global fall out. On the other side, these two sources show quite different ^{238}Pu/$^{239+240}$Pu activity ratio, 0.025-0.04 for the global fallout and 0.45 for that nuclear fuel, (Quinto, 2007).

- Another valuable tool to identify a nuclear reactor origin of a radionuclide contamination is ^{236}U/^{238}U isotopic ratio. The dominant ^{236}U mode of formation is the capture of a thermal neutron by ^{235}U, a secondary contribution being the alpha decay of ^{240}Pu. Its concentration in nature has been heavily increased as a consequence of irradiation of enriched uranium in nuclear reactors. Several orders of magnitude of difference between the ^{236}U/^{238}U isotopic ratios in naturally-occurring uranium (10^{-9} to 10^{-13}) and in spent nuclear fuel (10^{-2} to 10^{-4}) imply that also a small contamination from irradiated nuclear fuel in natural samples is able to increase significantly the ^{236}U/^{238}U ratio measured in the whole sample.

2.3 Needs for actinides monitoring

The nuclear safeguard system used to monitor compliance with the Nuclear Non-proliferation Treaty relies to a significant degree on the analysis of environmental samples. Undeclared nuclear activities and/or illegal use and transport of nuclear fuel can be detected through determination of the isotopic ratios of U and Pu in such samples. Accurate assessment and monitoring of every source of radioactive contamination are required from the point of view of the prevention from radiological exposure.

Both the operations of decommissioning of the existing NPPs and the possible future operation of new plants demand accurate investigations about the possible contamination by radioactive releases of nuclear sites and neighboring territory and of structural

materials of the reactors. The monitoring activity of surveillance institutions uses assessed radiometric techniques, but more and more ultrasensitive methodologies for the detection and quantification of ultralow activity radionuclides is requested at international level.
Most of U and Pu isotopes are long lived alpha emitters with very low specific activity: their detection and the measurement of their concentration and isotopic abundance demands very high sensitivity, so that they are included among the so called "hardly detectable" radionuclides. As it will be shown in the following, the required sensitivity is often not achieved using conventional analytical techniques, such as counting of the radiation emitted in the decay or conventional mass spectrometry. The main task of the present work is to illustrate an ultrasensitive methodology for the detection of ultralow level radionuclides belonging to the actinides subgroup of the periodic table.
The method is based on a combination of AS and AMS: the reason for such a combination lies in the fact that it may be necessary to be able to measure the Pu isotopes at the fg level and the U isotopes where the total uranium content may be at the ng level or with a sensitivity as low as 10^{-13} in the measurement of the $^{236}U/^{238}U$ isotopic ratio in samples incorporating a total of about 1 mg of U. AMS will be shown to be the only technique able to achieve such a sensitivity together with unparalleled suppression of molecular isobaric interferences for the detection of rare isotopes of elements with (quasi)stable isotopes many orders of magnitude more abundant, such as U.
Nevertheless, the measurement of ^{238}Pu abundance is heavily suffering interference from the atomic isobar ^{238}U, about seven orders of magnitude more abundant, and cannot be achieved by any mass spectrometric technique; on the other hand ultra-low activity AS can isolate this isotope, while alpha particles from the decay of ^{239}Pu and ^{240}Pu cannot be energetically resolved. Combination of the two techniques provides the determination of the abundances of the full suite of Pu isotopes. Moreover, AS plays an important role also for the calibration of the spikes used as carriers for the AMS measurements and as overall cross check of the employed methodologies. An important role in pursuing the goal of ultrasensitive detection of actinide isotopes is played by the sample preparation procedures, which has to be performed in a very clean environment with ultralow contamination. The procedure to be setup will be able to isolate the elements of interest and produce samples in the form suitable for both AS and AMS. In the first case very thin and uniform layers have to be achieved, while purification respect to elements which can produce molecular interferences is of paramount importance for AMS. Preliminary sampling and conditioning of a properly representative sample; uranium and plutonium are separated from the sample following a systematic chemical protocol of pre-enrichment/separation; fractions of U and Pu are purified from every possible element that could cause radiochemical interference to AS; fractions of U and Pu must be converted into useful chemical and physical-chemical forms (De Cesare, 2009; Quinto et al., 2009; Wilcken et al., 2007).
Finally, besides the application of the developed technique to the assessment of actinide contamination of the NPP site and plant, a more general objective is to provide an ultrasensitive diagnostic tool for a variety of applications to the national and international community. Applications range across a broad spectrum. Isotopes of plutonium are finding application in tracing the dispersal of releases from nuclear accidents and reprocessing operations, in studies of the biokinetics of the element in humans, and as a tracer of soil loss and sediment transport. ^{236}U has also been used to track nuclear releases, but additionally has a role to play in nuclear safeguards and in determining the extent of environmental contamination in modern theaters of war due to the use of depleted uranium weaponry.

2.4 Alpha-spectroscopy and mass spectroscopy

^{236}U and ^{239}Pu are present in environmental samples at ultra trace levels (^{236}U concentration is quoted to be in the order of pg/kg or fg/kg and ^{239}Pu around 100 pg/kg) and are long-lived radionuclides (Perelygin & Chuburkov, 1997).

If one considers alpha-spectroscopy for the detection of ^{239}Pu, assuming an efficiency of 50 % and a counting time of one month, one gets 64 counts (with a statistical uncertainty of 12%) with a total activity of 50 μBq, which correspond to about 40 million atoms, or about 15 fg. In addition, alpha-particle counting is unable to resolve the two most important plutonium isotopes, ^{239}Pu and ^{240}Pu, because their alpha-particle energies differ by only 11 keV in 5.25 MeV. Hence, the information on their isotopic ratio readily difficult to extract.

The $23{\cdot}10^6$ y half-life of ^{236}U limits the utility of alpha-particle spectroscopy for this isotope.

For the detection of such small amounts one can exploit the sensitivity of mass spectrometric techniques. Conventional Mass Spectrometry, CMS, methods give information on the ^{240}Pu/^{239}Pu ratio, and potentially have higher sensitivity than alpha-particle counting with values as low as 1 fg, but are sensitive to molecular interferences. Both ^{236}U and ^{x}Pu isotopes have been measured using either Thermal Ionization (TI-MS) or Inductively Coupled Plasma (ICP-MS) positive ion sources. For plutonium isotopes, abundance sensitivity is not a problem due to the absence of a relatively intense beam of similar mass. Molecular interferences such as $^{238}UH^-$, $^{208}Pb^{31}P$, etc. may be a problem (Fifield, 2008). For uranium, isotope variability both in the molecular ($^{238}UH^-$) and in tail contributions of main beam of ^{238}U limits the sensitivity of ICPMS to ^{236}U/^{238}U ratios of $\sim 10^{-7}$. TIMS ion sources, on the other hand, produce both much lower molecular beams and much less beam tail and so a sensitivity of $\sim 10^{-10}$ in the ^{236}U/^{238}U ratio is reached.

So that, the measurements of these isotopic ratios requires the resolution of mass spectrometric techniques, but only AMS allows the sensitivity needed e.g. ^{236}U/^{238}U ratios of $\sim 10^{-13}$, 0.1 fg of ^{236}U with about 1 mg of U, as well as for the ^{239}Pu.

Although AMS has advantages over the other techniques for ^{239,240,242,244}Pu, there are two other isotopes, ^{238}Pu and ^{241}Pu, which are of interest in some applications. Since the concentration of ^{238}U is seven orders of magnitude higher than that of ^{238}Pu, no chemical procedure is efficient to separate uranium and plutonium fractions to allow the mass spectrometric measurement of ^{238}Pu. So alpha-spectroscopy remains the only suitable technique for the measurement of ^{238}Pu concentration. The β^- emitter ^{241}Pu can be measured with either AMS or with liquid scintillation counting. Its short half-life of 14 years results, however, in higher sensitivity for the latter (Fifield, 2008).

2.5 AMS of actinides isotopes

Actinides AMS measurements were pioneered at the IsoTrace laboratory in Toronto (CA) (Zhao et al., 1994; 1997), where the ^{236}U content in an U ore was determined using the 1.6 MV AMS system. Moreover, the relative abundances of Pu isotopes were measured at 1.25 MV. Then, at the Australian National University (AUS) (Fifield et al., 1996; 1997) the utilization of a higher terminal voltage (4 MV) allowed to improve the sensitivity of the method, both for the detection limit as the minimum detectable number of U atoms in the sample, and for the lower limit of isotopic ratio measurable in samples at high concentration. Similar detection system have been developed at the Vienna Environmental Reasearch Accelerator (VERA - AU) (Steier et al., 2002), at the Lawrence Livermore National Laboratory (LLNL - USA) (Brown et al., 2004), at the Australian Nuclear Science and Technology Organisation (ANSTO - AUS) (Hotchkis, 2000), at much lower energies at the Eidgenössische Technische

Hochschule - ETH in Zurich (CH) (Wacker et al., 2005), at Munich facility (GE) (Wallner et al., 2000) and at the accelerator of Weizmann Institute, Israel (Berkovits et al., 2000). New AMS actinides line based on 1MV and 3 MV tandems have recently been and will be installed, respectively, in Seville (Spain) and in the Salento (Italy). In both cases, they will be upgraded to perform actnides AMS measurements, being the injection and the analyzing magnets overdimensioned.
Two recent review papers (Fifield, 2008; Steier et al., 2010) summarize the results obtained in the laboratories active in the fields of actinides AMS. Summarizing, the two systems aiming to the best isotopic ratio sensitivity (ANU and VERA) have shown that it is possible to reach a sensitivity of 10^{-13} for ^{236}U in samples including about 1 mg of U. The ANSTO and LLNL laboratories quote a sensitivity respectively of about 10^{-8} and 10^{-9} with U amounts of the order of 1 ng. In the case of plutonium, there is no stable abundance isotopes available; the plutonium isotopic ratio is not a problem and a ^{239}Pu concentration background of about 0.1 fg (2.5×10^5 atoms) is achieved, limited by the process blank count rate. In both cases these limits surmount by several orders of magnitude alpha spectrometry and conventional mass spectrometry. In nature, U stable abundant isotopes exist. For that reason, the sensitivity limit for the isotopic ratio depends on the U concentration in the sample. Thus, the AMS task is, for environmental samples, to push the sensitivity in the isotopic ratio measurement down to natural abundances (^{236}U/^{238}U$\simeq10^{-9}$-10^{-13}) in samples with sizeable amounts of U ($\sim$ 1 mg). On the other hand, for anthropogenically influenced samples, the required sensitivity for the measurement of the isotopic composition is alleviated, but significantly smaller amounts of U have to be used (down to 1 ng). For Pu, where no stable isotope interferences are present, the goal is the maximum possible detection efficiency, allowing few hundred counts from less than 1 million atoms in the sample.
The CIRCE laboratory is one of the few systems in the world able to perform such a measurement (De Cesare et al., 2010a) and the only one in Italy. Moreover it is 1 order of magnitude higher (De Cesare et al., 2010b) with respect to the 2 systems (ANU and VERA) providing the best ^{236}U/^{238}U isotopic ratio sensitivity of 10^{-13}, in samples including about 1 mg of U; it has low uranium contamination background, less then 0.4 μg of ^{236}U (De Cesare et al., 2011). The CIRCE actinides group aims to reach and to exceed the isotopic ratio sensitivity goal with the upgrade: the utilization of a TOF system and, in case, the installation of a magnetic quadrupole doublet. Regarding the Plutonium background results, the CIRCE is one of the best systems in the world (De Cesare, 2009).

3. AMS facilities

In this paragraph the facilities where the author was mainly involved will be illustrated: CIRCE and ANU AMS systems.

3.1 CIRCE system

CIRCE is a dedicated AMS facility based on a 3MV-tandem accelerator (Terrasi et al., 2007). In contrast to many nuclear physics applications, the pre-treated sample material (a few mg is pressed into a 1 mm diameter Al cathodes and put in the ion source) itself is analyzed by two mass spectrometers which are coupled to the tandem accelerator. A schematic layout of the CIRCE facility is shown in Fig. 1.
The caesium sputter ion source is a 40-sample MC-SNICS (Multi Cathode Source for Negative Ions by Cesium Sputtering). A total injection energy of 50 keV is used; 50-300 nA ^{238}U^{16}O$^-$ molecules are energy selected by a spherical electrostatic analyzer (nominal bending radius

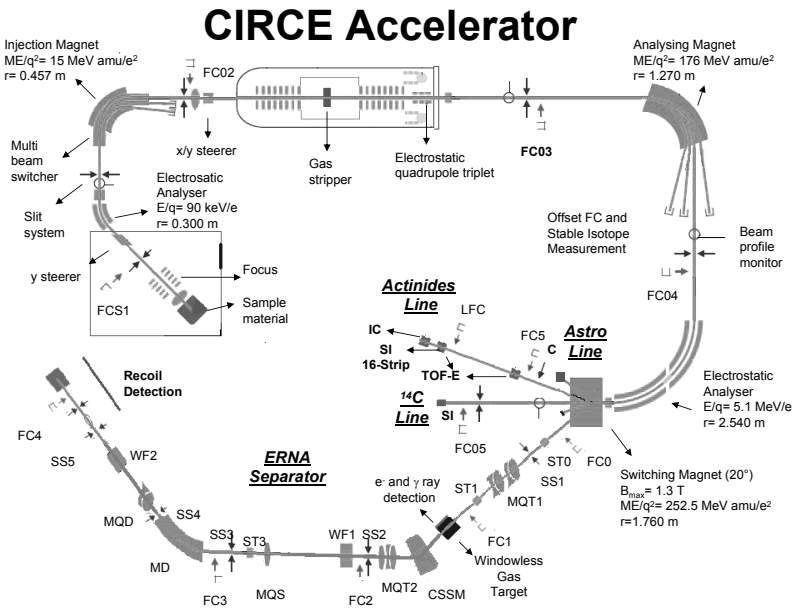


Fig. 1. Schematic layout of the CIRCE accelerator and of CIRCE accelerator upgrade with the actinides line and also the ERNA separator line, Astro line, besides the ^{14}C original line: the switching magnet already inserted and the start and the stop TOF-E detector not yet inserted. FC denotes Faraday Cup (LFC in the actinides line is Last Faraday Cup), C denotes the Collimator in the heavy isotope line and the arrows indicate a slit system. ERNA is the acronym of European Recoil separator for Nuclear Astrophysics.

r= 30 cm, plate gap= 5 cm) which cuts the sputter low energy tail of the beam, with a bending angle of $\pm 45^\circ$ and it is operated up to ± 15 kV. The maximum electric field strength is 6 kV/cm, resulting in an energy/charge state ratio of 90 keV/q. The 90° double focusing Low Energy (LE) injection magnet (r = 0.457 m, vacuum gap= 48 mm, ME/q^2 = 15 MeV·amu/e^2) allows high resolution mass analysis for all stable isotopes in the periodic table, mass resolution is M/ΔM$\sim$ 500 with the object and image slits set to $\pm$ 1 mm, (De Cesare et al., 2010a). The insulated stainless steel chamber (MBS) can be biased from 0 kV to +15 kV for beam sequencing (e.g. between $^{238}U^{16}O^-$, $^{236}U^{16}O^-$ and between $^{239}Pu^{16}O^-$, $^{240}Pu^{16}O^-$, $^{242}Pu^{16}O^-$).

The accelerator is contained inside a vessel filled with sulphur hexafluoride (SF_6) at a pressure of about 6 bar. Two charging chains supply a total charging current to the terminal; about 100 μA are delivered to the terminal for operation at 3.000 MV. Stabilization is achieved by GVM feedback on the charging system high voltage supply; the long term stability is about 1 kV peak-to peak. At the terminal the ions lose electrons in the gas stripper, where Ar is recirculated by two turbo-pumps. The working pressure is about 1.3 mTorr for $^{238}U^{5+}$ at 2.875 MV.

The ions with positive charge states are accelerated a second time by the same potential. The High Energy (HE) magnet, efficiently removes molecular break-up products (De Cesare et al., 2010a;b). The double focusing 90° HE bending magnet has r= 1.27 m, ME/q^2 = 176 MeV·amu/e^2 and M/ΔM = 725, with slit opening of ±1 mm both at object and image points. The two 45° electrostatic spherical analyzers (r = 2.54 m and gap = 3 cm) are operated up to ±60 kV; energy resolution is E/ΔE = 700 for typical beam size. A switching magnet (B_{max}= 1.3 T, r=1.760 m and ME/q^2= 252.5 MeV·amu/e^2 at the 20° exit) is positioned after the ESA. Finally the selected ions are counted in an appropriate particle detector, either a surface barrier detector or a telescopy ionization chamber. The control of the entire system, is handled by the AccelNet computer based system via CAMAC interfaces or Ethernet, and the acquisition system is ether AccelNet itself or FAIR (Fast Intercrate Readout) system, (Ordine et al., 1998).

3.1.1 CIRCE actinides measurement procedures

In this paragraph a description of the various steps of the ^{236}U and ^{x}Pu isotopes measurement are given. The relative abundance of ^{238}U in environmental samples is several order of magnitude (up to 13) larger than the ^{236}U. For this reason, while the number of events of ^{236}U are measured in the final detector, ^{238}U is measured as current in the high energy side. For the ^{x}Pu isotopes, since no natural and so abundant isotopes exist, they are all measure in the final detector. Before performing measurements of samples, a tuning of the transport elements up to the final detector is made by setting the accelerator parameters to the detection of ^{238}U. Then the MBS, the TV and the high energy ESA are scaled to select the rare isotopes. The sample preparation provides material that is sputtered as $^{x}U^{y}O^{-}$ and $^{z}Pu^{w}O^{-}$. The negative molecular ions, ex. $^{238}U^{16}O^{-}$, are accelerated to an injection energy of E_{inj}= 50 keV. To select different masses without changing the magnetic field, the energy of the ions inside the injection magnet is varied by applying an additional accelerating voltage to the bouncing system. The injected $^{238}U^{16}O^{-}$ ions are accelerated by the positive high voltage towards the stripper, where they loose electrons and gain high positive charge states. The positive ions are, then, accelerated a second time by the same potential in the high energy tube of the tandem. This for $^{238}U^{5+}$ results in an energy of E= 17.3 MeV with a terminal voltage of V= 2.900 MV. Ar is recirculated in the terminal stripper by two turbo-pumps; the working pressure is about 1.3 mTorr for $^{238}U^{5+}$ at 2.875 MV (De Cesare et al., 2010b) and the stripping yield achieved for $^{238}U^{5+}$ achieved is around 3.1%.

Molecular break-up products with mass over charge ratio (M/q) different from that of the wanted ion, are removed by the combination of the high energy (HE) magnet and an electrostatic analyzer (ESA) whose object point is the image point of the analyzing magnet. For heavy ion measurements, the object and image slits of the injection magnet are closed to ±1 mm, the slits of the analyzing magnet are closed to ±2 mm and a collimator of 4 mm diameter is positioned in the beam waist at the 20° beam line.

The tuning procedure at CIRCE is made by the optimization of HE magnet and ESA in the high-energy side: they are optimized by maximizing the $^{238}U^{5+}$ current in the Last Faraday Cup (LFC). The transmission efficiency between the HE magnet and LFC at 20° is 80 %, with the 4 mm collimator in.

Once the setup for the pilot beam $^{238}U^{5+}$ is found, the voltage at the chamber of the injection magnet, the terminal voltage and the voltage of the ESA are scaled to transmit $^{236}U^{5+}$. In order to measure the $^{236}U/^{238}U$ ratio, the measurement procedure is composed of three automatic steps:

1. measurement of $^{238}U^{5+}$ current at the high energy side in FC04.
2. the voltage on the magnet vacuum chamber, the terminal voltage and the ESA are then scaled to transmit $^{236}UO^{-}$ and a measurement of the count rate of $^{236}U^{5+}$ in the detector is performed.
3. repetition of step 1

Steps 1 and 3 are necessary to estimate, by linear interpolation, the value of $^{238}U^{5+}$ current at high energy side which would be measured simultaneously with $^{236}U^{5+}$ counting.
In order to measure the ^{x}Pu isotope ratios, the measurement procedure is composed of automatic steps:

1. tune the beam with the $^{238}U^{5+}$ current up to LFC.
2. the voltage on the magnet vacuum chamber, the terminal voltage and the ESA are then scaled to transmit $^{x}PuO^{-}$ and a measurement of the count rate of $^{x}Pu^{5+}$ in the detector is performed.
3. repetition of step 2 for all the plutonium isotopes are needed (ex. $^{242}Pu^{5+}$ spike for 18 s, $^{240}Pu^{5+}$ for 60 s and $^{239}Pu^{5+}$ for 30 s).
4. repetition of step 3 for 3 times.

3.1.2 CIRCE actinide results

Before the installation of a dedicated actinides beam line at CIRCE, preliminary results for the $^{236}U/^{238}U$ background ratio level at 0° line, rutinelly used for ^{14}C measurements, was of the order of $1{\cdot}10^{-9}$ (De Cesare et al., 2010a). The measurement was obtained with the "K. k. Uranfabric Joachimisthal" sample, the VERA in-house U standard, $(6.98\pm0.32)\times10^{-11}$ (Steier et al., 2008).
The main upgrade so far has been the addition of a switching magnet placed 50 cm after the exit of the high-energy ESA. The position of the magnet was decided by means of COSY infinity (Makino & Berz, 1999) magnetic optics simulation (De Cesare et al., 2010a), Fig. 2. This magnet provides a supplementary dispersive analyzing tool.
The abundance sensitivity results, using a 16-strip silicon detector, have shown that, in the upgraded CIRCE heavy ions beamline after the switching magnet installation, a background level $< 5.6\times10^{-11}$ has been reached, Fig. 3, compared to 3.0×10^{-9} obtained previously (De Cesare et al., 2010b; Guan, 2010).
Although most of the ^{238}U are suppressed at the injector side, by the analyzing magnet and electrostatic analyzer, a small fraction of this intense beam can still interfere with the ^{236}U measurement. The main reasons for this "leakage" of interfering ions are charge exchange processes due to residual gas in the system. Scattering on the residual gas, electrodes, slits or vacuum chamber walls can also allow the background to pass a filter. However, the scattering cross-section is in the order of 10^{-20} cm^2 whereas the cross section of charge changing is 10^{-16}-10^{-15} cm^2 (Betz, 1972; Vockenhuber et al., 2002).
Moreover, in the upgraded CIRCE heavy ions beamline, after the TOF-E installation, a background level of about 2.9×10^{-11}, summing over the central six strips, has been reached, compared to $\sim 5.6\times10^{-11}$ obtained with a 16 strip silicon detector alone. This small background reduction is attributed to the 1.6 ns time resolution mainly due to the thickness of the 4 μg/cm^2 LPA (Maier-Komor et al., 1997; 1999) carbon foil, (De Cesare, 2009).
The CIRCE laboratory is not so far from the two systems (ANU and VERA) that provide the best $^{236}U/^{238}U$ isotopic ratio sensitivity of 10^{-13}, in samples including about 1 mg of U.

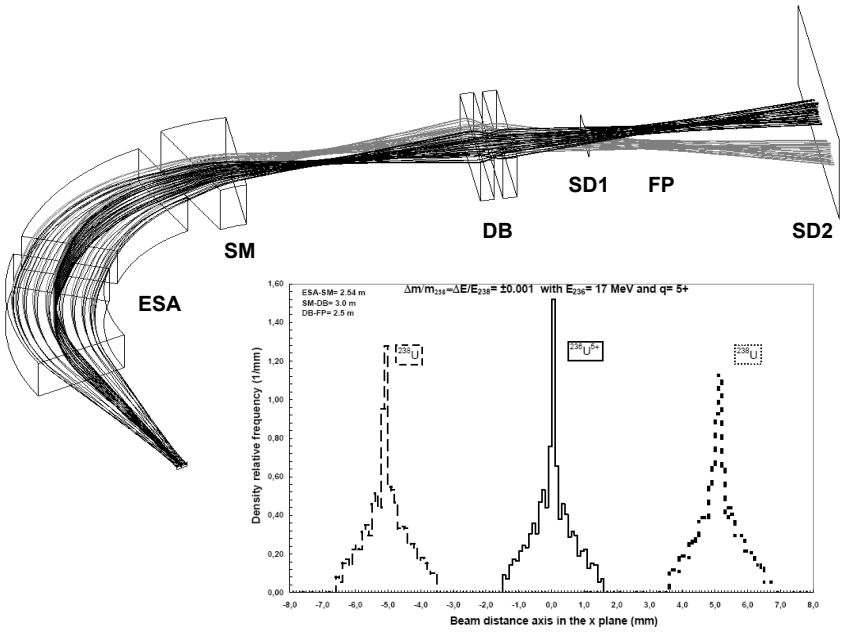


Fig. 2. The COSY infinity magnetic optics simulation is shown, where the development of two beams has been analyzed, starting from the waist of the high-energy magnet with a relative energy difference of $\Delta E/E = 0.001$ (corresponds to the resolution of the ESA). The adopted beam profiles are approximately Gaussian, with a halfwidth of 0.15 cm. A maximum divergence of 3 mrad was assumed. Simulations were performed for different geometric configurations. The distance ESA-SM (energy electrostatic analyzer-switching magnet), SM-DB (switching magnet-magnetic quadrupole doublet) and DB-FP (Focal Plane = the doublet focusing position) are shown in the upper part. The density relative frequency in function of beam distance in the x-plane is shown in the lower part. The central (solid line) peak is $^{236}U^{5+}$ and the dashed and dotted are the ^{238}U beams in the two opposite x positions, where the dashed one is not shown in the simulation

An overview of the planned upgrade of the CIRCE system using a TOF-E system, with a flight path of 3 m and a thinner DLC carbon foil, 0.6 $\mu g/cm^2$ is described in (De Cesare, 2009). Regarding the concentration sensitivity results, a 4μg uranium concentration sensitivity has been reached using only with the 16 strip silicon detector. That correspond to about 40 fg of ^{236}U and 10^8 ^{236}U atoms for a sample with isotopic ratio of 10^{-8} (De Cesare et al., 2011). For the ^{239}Pu concentration sensitivity results, the uranium background corresponding to the ^{239}Pu settings is at the level of 1 ppb. This is to be compared with the 10 ppm of ANSTO and 100 ppb of ANU. The CIRCE Lab. has at present a ^{239}Pu sensitivity level less than 0.1 fg, since 500 ng of uranium is required to produce an apparent ^{239}Pu concentration of 0.1 fg (De Cesare et al., 2011); for the Pu background level, CIRCE is one of the best system in the word.

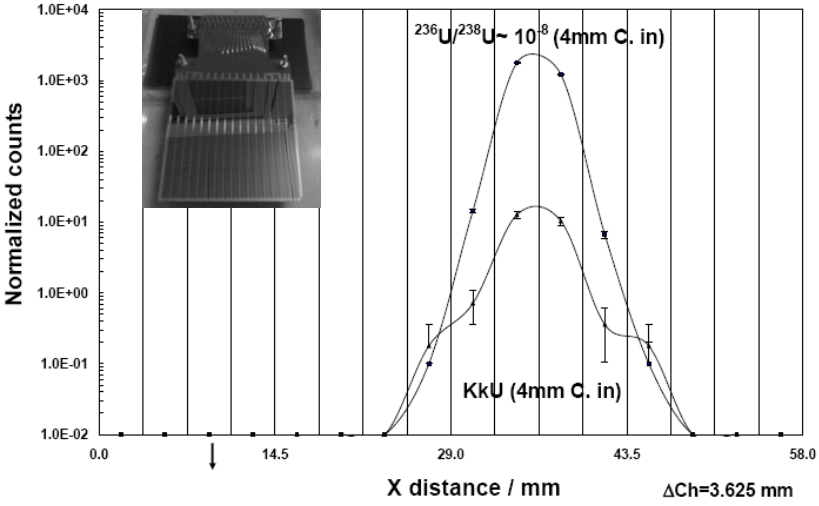


Fig. 3. Normalized counts (counts in the detector in 300 s over FC04 current corrected for the transmission $\sim$ 80% between FC04 and LFC) versus horizontal position of the 16-strip silicon detector. Ch= 3.625 mm is the distance between the center of two adjacent strips. A photo of the 16-strip detector is also shown. The bigger peak represents the position on the detector of the ^{236}U obtained with a spike sample; the nominal ratio is $^{236}U/^{238}U \sim 10^{-8}$. The lower ^{236}U peak is obtained with the KkU VERA in house U standard, see text. The arrow indicates that the normalized counts at that position are lower than 1×10^{-2} counts/nA.

3.2 ANU system

The ANU AMS system is based on a 15MV-tandem accelerator (Fifield et al., 1996). The high terminal voltage is required to apply certain techniques of isobar separation effectively, this makes the ANU tandem the best suited accelerator for the heavier isotopes e.g., ^{36}Cl and ^{53}Mn (Winkler, 2008). When the lower energy is necessary, for ^{236}U and ^{x}Pu isotopes, sections of the accelerator tube are shorted out, in order to optimize the ion optics for maximum transmission.

The pre-treated sample material (a few mg is pressed into a 1 mm diameter Al cathode and put in the ion source) itself is analyzed by two mass spectrometers which are coupled to the tandem accelerator. A schematic layout of the ANU 15 MV tandem facility is shown in Fig. 4. The caesium sputter ion source is a 32-sample MC-SNICS. This multi-cathode arrangement allows for measuring many samples without opening the source or employing a more complicated single cathode exchange mechanism. A total injection energy of 100 keV was used and $\sim$ 20 nA of $^{238}U^{16}O^{-}$ molecular ions are mass rigidity selected by the 90° double focusing Low Energy (LE) injection magnet (r = 0.83 m, B_{max} = 1.3 T, $ME/q^2 \simeq 56$ MeV·amu/e^2). This allows high resolution mass analysis for all stable isotopes in the periodic table. In contrast to the CIRCE system, there is no electrostatic analyzer, and hence the

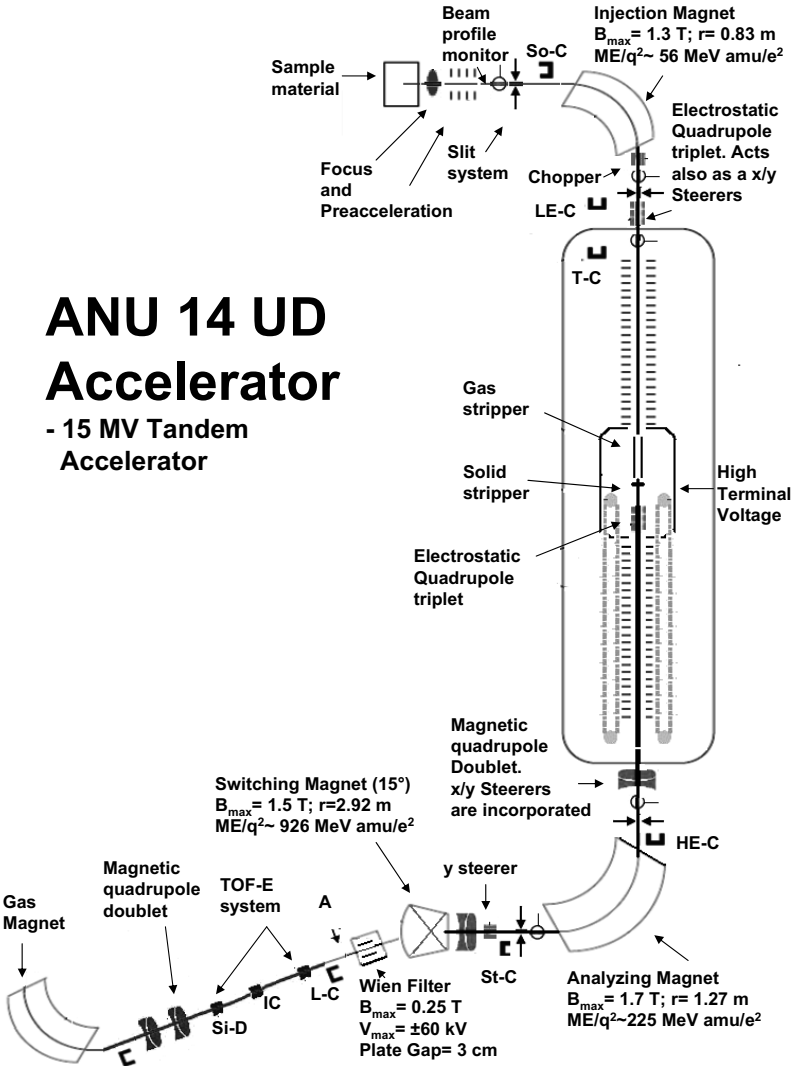


Fig. 4. Schematic lay out of the 15 MV ANU 14-UD (Units Doubled) Accelerator and the ^{236}U and ^{x}Pu isotopes detection line. The Switching Magnet, the Wien Filter, the start and the stop TOF-E detector, the Ionization Chamber, the magnetic quadrupole doublet and the Gas Magnet are shown in the line. C denotes the position-Faraday Cup, A denotes the Aperture of 1.5×4.0 mm^2 and the arrows indicate a slits system. The Accelerator is vertical up to the switching magnet that is indicate with a cross.

low-energy sputter tail is not removed prior to injection into the accelerator. For this reason, it is preferred to tune the system with $^{232}Th^{16}O^{-}$ rather than $^{238}U^{16}O^{-}$ (see next section).
A beam profile monitor (BPM) before the magnet and Faraday cups after the magnet (LE-Cup and Tank-Cup) are used to monitor the beam during the tuning. The injection beam line also features an electrostatic chopper that allows to reduce the beam intensity in cases where the beam currents are too high for injection into the tandem accelerator or counting rates that would be too high for the detector (e.g. ^{234}U). An electrostatic quadrupole and steerers are available to have the ions pass on an optimum trajectory for injection into the accelerator.
The terminal is charged by chains made of metal pellets which are isolated from each other by nylon links. The pellets supply a total charging current to the terminal of about 230 μA. The accelerator is contained inside a vessel filled with sulphur hexafluoride (SF_6) at a pressure of about 6 bar. The voltage is measured by a generating voltmeter. Regulation is achieved by employing a controlled corona discharge from ground to terminal. Both a gas stripper and a foil stripper are available at the terminal. At the terminal the ions lose electrons in the stripper, where O_2 is recirculated by two turbo-pumps; the working pressure is about 1 mTorr for $^{238}U^{5+}$ at 3.995 MV. Molecular ions are dissociated and the now atomic ions stripped to higher positive charge states (Litherland, 1980). The choice of charge state for heavy ions depends critically on a compromise between its stripping yield and the capability of the subsequent analyzing magnet to bend such ions. At $\sim$ 4 MV, ME/q^2 is $\sim$ 226 MeV·amu/e^2 for $^{238}U^{5+}$ and ME/q^2 is $\sim$ 293 MeV·amu/e^2 for $^{238}U^{4+}$; since the double focusing HE magnet reaches a maximal $ME/q^2\sim$ 225 MeV·amu/e^2, the 5+ represents the lowest charge state which can be bent by the HE magnet. Although the stripping yield to 4+ charge state is higher than 5+, it would be necessary to operate at lower terminal voltage in order to bend the ions. Since the transmission (due to the larger scattering angle) and the energy of the ions at this voltage is lower there is no gain to use the lower charge state.
The ions with positive charge states are accelerated a second time by the same potential. The High Energy (HE) magnet, efficiently removes molecular break-up products. The double focusing 90° HE bending magnet has r = 1.27 m, B_{max} = 1.7 T, $ME/q^2 \simeq$ 225 MeV·amu/e^2. A switching magnet (B_{max}= 1.5 T, r=2.92 m and $ME/q^2\simeq$ 926 MeV·amu/e^2 at the 15° exit) is positioned after the HE magnet. A Wien filter (B_{max}= 0.25 T, V_{max}= $\pm$ 60 kV with a Plate Gap= 3 cm) is employed to remove backgrounds.
Finally the selected ions are counted in a final detector. The control of the acquisition system is handled via Ethernet interfaces.

3.2.1 ANU actinides measurement procedures

In this paragraph a description of the various steps of the ^{236}U and ^{x}Pu isotope measurements will be given. The relative abundance of ^{238}U in environmental samples is many orders of magnitude (up to 13) larger than the ^{236}U. For this reason, while the number of events of ^{236}U are measured in the final detector, ^{238}U is measured as a current at the high energy side. Before performing measurements of samples, a tuning of the transport elements up to the final detector is required in order to maximize the ion optical transmission. The tuning is made by setting the parameters of the beam line to the detection of ^{232}Th. In order to have a good negative ion yield, molecular negative ions $^{232}Th^{16}O^{-}$ are extracted from the ion source. The negative molecular ions, $^{232}Th^{16}O^{-}$, are accelerated to injection energy of E_{inj}= 100 keV.
The injected ions are accelerated by the positive high voltage towards the gas stripper, where they lose electrons and gain high positive charge states. The positive ions are then accelerated a second time by the same potential in the high energy tube of the tandem. For $^{232}Th^{5+}$,

this results in an energy of E= 24.424 MeV with a terminal voltage of V= 4.098 MV. The stripping yield is the ratio between the $^{232}Th^{5+}$ beam current at the Faraday cup after the analyzing magnet (St-C) divided by 5 and the $^{232}Th^{16}O^{-}$ current measured at the entrance to the accelerator (T-C), and is about 3%. Molecular break-up products with mass over charge ratio M/q different from that of the wanted ion are removed by the analyzing magnet and switching magnet. The Wien filter is employed to remove backgrounds which have the same ME/q^2 as the ions of interest but different velocities in the actinides line. For heavy ion tuning, the object and image slits of the injection magnet are closed to $\pm$1 mm, the slits of the analyzing magnet are closed to $\pm$1.25 mm and an aperture of 1.5 $\times$ 4.0 mm^2 is used if high selectivity is required just after the Wien filter. For actual measurements, the object and image slits of the injection magnet are opened to $\pm$2 mm, the slits of the analyzing magnet are opened to $\pm$3 mm and the aperture is out.
For Uranium measurements, once the setup for the pilot beam $^{232}Th^{5+}$ is found, the fields of the injection magnet, the terminal voltage of the accelerator and the electric field of the Wien filter are scaled to $^{238}U^{5+}$ for a fine tuning and then to the other wanted masses. For $^{236}U/^{238}U$, the measurement procedure is composed of two loops of three steps. Each loop consists of integration of the $^{238}U^{5+}$ beam current for 10 s in the L-C, counting of $^{236}U^{5+}$ ions for 5 min in the TOF-E system and a final $^{238}U^{5+}$ integration. For ^{233}U (tracer), ^{234}U and ^{236}U, the measurement procedure is composed of two loops of four steps. The isotope sequence would usually start with the reference isotope ^{233}U followed by ^{234}U and ^{236}U, and finishing with ^{233}U. All of them are counted with the TOF-E system. The typical counting intervals were 1 minute for ^{233}U, 1 minute for ^{234}U and 5 minutes for ^{236}U.
For Plutonium measurements, once the setup for the pilot beam $^{232}Th^{5+}$ is found, since $^{238}U^{5+}$ may cause interference for $^{239}Pu^{5+}$, the fields of the injection magnet, the terminal voltage of the accelerator and the electric field of the Wien filter are scaled to the Pu wanted masses, ^{239}Pu, ^{240}Pu and ^{242}Pu (tracer). The measurement procedure is composed of two loops of four steps; the isotope sequence would usually start with the reference isotope ^{242}Pu followed by ^{240}Pu and ^{239}Pu, and finishing with ^{242}Pu. All of them are counted with a multiple electrode ionization chamber that is routinely used for measurements of ^{x}Pu isotopes. The typical counting intervals were 1 minute for ^{242}Pu, 5 minutes for ^{240}Pu and 3 minutes for ^{239}Pu.

3.2.2 Detection systems and ANU actinide results

Although most of the ^{238}U are suppressed at the injector side and by the analyzing magnet and Wien filter, a small fraction of this intense beam can interfere with the ^{236}U measurement even if the expected separation in the ion-optical filters is large, paragraph 3.1.2. For this reason the detection of the ^{236}U at ANU is made with a TOF-E detection. The configuration of the TOF detection system is as follows (Wilcken, 2006; Winkler, 2008); the start detector assembly is based on a MCP and a foil is placed at an angle of 45° to the beam. The MCP detects the backscattered electrons from a 0.6 μg/cm^2 thick diamond-like carbon (DLC) (Liechtenstein et al., 1999; 2002; 2004; 2006) foil that is used in the start detector to minimize scattering. The stop detector is a 200 mm^2 silicon surface barrier detector which also provides a total energy signal. The foils are mounted on a Cu mesh with a transparency of $\sim$ 75%. The MCP detector was operated with the anode at ground, the accelerating grid and the front face of the MCP at -1.8 kV, and the carbon foil at -2.8 kV. The presence of the foil, which is oriented at 45° to the beam, has two important consequences for the system. First, it causes scattering, which if through a large-enough angle can cause ions to miss the stop detector. This can be minimized by using the thinnest possible foil. Secondly, the 45° tilt introduces differences in path length

and therefore also in flight time due to the finite size of the beam at the start detector. The effect of the flight path variations on the resolution of the system is minimized by using an aperture that is 3.5 mm wide in the horizontal plane. This is attached on top of the grid-foil assembly.
For plutonium measurements no interfering ions exist; an ionization chamber is suitable for such a detection. The ANU configuration of the ionization chamber (Fifield et al., 1996; Wilcken, 2006; Wilcken et al., 2008) are the following; $\sim$ 50 torr of propane is used as the detector gas and the window is a 0.7 μm thick Mylar foil. Applied voltages are: cathode $\simeq$ -600 V, detector window $\simeq$ -300 V, first grid at ground, second grid at $\simeq$ +200 V and anode $\simeq$ +600 V. The energy of the $^{239}Pu^{5+}$ ions is $\sim$ 24.5 MeV. At this energy, the range of the plutonium ions in the ANU detector is $\sim$ 35 mm, which is roughly 18% of the length of the detector. The energy loss and straggling in the detector window are approximately 4.5 MeV and 450 keV, respectively. In addition, according to the manufacturer, a typical value for the surface roughness of the Mylar window is 38 nm, which is 5% of the thickness of the window and contributes an additional 140 keV of straggling. All of these result in an energy resolution of $\sim$ 4 %.
Regarding the abundance sensitivity results, the ANU is the best system in the word together with VERA laboratory (Steier et al., 2010). The ANU is able to obtain values of $^{236}U/^{238}U \simeq 10^{-13}$ (Wilcken et al., 2008), in samples including about 1 mg of U. This results is obtained with a time of flight of 2.3 m.
Preliminary results have been obtained with a 6 m flight path; the longer flight path confers a substantial improvement in the ability to separate ^{235}U and ^{236}U with little reduction in efficiency (Fifield, 2011).
The concentration sensitivity limit is of the order of $\sim$ 1 μg of uranium.
As regard the ^{239}Pu concentration sensitivity results, the uranium background at the ^{239}Pu settings is at the level of 100 ppb of the uranium concentration, i.e. 1 ng of uranium in the sample results in a background equivalent to 0.1 fg of ^{239}Pu (Fifield, 2008).

4. Summary and conclusion

The actinides detection technique described in this chapter can be applied in the assessment of contaminations from nuclear facility and used as sensitive fingerprints of programmed and accidental releases; a more general goal of this technique is to provide an ultrasensitive diagnostic tool for a variety of applications to the international community. Moreover the origin of actinides are discussed as well as the potential of actinides to serve as a tracer for geomorphologic processes.
The sensitivity of the different actinides measurements method and the peculiarity of the AMS technique with respect to AS and CMS techniques have been illustrated. Furthermore the principles and methodology of heavy-element AMS as applied to U and Pu isotopes, and the ways in which these have been implemented in various laboratories around the world, have been discussed. In particular the measurement procedures and the concentration and abundance sensitivity results of two systems, CIRCE and ANU, have been described in more details.
Those are two of the few systems in the world able to perform such measurements; the CIRCE is the only one in Italy.
The CIRCE system is at level of $\sim 10^{-12}$ $^{236}U/^{238}U$ isotopic ratio sensitivity which is still one order of magnitude higher then the ANU and VERA systems.

As future plan the CIRCE actinides group foresees to reach and exceed this sensitivity ratio goal with the new upgrade: the utilization of a TOF-E system with a thinner carbon foil and, if necessary, with a longer time of flight.
Regarding the Plutonium background results, the CIRCE is one of the best systems in the world; it is at the level of 1 ppb. This is to be compared with ANSTO where the uranium background is at the level of 10 ppm, and the ANU system where it is at the level of 100 ppb. The CIRCE laboratory has at present a ^{239}Pu sensitivity level less then 0.1 fg.

5. Acknowledgment

I kindly thank Prof. F. Terrasi, A. D'Onofrio, N. De Cesare, L. Gialanella, Dr. C. Sabbarese from SUN and Prof. L. K. Fifield, Dr. S. G. Tims from ANU and Dr. Y-J Guan from Guangxi University of Nanning and all the CIRCE actinides group who helped me to make this work possible.
Dr. P. Steier from VERA and Dr. D. Rogalla from Ruhr-Universität of Bochum and Dr. A. Di Leva from University of Naples and Dr. A. M. Esposito from SoGIN, for useful discussions and suggestions. This work was supported by SoGIN, Società Gestione Impianti Nucleari.

6. References

Beasley, T.M.; Kelley, J.M.; Maiti, T.C.; Bond, L.A. (1998). ^{237}Np/^{239}Pu Atom Ratios in Integrated Global Fallout: a Reassessment of the Production of ^{237}Np. *Journal of Environmental Radioactivity*, Vol. 38, pp 133-146

Berkovits, D.; Feldstein, H.; Ghelberg, S.; Hershkowitz, A.; Navon, E.; Paul, M. (2000). ^{236}U in uranium minerals and standards. *Nuclear Instruments and Methods in Physics Research B*, Vol. 172, pp 372-376

Betz, HD. (1972). Charge states and charge-changing cross sections of fast heavy ions penetrating through gaseous and solid media. *Reviews of Modern Physics*, Vol. 44, pp 465-539

Brown, T.A.; Marchetti, A.A.; Martinelli, R.E.; Cox, C.C.; Knezovich, J.P.; Hamilton, T.F. (2004). Actinide measurements by accelerator mass spectrometry at Lawrence Livermore National Laboratory. *Nuclear Instruments and Methods in Physics Research B*, Vol. 223Ű224, pp 788-795

Chiappini, R.; Pointurier, F.; Millies-Lacroix, J.C.; Lepitit, G.; Hemet, P. (1999). ^{240}Pu/^{239}Pu isotopic ratios and $^{239+240}$Pu total measurements in surface and deep waters around Mururoa and Fangataufa atolls compared with Rangiroa atoll (French Polynesia). *The Science of the Total Environment*, Vol. 237/238, pp 269-276

De Cesare, M. (2009). Accelerator Mass Spectrometry of actinides at CIRCE. *Phd Thesis*, Second University of Naples, Department of Environmental Sciense, Caserta (Italy)

De Cesare, M.; Gialanella, L.; Rogalla, D.; Petraglia, A.; Guan, Y.; De Cesare, N.; D'Onofrio, A.; Quinto, F.; Roca, V.; Sabbarese, C.; Terrasi, F. (2010). Actinides AMS at CIRCE in Caserta (Italy). *Nuclear Instruments and Methods in Physics Research B*, Vol. 268, pp 779-783

De Cesare, M.; Guan, Y.; Quinto, F.; Sabbarese, C.; De Cesare, N.; D'Onofrio, A.; Gialanella, L.; Petraglia, A.; Roca, V.; Terrasi, F. (2010). Optimization of ^{236}U AMS at CIRCE. *Radiocarbon*, Vol. 52, pp 286-294

De Cesare, M.; Fifield, L.K.; Sabbarese, C.; Tims, S. G.; De Cesare, N.; D'Onofrio, A.; D'Arco, A.; Esposito, A. M.; Petraglia, A.; Roca, V.; Terrasi, F. (2011), AMS12 conference

proceeding: Actinides AMS at CIRCE and ^{236}U and Pu measurements of structural and environmental samples from in and around a mothballed nuclear power plant.

Diamond, H.; Fields, P.R.; Stevens, C.S.; Studier, M.H.; Fried, S.M.; Inghram, M.G.; Hess, D.C.; Pyle, G.L.; Mech, J.F.; Manning W.M.; Ghiorso, A.; Thompson, S.G.; Higgins, G.H.; Seaborg G.T.; Browne, C.I.; Smith, H.L.; Spence, R. W. (1960). Heavy Isotope Abundances in Mike Thermonuclear Device. *Physical Review*, Vol. 119, 2000-2004

Fifield, L.K.; Cresswell, R.G.; Tada, M.L.D.; Ophel, T.R.; Day, J.P.; Clacher, A.P.; King, S.J.; Priest, N.D. (1996). Accelerator mass spectrometry of plutonium isotopes. *Nuclear Instruments and Methods in Physics Research B*, Vol. 117, pp 295-303

Fifield, L.K.; Clacher, A.P.; Morris, K.; King, S.J.; Cresswell, R.G.; Day, J.P.; Livens, F.R. (1997). Accelerator mass spectrometry of the planetary elements. *Nuclear Instruments and Methods in Physics Research B*, Vol. 123, pp 400-404

Fifield L.K. (2008). Accelerator mass spectrometry of the actinides. *Quaternary Geochronology*, Vol. 3, pp 276-290

Fifield, L.K.; Tims, S.G.; Stone, J.O.; Argento, D.C.; De Cesare, M. (2011), AMS12 conference proceeding: Ultra-sensitive measurements of ^{36}Cl and ^{236}U at the Australian National University

Guan, Y.G.; De Cesare, M.; Terrasi, F.; Quinto, F.; Sabbarese, C.; De Cesare, N.; D'Onofrio, A.; Wang, H. J. (2010). ^{236}U AMS measurement at CIRCE. *Chinese Physics C*, Vol. 34, pp 1729-1732

Hotchkis, M.A.C.; Child, D.; Fink, D.; Jacobsen, G.E.; Lee, P.J.; Mino, N.; Smith, A.M.; Tuniz, C. (2000). Measurement of ^{236}U in environmental media. *Nuclear Instruments and Methods in Physics Research B*, Vol. 172, pp 659-665

Hrneceka, E.; Steier, P.; Wallnerbet, A. (2005). Determination of plutonium in environmental samples by AMS and alpha spectrometry. *Applied Radiation and Isotopes*, Vol. 63, pp 633-638

Ketterer, M.E.; Hafer, K.M.; Link, C.L.; Royden, C.S.; Hartsock, W.J. (2003). Anthropogenic ^{236}U at rocky flats, Ashtabula river harbor, and Mersey estuary: three case studies by sector inductively coupled plasma mass spectrometry. *Journal of Environmental Radioactivity*, Vol. 67, pp 191-206

Ketterer, M. E. & Szechenyi, S.C. (2008). Determination of plutonium and other transuranic elements by inductively coupled plasma mass spectrometry: A historical perspective and new frontiers in the environmental sciences. *Spectrochimica Acta Part B*, Vol. 63, pp 719-737

Koide, M.; Bertine, K.K.; Chow, T.J.; Goldberget, E.D. (1985). The ^{240}Pu/^{239}Pu ratio, a potential geochronometer. *Earth and Planetary Science Letters*, Vol. 72, pp 1-8

Liechtenstein, V.Kh.; Ivkova, T.M.; Olshanski, E.D.; Baranov, A.M.; Repnow, R.; Hellborg, R.; Weller, R.A.; Wirth, H.L. (1999). Preparation and comparative testing of advanced diamond-like carbon foils for tandem accelerators and time-of-Fight spectrometers. *Nuclear Instruments and Methods in Physics Research A*, Vol. 438, pp 79-85

Liechtenstein, V.Kh.; Ivkova, T.M.; Olshanski, E.D.; Repnow, R.; Levin, J.; Hellborg, R.; Persson, P.; Schenkel, T. (2002). Advances in targetry with thin diamond-like carbon foils. *Nuclear Instruments and Methods in Physics Research A*, Vol. 480, pp 185-190

Liechtenstein, V.Kh.; Ivkova, T.M.; Olshanski, E.D.; Golser, R.; Kutschera, W.; Steier, P.; Vockenhuber, C.; Repnow, R.; von Hahn, R.; Friedrich, M.; Kreissig, U. (2004). Recent investigations and applications of thin diamond-like carbon (DLC) foils, *Nuclear Instruments and Methods in Physics Research A*, Vol. 521, pp 197-202

Liechtenstein, V.Kh.; Ivkova, T.M.; Olshanski, E.D.; Repnow, R.; Steier, P.; Kutschera, W.; Wallner, A.; von Hahn, R. (2006). Preparation and investigation of ultra-thin diamond-like carbon (DLC) foils reinforced with collodion. *Nuclear Instruments and Methods in Physics Research A*, Vol. 561, pp 120-123

Litherland A.E. (1980). Ultrasensitive Mass Spectrometry with Accelerators. *Annual Review of Nuclear and Particle Science*, Vol. 30, 437-473

Maier-Komor, P.; Bergmaier, A.; Dollinger, G.; Frey, C.M.; Krner, H.J. (1997). Improvement of the preparation procedure of carbon stripper foils from the laser ablation-deposition process. *Nuclear Instruments and Methods in Physics Research A*, Vol. 397, pp 131-136

Maier-Komor, P.; Dollinger, G.; Krner, H.J. (1999). Reproducibility and simplifcation of the preparation procedure for carbon stripper foils by laser plasma ablation deposition. *Nuclear Instruments and Methods in Physics Research A*, Vol. 438, pp 73-78

Makino, Kyoko & Berz, Martin (1999). COSY INFINITY version 8. *Nuclear Instruments and Methods in Physics Research A*, Vol. 427, pp 338-343

O'Donnell, R.G.; Mitchell, P.I.; Priest, N.D.; Strange, L.; Fox, A.; Henshaw, D.L.; Long, S.C.; (1997). Variations in the concentration of plutonium, strontium-90 and total alpha-emitters in human teeth collected within the British Isles. *Science of the Total Environment*, Vol. 201, pp 235-243

Ordine, A.; Boiano, A.; Vardaci, E.; Zaghi, A.; Brondi, A. (1998). A new fast trigger and readout bus system. *Nuclear Science*, Vol. 45, pp 873-879

Perelygin, V.P. & Chuburkov, Yu.T. (1997). Man-made plutoniumŮ possible serious hazard for livingspecies. *Radiation Measurements*, Vol 28, pp 385-392

Quinto, F. (2007). Assessing radioactive contamination in the environment around the Garigliano Nuclear Power Plant. *Phd Thesis*, Second University of Naples, Department of Environmental Sciense, Caserta (Italy)

Quinto, F.; Steier, P.; Wallner, G.; Wallner, A.; Srncik, M.; Bichler, M.; Kutschera, W.; Terrasi, F.; Petraglia, A.; Sabbarese, C. (2009). The first use of ^{236}U in the general environment and near a shutdown nuclear power plant. *Applied Radiation and Isotopes*, Vol. 67, pp 1775-1780

Richter, S.; Alonso, A.; Bolle, W.D.; Wellum, R.; Taylor, P.D.P. (1999). Isotopic ŚŚfingerprintsŠŠ for natural uranium ore samples. *International Journal of Mass Spectrometry*, Vol. 193, pp 9-14

Roca, V.; Napolitano, M.; Speranza, P.R.; Gialanella, G. (1989). Analysis of radioactivity levels in soils and crops from the Campania region (South Italy) after the Chernobyl accident. *Journal of Environmental Radioactivity*, Vol. 9, pp 117-129

Sakaguchi, A.; Kawai, K.; Steier, P.; Quinto, F.; Mino, K.; Tomita, J.; Hoshi, M.; Whitehead, N.; Yamamoto, M. (2009). First results on ^{236}U levels in global fallout. *Science of the Total Environment*, Vol. 407, pp 4238-4242

Sanchez, A.M.; Tome, F.V.; Bejarano, J.D.; Vargas, M.J. (1992).A rapid method for determination of the isotopic composition of uranium samples by alpha spectrometry. *Nuclear Instruments and Methids in Physics Research A*, Vol. 313, pp 219-226

Steier, P.; Golser, R.; Kutschera, W.; Liechtenstien, V.; Priller, A.; Valenta, A.; Vockenhuber, C. (2002). Heavy ion AMS with a "small" accelerator. *Nuclear Instruments and Methods in Physics Research B*, Vol. 188, pp 283-287

Steier, P.; Bichler, M.; Fifield, L.K.; Golser, R.; Kutschera, W.; Priller, A.; Quinto, F.; Richter, S.; Srncik, M.; Terrasi, F.; Wacker, L.; Wallner, A.; Wallner, G.; Wilcken, K.M.; Wild,

E.M. (2008). Natural and anthropogenic ^{236}U in environmental samples. *Nuclear Instruments and Methods in Physics Research B*, Vol. 266, pp 2246-2250

Steier, P.; Dellinger, F.; Forstner, O.; Golser, R.; Knie, K.; Kutschera, W.; Priller, A.; Quinto, F.; Srncik, M.; Terrasi, F.; Vockenhuber, C.; Wallner, A.; Wallner, G.; Wild, E.M. (2010). Analysis and application of heavy isotopes in the environment. *Nuclear Instruments and Methods in Physics Research B*, Vol. 268, pp 1045-1049

Terrasi, F.; Rogalla, D.; De Cesare, N.; DŠOnofrio, A.; Lubritto, C.; Marzaioli, F.; Passariello, I.; Rubino, M.; Sabbarese, C.; Casa, G.; Palmieri, A.; Gialanella L.; Imbriani, G.; Roca, V.; Romano, M.; Sundquist, M.; Loger, R. (2007). A new AMS facility in Caserta/Italy. *Nuclear Instruments and Methods in Physics Research B*, Vol. 259, pp 14-17

Vockenhuber, C.; Golser, R.; Kutschera, W.; Priller, A.; Steier, P.; Winkler, S.; Liechtenstein, V. (2002) Accelerator mass spectrometry of heaviest long-lived radionuclides with a 3-MV tandem accelerator, *Pramana-Journal of Physiscs*, Vol. 59, pp 1041-1051

Wacker, L.; Chamizo, E.; Fifield, L.K.; Stocker, M.; Suter, M.; Synal, H.A. (2005). Measurement of actinides on a compact AMS system working at 300 kV. *Nuclear Instruments and Methods in Physics Research B*, Vol. 240, pp 452-457

Wallner, C.; Faestermann, T.; Gerstmann, U.; Hillebrandt, W.; Knie, K.; Korschinek, G.; Lierse, C.; Pomar, C.; Rugel, G. (2000). Development of a very sensitive AMS method for the detection of supernovaproduced longliving actinide nuclei in terrestrial archives. *Nuclear Instruments and Methods in Physics Research B*, Vol. 172, pp 333-337

Wilcken, K.M. (2006). Accelerator Mass Spectrometry of natural ^{236}U and ^{239}Pu with enphasis on nucleogenic isotope production. *Phd Thesis*, Australian National University, Department of Nuclear Physics, Canberra (Australia)

Wilcken, K.M.; Barrows, T.T.; Fifield, L.K.; Tims, S.G.; Steier, P. (2007). AMS of natural ^{236}U and ^{239}Pu produced in uranium ores. *Nuclear Instruments and Methods in Physics Research B*, Vol. 259, pp 727-732

Wilcken, K.M.; Fifield, L.K.; Barrows, T.T.; Tims, S.G.; Gladkis L.G. (2008). Nucleogenic ^{36}Cl, ^{236}U and ^{239}Pu in uranium ores. *Nuclear Instruments and Methods in Physics Research B*, Vol. 266, pp 3614-3624

Winkler, S.R. (2008). Acclerator Mass Spectrometry of heavy radionuclides with special focus on ^{182}HF. *Phd Thesis*, Australian National University, Department of Nuclear Physics, Canberra (Australia)

Wyse, E.J.; Lee, S.H.; Rosa, J.L.; Povinec, P.; Mora, S.J.D. (2001). ICPsector field mass spectrometry analysis of plutonium isotopes: recognizing and resolving potential interferences. *Journal of Analytical Atomic spectrometry*, Vol. 16, pp 1107-1111

Zhao, X.L.; Nadeau, M.J.; Kilius, L.R.; Litherland, A.E. (1994). The first detection of naturally-occurring ^{236}U with accelerator mass spectrometry. *Nuclear Instruments and Methods in Physics Research B*, Vol. 92, pp 249-253

Zhao, X.-L., Kilius, L.R.; Litherland, A.E.; Beasley, T. (1997). AMS measurement of environmental U-236. Preliminary results and perspectives. *Nuclear Instruments and Methods in Physics Research B*, Vol. 126, pp 297-300

Part 2

Reliability and Failure Mechanisms

10

Evaluation of Dynamic J-R Curve for Leak Before Break Design of Nuclear Reactor Coolant Piping System

Kuk-cheol Kim, Hee-kyung Kwon, Jae-seok Park and Un-hak Seong
Doosan Heavy Industries & Construction Co. Ltd.
Korea

1. Introduction

Because safety is of paramount importance in the nuclear industry, numerous efforts have been made to guarantee structural integrity against sudden accidents. In the past, design against a Double Ended Guillotine Break (DEGB) was accomplished through the construction of massive pipe whip restraints and jet impingement shields to minimize the secondary damage to other structures in close proximity to ruptured piping. However, through long-term operating experience, the commercial nuclear industry has recognized that, for most damaged piping, fluid leakage from through-wall cracks occurs prior to a DEGB accident. Hence, if the leakage can be detected reliably at an early stage of fracture, a DEGB accident can be prevented by shutting down the reactor prior to the DEGB. Leak-Before-Break (LBB) design is based on this concept. For a piping system where LBB design is applied, a leak detection monitoring system must be installed to detect crack initiation while construction of massive pipe whip restraints and jet impingement shields become unnecessary. Thus, LBB design focuses on the ability to detect cracks for structural integrity while DEGB design focuses on preventing secondary damage. Since the mid-1980s, the LBB design concept has been widely applied on nuclear high energy piping systems. In Korea, the LBB design concept based on U.S. nuclear regulatory commission (USNRC) standard review plan 3.6.3 and NUREG-1061 has been applied to reactor coolant piping systems ever since the Yong-Gwang units 3 & 4 nuclear power plants were approved in 1994 (J.B.Lee & Choi, 1999).

The LBB design applied to nuclear piping systems is based on the premise that a piping break accident can be prevented by detecting leakage from a through-wall crack by leak detection instrumentation prior to a DEGB accident. To meet LBB design criteria, the nuclear piping material must have excellent fracture toughness characteristics so that a sudden break will not occur even if the piping has a large through-wall crack that corresponds to a detectable leakage rate. For LBB design, material properties for stress – strain curves and J-R curves as a function of resistance to stable crack extension at service temperatures are needed. The stress – strain curve is for use in the determination of detectable leakage crack length and the elastic-plastic finite element analysis of the piping with a through-wall crack. The J-R curve is for use in the crack stability evaluation of piping under normal operating loads and safe shutdown earthquake loads. In the Korean standard nuclear power plant, shown in Fig. 1, carbon steel with stainless steel cladding is used for the hot leg pipe and the

cold leg pipe of the reactor coolant piping system. For carbon steel, it is reported that fracture toughness is dependent on loading speed due to dynamic strain aging (J.W.Kim & I.S.Kim, 1997). In addition to static J-R curve testing, the dynamic J-R curve, which is a part of facture toughness data, is also required to verify satisfaction of LBB when applying seismic loading for carbon steel nuclear piping. However, until now it has been difficult to obtain a reliable dynamic J-R curve for ferritic steel due to the fast loading condition. In this paper, the measurement method for obtaining a reliable dynamic J-R curve for integrity analysis of nuclear piping systems is proposed and discussed.

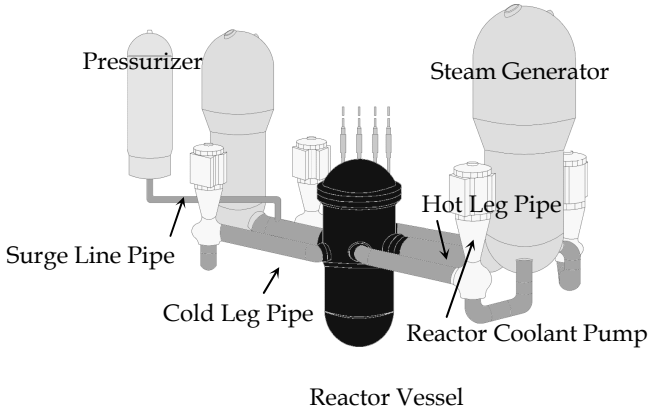


Fig. 1. Reactor coolant piping system

2. Dynamic J-R curve using DCPD and normalization methods

A dynamic J-R curve can be obtained by two different test methods; direct current potential drop (DCPD) (Joyce, 1996) and the Normalization method (Landes et al., 1991; ASTM, 2001). With DCPD on ferritic steel, a pulse drop phenomenon of output voltage occurs due to its ferromagnetic characteristics, making it difficult to determine a reliable J-R curve. On the other hand, the Normalization method, which was recently designated by the American Society for Testing and Materials (ASTM) code, has its strong point in that the J-R curve can be obtained by load - displacement curve without additional crack length measurement instrumentation such as needed by DCPD. In Korea, dynamic J-R curves have been obtained for piping materials in several nuclear power plants, and a database has been developed for dynamic J-R curves on each material based on these test results. According to the ASTM code at the time, the dynamic J-R curves were obtained by DCPD, but more recently, they are obtained by the Normalization method for newly constructed power plant projects. To utilize previous dynamic J-R curve data obtained by DCPD for piping material, the effect of test methods was investigated.

2.1 Experimental procedure

To compare the dynamic J-R curves between the DCPD and normalization methods, dynamic J-R curve testing was performed for base and weld metals of reactor coolant piping systems. Test specimens were 1 inch compact tension specimens. A test speed of 1,000

mm/min for dynamic J-R testing was determined on the basis of the natural frequency method proposed at Battelle (Scott et al., 2002) according to Eq. (1)

$$V_{LL} = 4 \times \text{natural frequency (mode 1)} \times D_i \quad (1)$$

where D_i is the load line displacement at crack initiation of the static J-R curve testing. This test speed also satisfies the criterion of ASTM E1820 A14 (Nakamura et al., 1986; ASTM, 2009) in which test time t_Q should be longer than minimum test time t_w

$$t_w = \frac{2\pi}{\sqrt{k_s / M_{eff}}} \quad (2)$$

where k_s is specimen load line stiffness in N/m, M_{eff} is effective mass of the specimen, taken here to be half of the specimen mass in kg.

Table 1 represents tested materials for each pipe and number of tests. Each hot leg is a 42 inch inner diameter pipe of SA508 Cl.1a material with a 3-½ inch nominal wall thickness. The cold leg is a 30 inch inner diameter pipe of SA508 Cl.1a material with a 3 inch nominal wall thickness. The elbow is SA516 Gr.70. The straight pipe and elbow are welded by submerged arc welding (SAW) and shielded metal arc welding (SMAW). Table 2 shows the chemical composition of the tested material and weld deposit. The comparison between DCPD and the Normalization method is summarized in Table 3. For DCPD, potential drop instrumentation was used for crack length measurement during the experiment but for the Normalization method, J-R curve was estimated only by the load – displacement curve without any crack length measurement device during the test. Therefore, in this study, dynamic J-R curve testing was performed using DCPD and analyzed by both DCPD and Normalization methods for each specimen with the test results compared between the two methods. Comparison tests were performed on two power plants, Shin-Kori units 3 & 4 and Shin-Wolsung units 1 & 2. For Shin-Kori, physical crack extension length did not exceed the lesser of 4mm or 15% of the initial uncracked ligament in accordance with normalization method. For Shin-Wolsung, tests were performed until full coverage of crack opening displacement (COD) gage, 10mm in accordance with previous DCPD method as performed at our test laboratory. Test temperature was 316°C; same as the operating temperature of the piping system. Additionally, in the case of Shin-Wolsung, tests were performed at hot standby temperature, 177°C. Table 1 shows the number of test specimens and test temperatures for dynamic J-R curve testing.

Item			Material	Dynamic J-R curve testing		
				Shin-Kori units 3 & 4	Shin-Wolsung units 1 & 2	
				316°C	177°C	316°C
Base metal	Main loop piping	Hot leg	SA508 Cl. 1a	1	1	1
		Cold leg	SA508 Cl. 1a	1	1	1
		Elbow	SA516 Gr. 70	1	1	1
Weld metal	Main loop piping segments		SMAW	1	1	1
			SAW	1	1	1
Total				15		

Table 1. Fracture toughness test conditions of the coolant piping

Pipe	C	Si	Mn	Cu	Mo	V	Ni
Hot leg & cold leg	<0.30	0.15~0.40	0.70~1.35	<0.2	<0.1	<0.03	<0.4
Elbow	<0.30	0.15~0.40	0.85~1.20	<0.4	<0.12	<0.03	<0.4
SMAW	<0.17	<0.75	<1.60	-	<0.30	<0.08	<0.30
SAW	<0.15	<0.80	1.25~2.10	<0.06	0.40~0.65	<0.03	<0.20

Table 2. Chemical composition of base materials and weld joints for reactor coolant piping (%, wt)

Item	DCPD method	Normalization method
Crack length measurement device	DCPD	N/A
Crack length estimation method during the test	By variation of output voltage when constant current is applied to specimen	By only load-displacement record
Effective crack extension length	Not more than 4mm or 15% of the initial uncracked ligament, whichever is less as physical crack extension length	Not more than 25% of the initial uncracked ligament as effective data region at data analysis

Table 3. Comparison of dynamic J-R curve testing method

2.1.1 DCPD method

The schematic diagram of the dynamic J-R curve testing apparatus is shown in Fig. 2. The specimen was isolated from the load frame by inserting Bakelite plates between the connecting rods, and constant current was applied to the specimen using a power supply in order to measure crack growth length during the test. A sufficiently high current of 100 amperes was used to minimize error due to ferromagnetic phenomenon. (Landow & Marschall, 1991; B.S.Lee et al., 1999) Current input wires were mechanically fastened to both sides of the specimen with screws at points A and B in Fig. 3, and voltage measurement wires, 0.7mm in diameter were spot welded at the points C and D. Using high-speed data acquisition, the variation of load, crack opening displacement (COD) value and output voltage were acquired digitally during the test. Prior to the dynamic J-R curve testing at high temperature, to compensate for the thermal effect, the reference voltage was measured from the specimen with current off at the test temperature. Voltage measurement was normalized by subtracting the reference voltage from measured voltage during the dynamic J-R tests. The variation of crack length was calculated based on Johnson's equation, Eq. (3) (Johnson, 1965).

$$\frac{a}{W}=\frac{2}{\pi}\cos^{-1}\left[\frac{\cosh\left(\pi y/2W\right)}{\cosh\left[\left(U/U_0\right)\cosh^{-1}\left[\cosh\left(\pi y/2W\right)/\cos\left(\pi a_0/2W\right)\right]\right]}\right] \tag{3}$$

where U_0 and a_0 are initial output voltage and initial crack length, respectively. According to the ASTM code (ASTM, 2009), as shown in Fig. 4(a), crack initiation point is determined as the intersection point of the measured DCPD curve and the 5% offset line based on a linear best-fit line of the data over the range from 0.1~0.5 P_{max}. However, as shown in Fig. 4(b), in the case of the tested ferritic steel, pulse drop phenomenon in the early loading stage of testing occurs due to the sudden reorientation of ferromagnetic domain nearby the crack tip (Hackett et al., 1986).

This pulse drop phenomenon makes it difficult to determine the crack initiation point. To resolve this problem, a backtracking technique proposed by Oh (Oh et al., 2002) was selected.

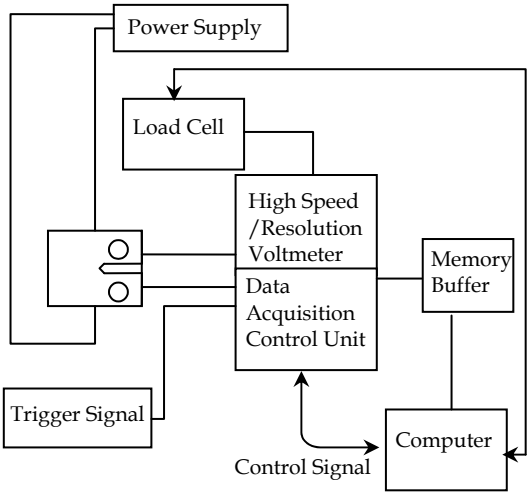


Fig. 2. Data acquisition system for dynamic J-R curve testing

In the backtracking technique, the crack initiation point is estimated by using final crack length measured in the fractured specimen. The backtracking technique is as follow; First, prior to crack initiation, it is assumed that crack extension length is in accordance with the standard blunting relation of $\Delta a=J/(2\sigma_Y)$, namely, a_0 in Eq. (3) is substituted for $a_0+J_B/(2\sigma_Y)$ where $J_B=J$ at crack initiation. Next, with changing U_0, the variation of crack length for each loading point can be obtained. Through this iterative process, U_0 is obtained such that the calculated final crack length is in agreement with the measured final crack length. Finally, the J-R curve is calculated using U_0.

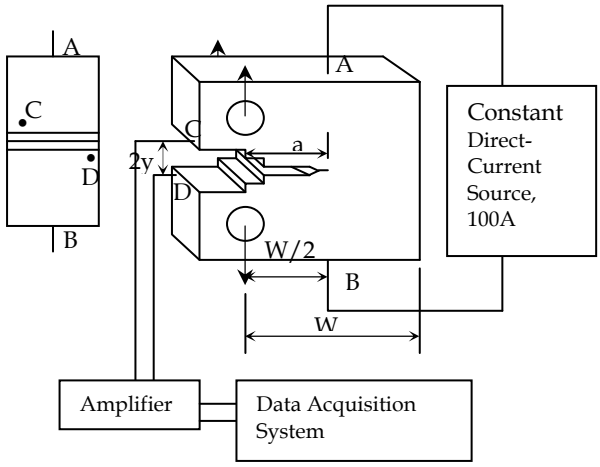
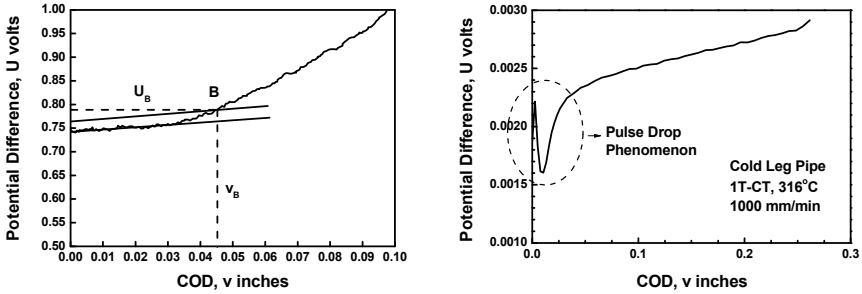


Fig. 3. Specimen geometry for dynamic J-R curve testing



(a) HY130 steel (Joyce, 1996) (b) SA508 Gr.1a steel

Fig. 4. Potential rise versus crack opening displacement

2.1.2 Normalization method

In the Normalization method (ASTM, 2009), dynamic J-R curve can be estimated using load - displacement data pairs. At first, load - displacement data is normalized by considering specimen size and crack length per Eqs. (4) and (5)

$$P_{Ni} = \frac{P_i}{WB\left[\frac{W - a_{bi}}{W}\right]^{\eta_{pl}}} \tag{4}$$

$$v'_{pli} = \frac{v_{pli}}{W} = \frac{v_i - P_i C_i}{W} \tag{5}$$

where $a_{bi}=a_0+J_i/(2\sigma_Y)$, P_{Ni} is normalized load, P_i is load, W is specimen width, B is specimen thickness, η_{pl} is plastic η factor, v'_{pli} is normalized displacement, v is load line displacement, and C_i is compliance. Using final crack length measured at the broken specimen surface, final normalized load displacement pair can be obtained from Eqs (4) and (5). Fitting coefficients a, b, c, d are obtained by curve fitting with Eq. (6) for effective data pair (P_{Ni}, v'_{pli}) including final normalized load displacement pair designated in ASTM E1820 A15.

$$P_N = \frac{a + bv'_{pl} + cv'^{2}_{pl}}{d + v'_{pl}} \tag{6}$$

The crack length a_i coinciding with P_{Ni} in Eq.(4) and with P_N in Eq.(6) is calculated for each v'_{pli} by checking with slightly increasing crack lengths from initial crack length a_0. Finally, using load, load line displacement and the calculated crack length data, the J-R curve can then be calculated.

2.2 Test results and discussion

Figure 5 shows the comparison of dynamic J-R curve between DCPD and the Normalization method for Shin-Kori units 3 & 4 when testing in accordance with crack extension criteria of the Normalization method. The dynamic J-R value at given crack extension length is well within the deviation range of ±5%. In the case of hot leg pipe and SAW, the dynamic J-R data using DCPD method tend to be 10% higher at the crack initiation point compared to that using the Normalization method. However, in the case of other materials, the dynamic

J-R curves are coincident with each other. Figure 6 shows the comparison of dynamic J-R curve between DCPD method and normalization method for Shin-Wolsung when testing until load line displacement of 10mm is reached. Note: hereafter short crack extension means the crack extension length is not more than 4mm or 15% of the initial uncracked ligament, whichever is less, and long crack extension means the crack extension length is over 4mm and 15% of the initial uncracked ligament. Except cold leg pipe material at 177°C and 316°C and elbow material at 316°C with long crack extension, the dynamic J-R curve is coincident for different test methods. Therefore we know that for short crack extension, the dynamic J-R curve is coincident for different test methods, but for long crackextension, the J-R curve using DCPD is estimated to 10~30% higher than that using normalization method.

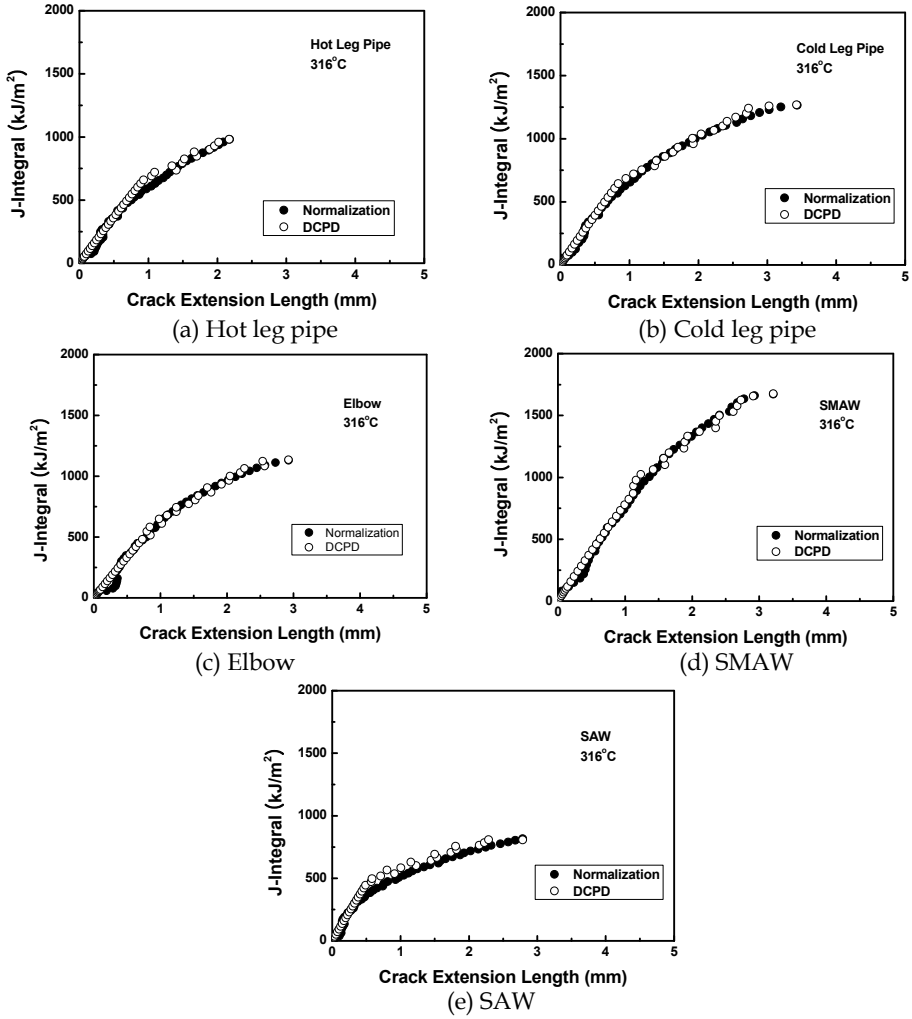


Fig. 5. Comparison of dynamic J-R curve between DCPD and normalization method when testing in accordance with crack extension criteria of normalization method

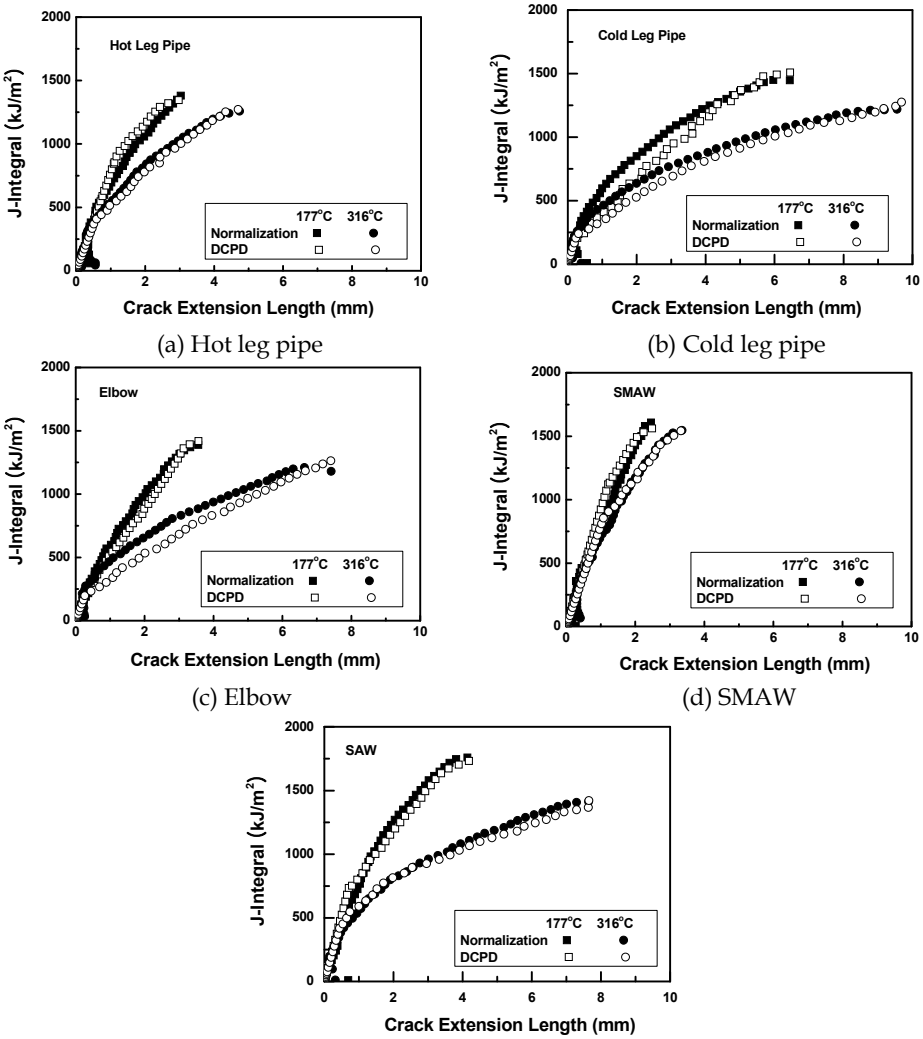


Fig. 6. Comparison of dynamic J-R curve between DCPD and normalization method when testing until load line displacement of 10mm

When applying DCPD, the pulse drop phenomenon on the displacement versus DCPD output voltage relationship makes it difficult to determine an accurate dynamic J-R curve. The output voltage increases slightly, decreases sharply and then recovers in a early loading stage for this ferritic steel as shown in Fig.4(b) by superimposition of the induced voltage due to sudden reorientation of ferromagnetic domain nearby the crack tip. Since Johnson's equation, Eq.(3), considers only the variation of output voltage with specimen geometry including crack length, some errors for estimation of crack length can occur in this case where output voltage includes the induced voltage. Despite this problem with DCPD, for short crack extension, dynamic J-R

curve using DCPD is similar to that using normalization. On the other hand, at long crack extension a difference in dynamic J-R curve between two test methods appears. However, in this case, normalization method is also not effective since a crack extension criterion is violated according to ASTM code. The difficulty of obtaining reliable J-R curve for long crack extension can be explained as follow; For normalization method, the normalization is based on the principle of load separation following Eq.(7) (Sharobeam & Landes, 1991; Landes et al., 1991)

$$P_N = \frac{P}{G(a/W)} = H\left(v_{pl}/W\right) \tag{7}$$

where P_N is normalized load, P is load, a is crack length, v_{pl} is plastic displacement and W is specimen width. In Eq.(7), if crack length is fixed, normalized load P_N value is easily calculated. However, to obtain J-R curve, normalized load in accordance with normalized function of Eq.(3) should be calculated based on actual crack length variation instead of fixed initial crack length. When load - displacement curve is normalized as fixed specimen geometry and crack length, the normalized curve is described by the open symbols in Fig. 7(a). According to ASTM code, to obtain the normalized load - displacement curve considering the variation of actual crack length, final data pairs estimated from measured final crack length and effective data pairs prior to crack initiation point in accordance with the method designated in ASTM E1820 code were selected. By performing a best fit for selected data pairs using Eq.(6), the normalized load - displacement curve can be estimated reflecting crack length variation. Crack extension length is estimated from the difference of two normalized curves as shown in Fig. 7(b), so for estimation of dynamic J-R curve, it is important to estimate reliable normalized load - displacement curve. In normalization method, normalized load - displacement curve at crack propagation region is estimated by interpolation using Eq.(6). In the case of small crack extension, this interpolation is reasonable because the region to be interpolated is narrow but in the case of long crack extension, interpolation errors can occur. If the position of normalized data pairs at the middle point between crack initiation point and final point is incorrectly estimated, the final estimated J-R curve is also in error from the actual J-R curve. It is therefore important to estimate the middle point exactly between crack initiation point and final point in the long crack extension case. If the position of middle point can be measured exactly through experiment, the reliable J-R curve will be able to evaluate for long crack extension beyond the crack extension length designated at ASTM code.

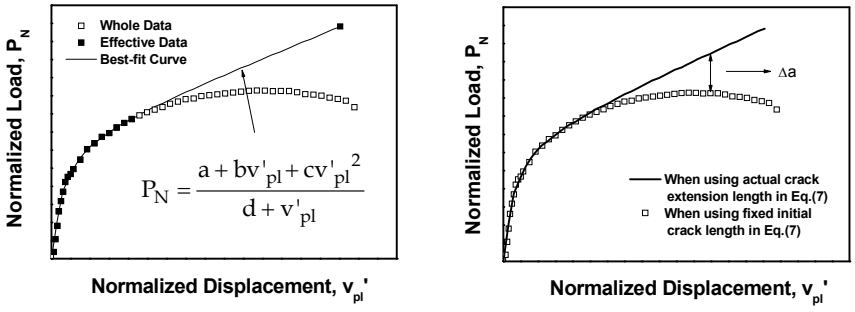


Fig. 7. J-R curve estimation concept on normalization method (a) Optimal best-fit for effective data pair (b) Crack extension length estimation from normalized curve

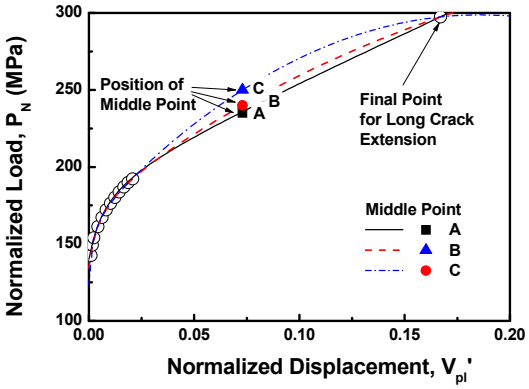


Fig. 8. The illustration diagram for the variation of the normalized load - displacement curve with the position of middle point

3. Proposal of modified normalization method for measurement of dynamic J-R curve with long crack extension

3.1 Importance of J-R curve for long crack extension

In LBB analysis, as an integrity analysis method against instability fracture of cracked piping, J-integral, tearing modulus (J/T) method (Ernst et al., 1979, 1981) and the limit load method were used. While the limit load method is appropriate for the analysis of stainless steel piping, J/T method is appropriate for the analysis of both carbon and stainless steel piping. For ductile material, final instability rupture occurs after stable crack extension with increasing load value. This instability point where piping rupture occurs can be determined using J/T method based on J-integral parameter. The stable growth criterion is

$$T_{app}\left(=\frac{E}{\sigma_f^2}\frac{dJ}{da}\right) < T_{mat}\left(=\frac{E}{\sigma_f^2}\frac{dJ_R}{da}\right) \tag{8}$$

where T_{app} is applied tearing modulus, T_{mat} is material tearing modulus, E is elastic modulus, σ_f is effective yield strength as defined by the average value of tensile strength and yield strength, J is J-integral value to be calculated from finite element analysis for the cracked piping and J_R is J-integral value to be obtained from J-R curve testing.

As shown in Fig. 9, tearing instability point is determined from intersection point of two J/T curves. From this instability point, critical load P_{max} value is determined, and the safety factor is defined as the ratio between the critical load P_{max} and applied load P. If the safety factor is larger than 1, the structural integrity can be verified by Eq. (8). For reliable stability analysis, T_{mat} curve should be evaluated experimentally to determine the intersection point between T_{app} curve and T_{mat} curve. However, when testing using normalization method for the dynamic J-R curve, T_{mat} curve can not sufficiently be measured due to the restriction of crack extension length. In this case, T_{mat} curve corresponding to long crack extension should be estimated from the limited T_{mat} curve with short crack extension by extrapolation as shown in Fig. 9. As an analytical approach of extrapolation, Wallen, K (Wallen, 2009) suggested additional two applicable best-fit methods from limited T_{mat} curve in addition to

conventional fitting method for tearing modulus curve. However, analytical approach has uncertainty basically by fitting. In this paper, to evaluate reliable T_{mat} curve at long crack extension region experimentally, we have researched the method for measurement of dynamic J-R curve with crack extension as long as possible.

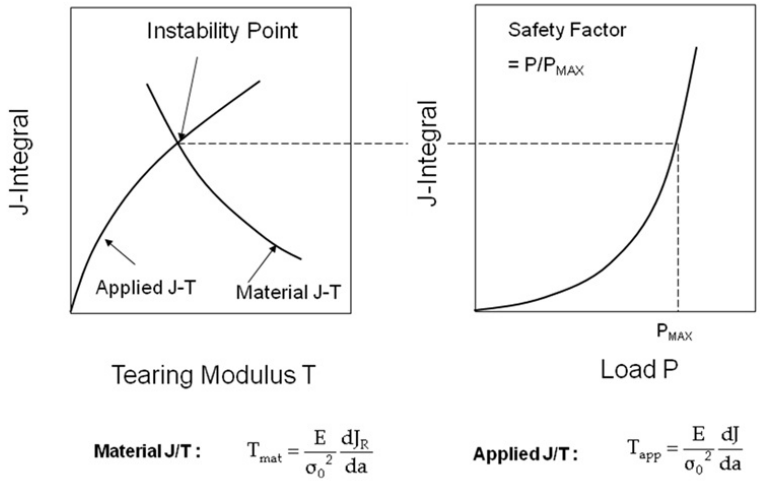


Fig. 9. Graphical illustration of J/T method

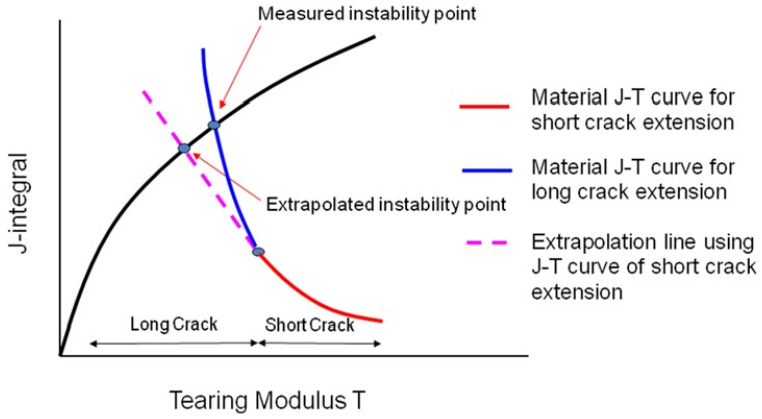


Fig. 10. The illustration diagram for estimation of crack instability point for J/T method

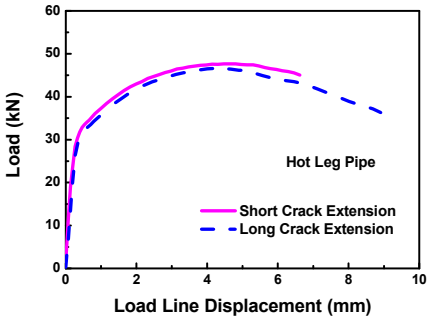
3.2 Dynamic J-R curve testing for long crack extension

To obtain the effective J-R curve under the condition of long crack extension, two specimens were used where one is for short crack extension and the other is for long crack extension. By using two test data, the dynamic J-R curve was evaluated over the crack extension length range according to ASTM code. Table 1 shows test matrix for reactor coolant piping base metal for Shin-Wolsung.

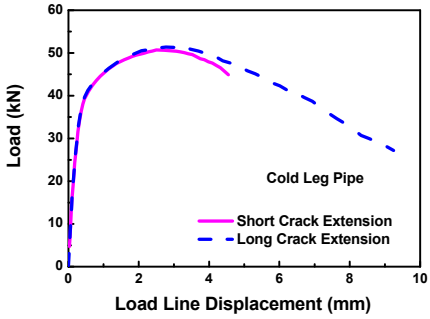
Item		Material	Pipe size (Inner Dia.)	Number of test	
				Short crack extension	Long crack extension
Main Loop Piping	Hot Leg	SA508 Gr. 1a	42 in.	1	1
	Cold Leg	SA508 Gr. 1a	30 in.	1	1
	Elbow	SA516 Gr. 70		1	1

Table 4. Dynamic J-R test conditions for short and long crack extension conditions

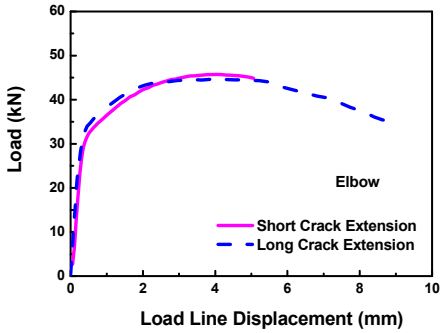
The load - displacement curve for each piping material is shown in Fig 11. In the dynamic J-R curves obtained by normalization method, for hot leg pipe and elbow materials, dynamic J-R curves were similar regardless of crack extension length; whereas for cold leg piping material, J-R curve for short crack extension length was lower than that for long crack extension length as shown in Fig.12. To analyze the reason for the difference between short and long crack extension for cold leg pipe, normalized load-displacement curve is described in Fig. 13. Normalized load-displacement curve, P_N - v'_{pl} curve shows different shape between two tests with different crack extension length. In general, normalized load – displacement curve should maintain a constant shape regardless of crack extension size. Therefore, optimal normalized P_N - v'_{pl} curve should be calculated by considering both P_{Ni} - v'_{pli} data pair for short and long crack extension.



(a) Hot leg pipe

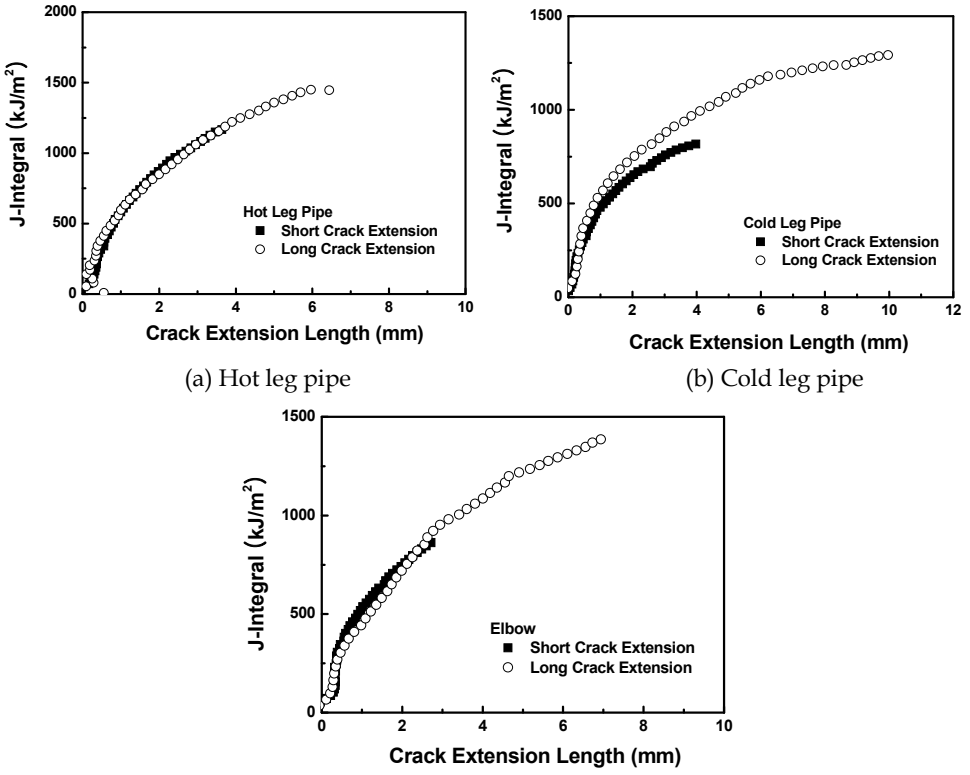


b) Cold leg pipe



(c) Elbow

Fig. 11. The load versus load line displacement curves for each material



(a) Hot leg pipe (b) Cold leg pipe

(c) Elbow

Fig. 12. The comparison of dynamic J-R curve by normalization method between the tests for short and long crack extension

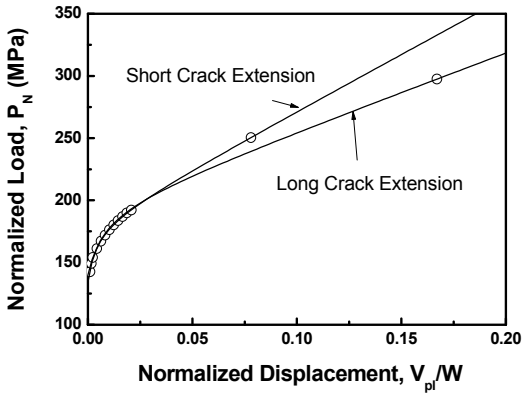


Fig. 13. Normalized load, displacement data pair and its each fitting curve for short and long crack extension of cold leg piping material

3.3 Combined analysis

Based on this concept, combined analysis is proposed as the evaluation method of J-R curve to long crack extension using the test results with two different crack extensions. The procedure is as follows; At first, the P_{Ni} - v'_{pli} data pair is obtained by using load – load line displacement curve for long crack extension length in accordance with Eqs.(9) and (10), and final P_{Ni} - v'_{pli} data pair is obtained for two specimens respectively, where final P_{Ni} - v'_{pli} values are

$$\text{Final } P_{Ni} = \frac{P_f}{WB\left[\frac{W-a_f}{W}\right]^{\eta_{pl}}} \tag{9}$$

$$\text{Final } v'_{pli} = \frac{v_f - P_f C_f}{W} \tag{10}$$

A line is drawn from the final P_{Ni} - v'_{pli} data pair of short crack extension tangent to the P_N - v'_{pl} curve of long crack extension. The right side data to the tangent point and data with v'_{pli} <0.001 are excluded from effective P_{Ni} - v'_{pli} data pair. The coefficients of the fitting function of Eq.(11) instead of Eq.(6) are calculated for two final P_{Ni} - v'_{pli} values and the effective P_{Ni} - v'_{pli} data pair.

$$P_N = \frac{a + bv'_{pl} + cv'^{2}_{pl} + dv'^{3}_{pl}}{e + v'_{pl}} \tag{11}$$

The following least square method is used for curve fitting of the function of Eq.(11).

$$z = \sum \left\{ P_N \left(e + v'_{pl} \right) - \left(a + bv'_{pl} + cv'^{2}_{pl} + dv'^{3}_{pl} \right) \right\}^2 = \min. \tag{12}$$

The coefficient values, a, b, c, d, e can be calculated directly by Eq.(13).

$$\begin{bmatrix} n & \sum v'_{pl} & \sum v'^{2}_{pl} & \sum v'^{3}_{pl} & -\sum P_N \\ \sum v'_{pl} & \sum v'^{2}_{pl} & \sum v'^{3}_{pl} & \sum v'^{4}_{pl} & -\sum v'_{pl} P_N \\ \sum v'^{2}_{pl} & \sum v'^{3}_{pl} & \sum v'^{4}_{pl} & \sum v'^{5}_{pl} & -\sum v'^{2}_{pl} P_N \\ \sum v'^{3}_{pl} & \sum v'^{4}_{pl} & \sum v'^{5}_{pl} & \sum v'^{6}_{pl} & -\sum v'^{3}_{pl} P_N \\ \sum P_N & \sum v'_{pl} P_N & \sum v'^{2}_{pl} P_N & \sum v'^{3}_{pl} P_N & -\sum P_N^2 \end{bmatrix} \begin{Bmatrix} a \\ b \\ c \\ d \\ e \end{Bmatrix} = \begin{Bmatrix} \sum v'_{pl} P_N \\ \sum v'^{2}_{pl} P_N \\ \sum v'^{3}_{pl} P_N \\ \sum v'^{4}_{pl} P_N \\ \sum v'_{pl} P_N^2 \end{Bmatrix} \tag{13}$$

Figure 14 shows normalized load - displacement curve best-fit by Eq.(11) for two final points of short and long crack extension cases and the effective P_{Ni} - v'_{pli} data pair. Next, the crack length a_i coinciding with P_{Ni} in Eq.(4) and with P_N in Eq.(11) is calculated for each v'_{pli} by checking with slightly increasing crack lengths from initial crack length a_0, where load - displacement curve for long crack extension length is used. However, J-R curve obtained using combined analysis was deviated from individual J-R curve for short and long crack extension respectively in the case of hot leg pipe material as shown in Fig. 14. This reason is

that load - displacement curve between short and long crack extension have slightly different shape as shown in Fig. 11. Therefore, it is needed to adjust the position of middle point by reflecting the characteristics of J-R curves for short and long crack extension. To do so, the coincidence level is evaluated by comparing the J-R curves between normalization analysis by only short crack extension and combined analysis. As a method of evaluation for coincidence, best fit curve of Eq.(14) for the J-R curve of short crack extension is used.

$$J = C(\Delta a)^m \tag{14}$$

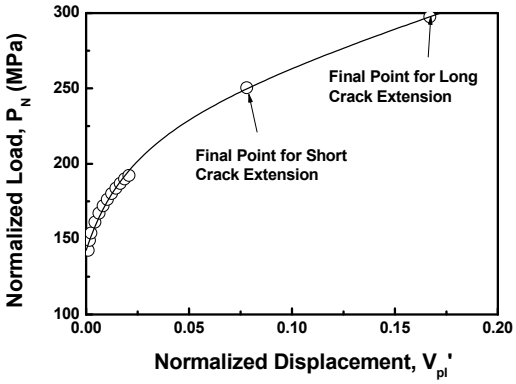


Fig. 14. The best fit curve by Equation (11) on effective data pair for combined analysis

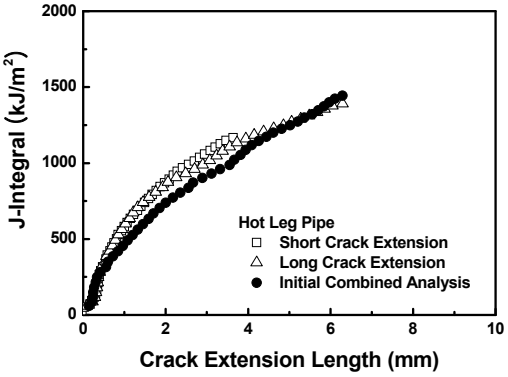
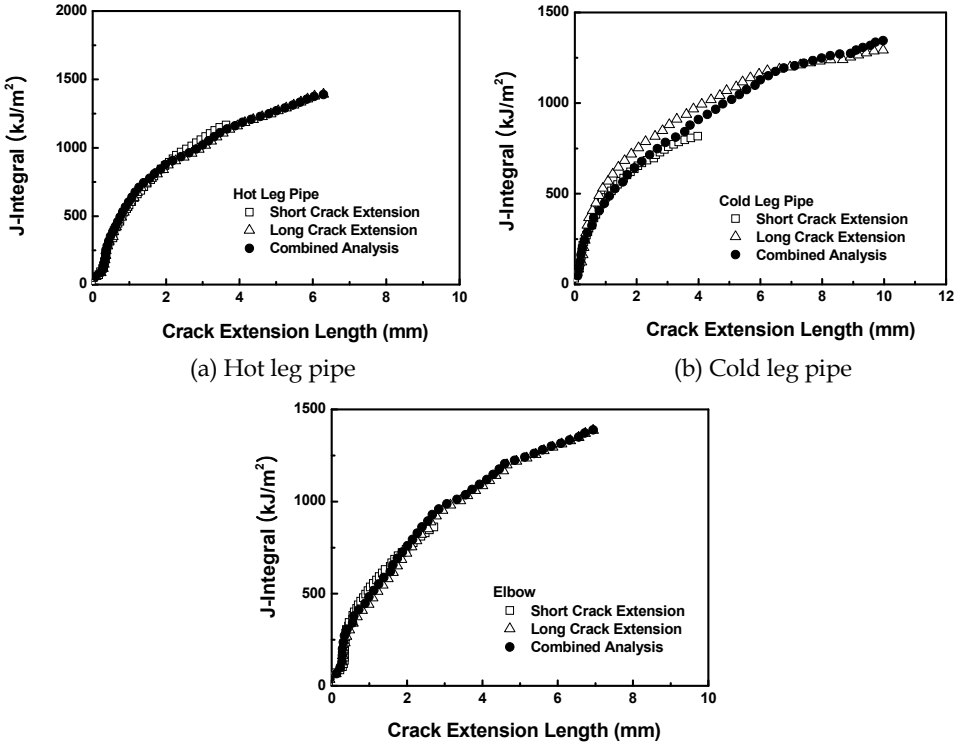


Fig. 15. Dynamic J-R curve for hot leg pipe material prior to adjustment of middle point on normalized load versus displacement curve in combined analysis

Next, the standard deviation σ of Eq.(15) is calculated from J value by combined analysis and J value obtained by J-R curve of Eq.(14). Such that, the data of combined analysis to short crack extension are used in calculating σ

$$\sigma = \sqrt{\frac{\sum (J_{fit} - J_{combined})^2}{n-1}} \tag{15}$$

where J_{fit} is J value obtained by fitting function of Eq.(14) $J_{combined}$ is J value obtained by combined analysis and n is the number of effective J-R data to short crack extension. Optimal middle point on the normalized load-displacement relationship is determined as a point when standard deviation σ value of Eq.(15) is reached to minimize by adjusting P_N value at v'_{pl} value at final point of short crack extension. Using the optimal middle point, final P_{Ni} - v'_{pli} data pair of long crack extension and effective P_{Ni} - v'_{pli} data pairs, J-R curve can be estimated. Figure 9 shows the comparison of dynamic J-R curve among the combined method and normalization method of short and long crack extension. For all three kinds of piping, dynamic J-R curve by combined analysis is well described with the behavior of that for two different crack extensions. From this combined analysis, we could obtain reasonable dynamic J-R curve until long crack extension for nuclear piping materials. In combined analysis, one J-R curve is obtained using two specimens. Therefore, the scatter of material properties with the position of taking specimen is required not to be large. In LBB analysis, the lowest material property is used among three test results for material property scatter. In this approach, the J-R curve tends to be estimated as an average J-R data for two test results. Further investigation is therefore needed for low bound curve of J-R curve with long crack extension effectively based on the statistical concept.



(a) Hot leg pipe (b) Cold leg pipe

(c) Elbow

Fig. 16. The dynamic J-R curve by combined analysis for each material

4. Conclusion

From the comparison test results between DCPD and normalization method as a dynamic J-R curve testing method, short crack extension, dynamic J-R curves were similar but, for long crack extension, J-R curve estimated by normalization was higher by 10~30% at the initial loading stage than that by DCPD. For reliable J/T analysis for LBB design of nuclear piping, material J-R curve for long crack extension is needed. However, normalization method is applicable for only short crack extension. To overcome this problem, combined analysis based on normalized method was proposed. In combined analysis, dynamic J-R curve with long crack extension is estimated by two dynamic J-R curve tests with different crack extension length. The dynamic J-R curve beyond the crack extension length range designated by ASTM code could be estimated using the combined analysis.

5. References

ASTM (2009). ASTM E1820-09e1 Standard Test Method for Measurement of Fracture Toughness, In: *Annual Book of ASTM Standard*, Vol. 03.01, ASTM International, West Conshohocken, Pennsylvania, USA

Ernst, H.A., Paris, P.C., Rowssow, M. & Hutchinson, J.W. (1979). Analysis of Load Displacement Relationship to Determine J-R Curve and Tearing Instability Material Properties. In: *ASTM STP 677 Fracture Mechanics*, Smith, C.W. (Ed.), pp. 581-599, ASTM International, ISBN EB 978-0-8031-4746-1, West Conshohocken, Pennsylvania, USA

Ernst, H.A., Paris, P.C. & Landes, J.D. (1981). Estimations on J-integral and Tearing Modulus T from a Single Specimen Test Record. In: *ASTM STP 743 Fracture Mechanics*, Roberts, R. (Ed.), pp. 476-502, ASTM International, ISBN EB 978-0-8031-4809-3, West Conshohocken, Pennsylvania, USA

Hackett, E.M., Kirk, M.T. & Hays, R.A. (1986). *NUREG/CR-4550 : An Evaluation of J-R Curve Testing of Nuclear Piping Materials Using the Direct Current Potential Drop Technique*, U.S. Nuclear Regulatory Commission

Johnson, H.H. (1965). Calibrating the Electric Potential Method for Studying Slow Crack Growth. *Materials Research and Standards*, (September 1965), Vol.5, No.9, pp. 442-445, ISSN 0025-5394

Joyce, J.A. (1996). *Manual on Elastic-Plastic Fracture Laboratory Test Procedures*, ASTM International, ISBN 0-8031-2069-9, West Conshohocken, Pennsylvania, USA

Kim, J.W. & Kim, I.S. (1997). Investigation of Dynamic Strain Aging on SA106-Gr.C Piping Steel. *Nuclear Engineering and Design*, Vol. 172, No. 1-2, (July 1997), pp. 49-59, ISSN 0029-5493

Scott, P.M., Olson, R.J. & Wilkowski, G.M. (2002). *NUREG/CR-6765: Development of Technical Basis for Leak-Before-Break Evaluation Procedures*, U.S. Nuclear Regulatory Commission

Landow, M.P. & Marschall, C.W. (1991). Experience in Using Direct Current Electric Potential to Monitor Crack Growth in Ductile Metals, In: *ASTM STP 1114 Elastic-Plastic Fracture Test Methods*, Joyce, J.A. (Ed.), pp. 163-177, ASTM International, ISBN-EB 978-0-8031-5172-7, West Conshohocken, Pennsylvania, USA

Landes, J.D., Zhou, Z., Lee, K. & Herrera.,R. (1991). Normalization Method for Developing J-R Curve with the LMN Function. *Journal of Testing and Evaluation*, Vol. 19, No. 4, (July 1991), pp. 305-311, ISSN 0090-3973

Lee, B.S., Yoon, J.H., Oh, Y.J., Kuk, I.H. & Hong, J.H. (1999). Static and Dynamic J-R Fracture Characteristics of Ferritic Steels for RCS Piping, *15th International Conference on Structural Mechanics in Reactor Technology*, Vol. V, pp. 297-302, ISBN 89-88819-05-5 94500, Seoul, Korea, August 1999

Lee, J.B. & Choi, Y.H. (1999). Application of LBB to High Energy Pipings of a Pressurized Water Reactor in Korea, *Nuclear Engineering and Design*, Vol.190, No.1-2, (June 1999), pp.191~195, ISSN 0029-5493

Nakamura, T., Shih, C.F. & Freund, L.B. (1986). Analysis of a Dynamically Loaded Three-Point-Bend Ductile Fracture Specimen, *Engineering Fracture Mechanics*, Vol. 25, No. 3, pp. 323-339, ISSN 0013-7944

Oh, Y.J, Kim, J.H. & Hwang, I.S. (2002). Dynamic Loading Fracture Tests of Ferritic Steel Using Direct Current Potential Drop Method. *Journal of Testing and Evaluation*, Vol. 30, No. 3, (May 2002), pp. 221-227, ISSN 0090-3973

Sharobeam, M.H. & Landes, J.D. (1991). The Separation Criterion and Methodology in Ductile Fracture Mechanics. *International Journal of Fracture*, Vol. 47, No.2, (January 1991), pp. 81-104, ISSN 0376-9429

Wallen, K. (2009). Extrapolation of Tearing Resistance Curves. *2009 Proceeding of the ASME Pressure Vessel and Piping Conference*, Vol.3, pp. 281-286, ISBN 978-0-7918-4366-6, Prague, Czech Republic, July 2009

11

Feed Water Line Cracking in Pressurized Water Reactor Plants

Somnath Chattopadhyay
Georgia Southern University, Statesboro, Georgia, USA

1. Introduction

As early as 1979, a through wall crack was detected in a pressurized water reactor (PWR) plant. This crack initiated at the counter bore region of the pipe, adjacent to the weld joint attaching the pipe to the steam generator feed water nozzle. Subsequent inspection of the remaining feed water piping revealed cracking in the same vicinity but these were limited to partial wall penetration. As a result of this incident, the US Nuclear Regulatory Commission issued a directive to all operating plants requiring them to perform inspection of their feed water lines. The cracks were subsequently detected in the immediate vicinity of the steam generator nozzles in a number of plants. An exhaustive investigation was undertaken subsequently and this revealed that the primary cause of cracking was due to a fatigue loading mechanism induced by thermal stratification and high cycle thermal oscillations (striping) during low flow conditions.

Thermal stratification phenomenon results from a temperature differential across the pipe cross section with the top fluid stream hot and bottom stream relatively cold. During normal plant operations at low flow conditions, when the feed water nozzle is not completely full, hot water from the steam generator remains in the nozzle to fill up the rest of the volume. The difference in buoyancy between the hot and cold fluids inhibits their mixing so that the feed water becomes and remains thermally stratified. Separation of these two flow regions is due to the density difference in the hot and cold streams. The stratified temperature conditions can produce very high stresses, and can occur may times during normal low power operations; therefore this has the potential to initiate cracks in a relatively short period of time. Thermal striping is a local phenomenon that occurs at the interface between hot and cold flowing fluids. The interface level oscillates with periods ranging from 0.1 to 10 seconds. The oscillating fluid temperature gives rise to fluctuating stresses. The magnitudes of the striping stresses are not as high as those due to stratification itself, but the number of cycles is so large that they contribute significantly to fatigue crack initiation.

During normal plant operation, a series of temperature measurements has been taken around the pipe circumference at the vicinity of the of the feed water nozzle/pipe weld. Analysis of the data indicates that the stratified temperature distributions may be grouped into a handful of basic profiles corresponding to different levels of the interface between the hot and cold fluids. For analysis purposes these profiles could be assumed to be at steady state conditions because of their long duration observed during the tests. Nuclear piping systems (Class 1) are designed according to the rules of NB 3600 of the ASME Boiler and

Pressure Vessel Code, Section III. The loads producing the stresses originate from the internal pressure, mechanical loads due to deadweight, seismic and thermal expansion and the operating thermal transients. Normally piping systems are not designed for circumferential temperature variation. The effect of the thermal stratification on the state of stress in the pipe is manifested in two ways: (a) the difference in temperature between the top and bottom of the pipe causes greater thermal expansion at the top tending to bow the pipe. When such bowing is restrained global bending stresses result; (b) the interface between the two fluid layers causes a local stress in the pipe due to thermal discontinuity across the pipe section. The fatigue damage produced by thermal stratification and the associated thermal striping are a good indication of the contribution of these phenomena to the observed feed water line cracking.

A detailed finite element stress analysis has been carried out using a three dimensional model that includes the steam generator shell, the feed water nozzle, and the elbow/pipe. The shell nozzle/elbow model contains three distinct regions with different heat transfer characteristics between the metal and the adjacent fluid. Each of the stratification profiles produces a complex state of stress throughout the nozzle and the elbow (pipe). Different levels of interface produce peak stresses at different locations around the circumference. Since the interface level varies during low flow operating conditions, each point in the counter bore area is subjected to a state of varying stresses of large magnitudes. A maximum range of stress intensity analysis was carried out prior to fatigue evaluation to determine whether the simplified elastic plastic analysis procedure would be required, and if so, to calculate the plastic intensification factor K_e by which the peak alternating stresses would be multiplied. The analysis predicted crack locations that that correlated well with the observed cracking.

The major cause of growth of the cracks is due to the thermal stratification cycles, which occur during low flows, primarily at hot standby. The thermal striping phenomenon or the oscillations occurring at the interface between hot and cold fluids has some influence on the crack growth, but it certainly impacts the crack initiation predictions. Thermal stratification causes a stress distribution in a pipe that is similar to what happens in a bimetallic strip. In the hot upper region compressive stresses develop as a result of constrained expansion, with the tensile stresses occurring in the lower region. This has been demonstrated using a simplified 2-dimensional finite element model. These are essentially the membrane stresses in the axial direction. Since the piping is flexible, the thermal moment gives rise to a bending stress that is added to the membrane stresses to obtain the total stresses.

It is suggested that the equations for obtaining stresses in piping systems as outlined in the ASME Code contain a term addressing circumferential temperature gradients in the pipe. A number of remedial measures have been implemented or suggested in operating power plants to minimize the stress amplitudes and frequency of load cycling during the stratification events.

In recent years, thermal stratification phenomenon has been observed to exist on several piping systems in pressurized water reactors. Damages have been observed in the main feed water lines, pressurizer spray lines, unisolable branch piping connected to reactor coolant piping, and pressurizer surge lines, with evidence linked to thermal stratification. The stratification phenomenon results from a temperature differential across the pipe cross-section with the top fluid stream hot and the bottom stream relatively cold. This condition occurs under relatively low flow conditions by cold feed injection into a stagnant hot pipe region or vice versa. Separation of two fluid flow areas is due to density differences in the

hot and cold streams. This gives rise to gross thermal bending moments across the pipe section resulting in bowing deformation of the pipe.
In May 1979, a pressurized water reactor plant in operation approximately a year developed a through-wall crack in one of its feed water lines at the entrance to the steam generator. Subsequent investigation of the remaining lines revealed cracking in the same vicinity but limited to partial wall penetration. As a result of this incident, the United States Nuclear Regulatory Commission submitted a directive to all PWR operating plants to perform inspection of their feed water lines. A number of plants produced same degree of cracking in the same general area with wide variety of size, orientation and length of plant operation. Because of the involvement of many variables, it was impossible to immediately identify the specific mechanisms of crack initiation and growth. A number of activities were initiated to investigate the structural, thermal, hydraulic, operational and environmental conditions which individually or collectively contributed to the observed cracking.

2. Observed crack locations

Figure 1 illustrates the feed water pipe to steam generator nozzle junction where majority of cracking occurred. Cracks were found to be oriented circumferentially and located in the base metal outside the heat affected zone. There were intermittent pitting throughout the inside surface. The deepest cracks were found at the base of the counter-bore transition

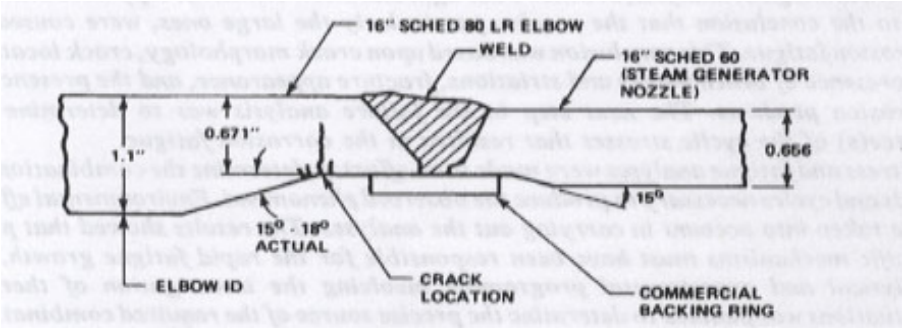


Fig. 1. Location of Cracks in PWR Feed water Pipe to Nozzle Attachment Region [1]

Typically a majority of PWR plants produced the circumferential cracking, the pattern of depth orientation varied considerably for different plants. Generally the deepest cracking was observed at the top, although in a number of plants this was found to occur at the sides, as well as the bottom. With the exception of one through-wall condition, most plants produced relatively small shallow cracks.

3. Metallurgical studies

The metallurgical investigations revealed that although corrosion may have been a major factor in initiating the cracks, the primary driving force for crack growth was mainly mechanical in nature. The corrosion fatigue may have resulted the cracking; both high and low cycle fatigue were involved, with high cycle initiating the fatigue and the low cycle propagating it. The fracture appearances were studied at high magnification by electron microscopy. Striations were found (Figure 2) substantiating the evidence that crack growth

was taking place by fatigue, although the striation spacing was unreliable as a measure of the growth rate, since a large range of temperatures were involved (200 - 450°F).

Fig. 2. Fractographs of the tip of a Deep Crack [1]

Instrumentations were installed at various plants to measure vibration and displacements of the feed water piping as well as temperatures in the vicinity of pipe to nozzle junction. The plants (both with and without observed cracking) were surveyed to determine their transient operation history and chemistry control. Particular attention was paid to the feed water oxygen content because of the presence of pitting. Thermocouple data of the on-site testing demonstrated the existence of persistent pipe thermal stratification during low feed water flow operations such as feed water makeup cycling during hot standby.

4. Flow model studies

Based on flow model tests it was shown that the temperature profile in a stratified cross section is mainly correlated with two thermal hydraulic parameters: (a) the flow rate in the line, and (b) the temperature difference between the top and the bottom of the pipe cross-section under consideration. The flow model test was a full scale feed line and nozzle assembly made of Plexiglas for visual observation and fluid temperature measurement

(Figure 3). The test was designed to establish the temperature profile of the stratified water more accurately than the field measurements and to determine that thermal striping exists at the stratified interface, and if so determine the magnitudes and frequencies.

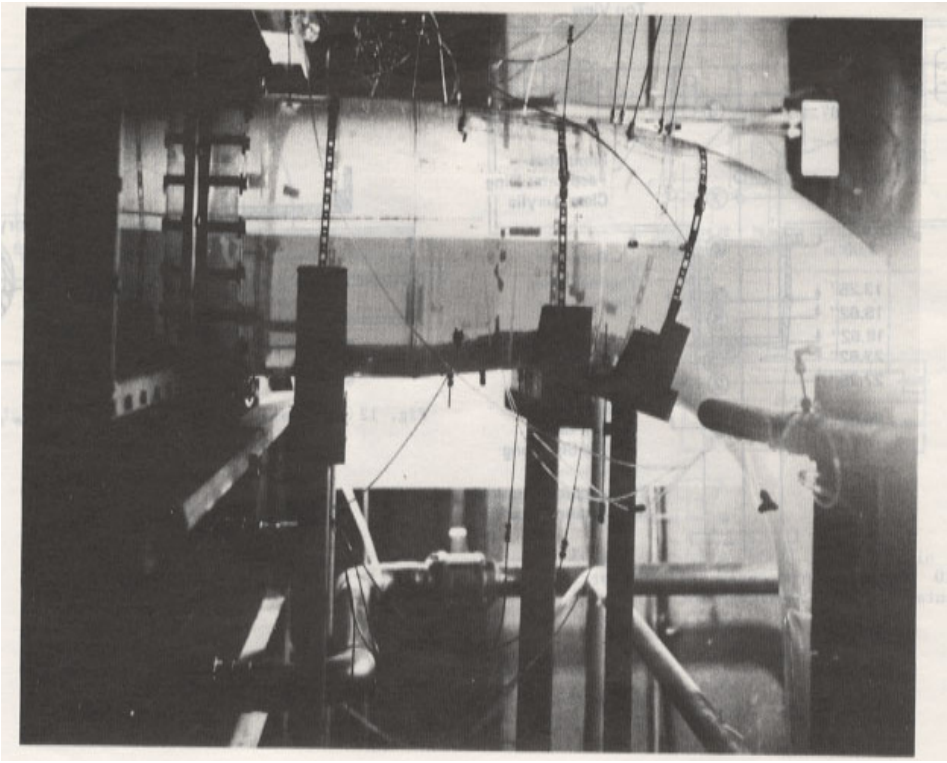


Fig. 3. Flow Model Test showing Stratification: upper clear layer hot water, lower gray layer cold saline solution [2]

The fluid temperature oscillations were recorded and it was subsequently confirmed that thermal striping mechanism led to feed line thermal fatigue.

5. Structural analysis

`During normal plant operation at low power conditions water is supplied to the steam generators at very low flow rates. When the flow rate is not high enough to completely fill the nozzle, hot water from the steam generator remains in the nozzle to fill up the rest of the volume. The difference in buoyancy between the hot and cold fluids inhibits their mixing so that the feed water becomes and remains thermally stratified as long as the flow rate is less that that required to completely fill the nozzle. During normal plant operation a series of temperature measurements was taken around the pipe circumference at the vicinity of the pipe weld. Analysis of the test data indicated that the stratified temperature distributions may be grouped into six basic profiles corresponding to different levels of the interface between the hot and cold fluids and are shown in Figure 4.

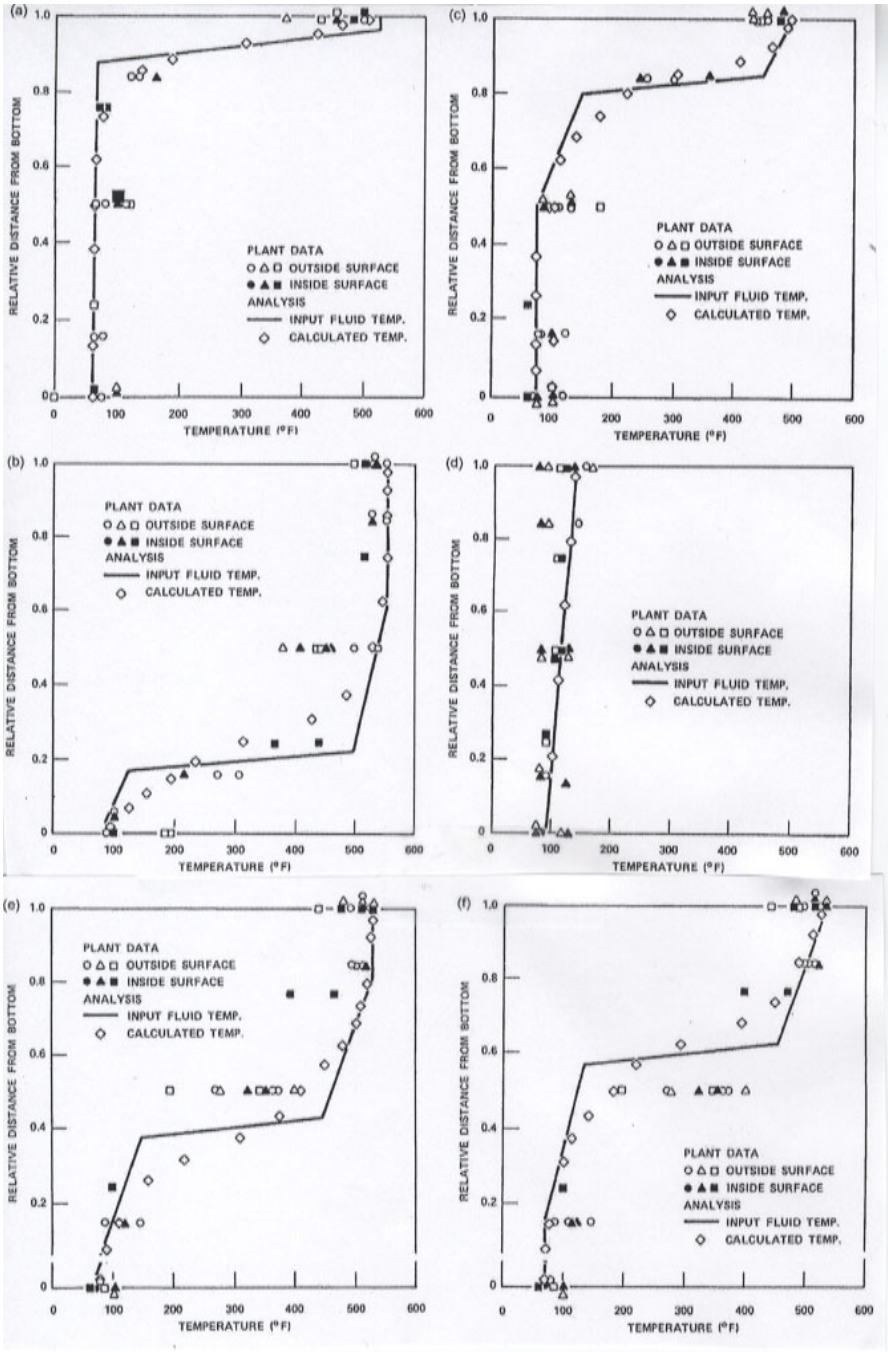


Fig. 4. Stratified Temperature Profiles [3]

A finite element model has been prepared that includes a part of the steam generator shell, the feed water nozzle and the connecting elbow. The model uses 20-node isoparametric solid elements, two elements through the thickness and twelve around the circumference of the model.
The shell/nozzle/elbow model contains three distinct regions with different heat transfer characteristics between the metal and the adjacent fluid. The first region is that of the inside of the steam generator shell exposed to slowly moving hot water. The other regions are the section of the nozzle under the thermal sleeve, and the rest of the nozzle and the elbow.
Each of the stratification profiles produces a complex stress state throughout the nozzle and the elbow. The highest stresses occur in the weld counter bore region at the root of the elbow transition. For each profile there is a zone of compressive stress above the hot/cold interface and a region of tensile stress below it. Different interface levels produce peak stresses at different locations around the circumference. Since the interface level varies during low flow operating conditions, each point in the counter bore area is subjected to varying stress state.
Fatigue evaluations have been performed around the circumference for the counter bore transition root and along the top and side of the counter bore region. The load conditions and the number of cycles were combined with a pressure of 7.6 MPa. A maximum range of stress intensity analysis was performed prior to each fatigue evaluation to determine whether the simplified elastic plastic analysis procedure would be required and if so, to calculate the plasticity intensification factors, K_e factors by which the peak alternating stresses are to be multiplied.
The results for a typical plant fatigue evaluation [3] indicate that the peak usage factors are well above 1.0 and occur at the top and sides. These correlate with the observed locations of the deepest cracks for that plant. The high usage factors conclusively implicate thermal stratification and thermal striping during low flow conditions as prime contributors to the observed feed line cracking.

6. Analytical studies

An analytical technique has been developed to evaluate the stresses due to circumferential temperature gradient during thermal stratification. The associated numerical solution is an approximate one that uses the standardized profiles of Figure 4. The mean temperatures at various circumferential pipe segments are calculated and shown in Figure 5.
The stress distributions for the Profiles 1 through 6 has been computed using the approximate numerical model and are shown in Figure 6.
The maximum range of stresses occurs at the top of the pipe and equals 72-(-124) = 196 MPa (based on profiles 2 and 1). Although the peak stress due to through the thickness temperature has not been explicitly considered, a conservative value of 2.0 is used. This makes the alternating stress amplitude as 196 MPa, which gives the allowable number of cycles about 30,000 using the design curve of [5]. The plant data in [4] indicates a comparable number of stratification temperature excursions. This leads to a significant fatigue usage factor at the top of the pipe that correlates with the fatigue cracks observed at this location.

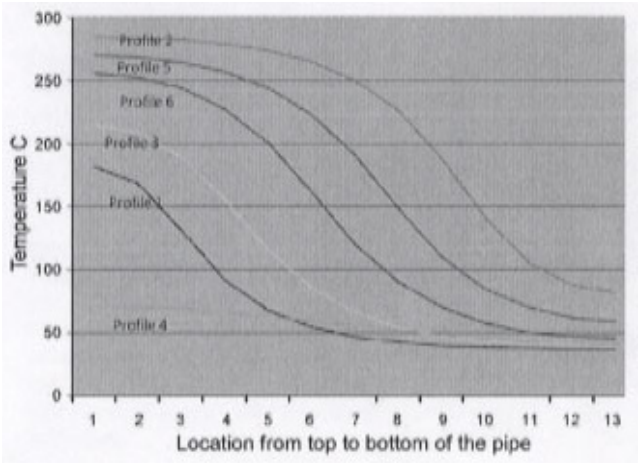


Fig. 5. Calculated Temperature input to the Approximate Numerical Model [4]

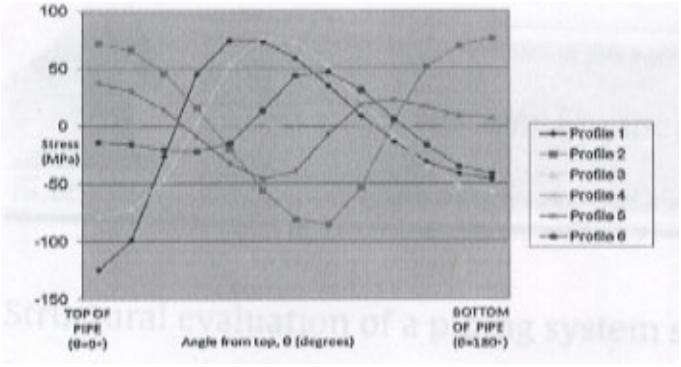


Fig. 6. Stress Distribution across Pipe Diameter for Profiles 1 through 6[4]

7. References

[1] Enrietto, J.F., Bamford, W. H., and White, D. H. (1981), "Preliminary Investigation of PWR Feed water Line Cracking, *International Journal of Pressure Vessels and Piping*, 9, pp. 421-443.

[2] Hu, M. H., Houtman, J. L., and White, D. H. (1981) "Flow Model test for the Investigation of Feed water Line Cracking for PWR Steam Generators, *ASME Paper 81-PVP-4.*

[3] Thurman, A. L., Mahlab, M. S., and Boylstein, R. E. (1981), "3-D Finite Element Analysis for the investigation of Feed water Line Cracking in PWR Steam Generators, *ASME Paper 81-PVP-3.*

[4] Chattopadhyay, S. (2009), "Structural Evaluation of a Piping System Subjected to Thermal Stratification," *Nuclear Engineering and Design*, 239, pp. 2236-2241.

[5] ASME Boiler and Pressure Vessel Code, 2010, Section III, Nuclear Power Components, American Society of Mechanical Engineers, New York.

12

Degradation Due to Neutron Embrittlement of Nuclear Vessel Steels: A Critical Review about the Current Experimental and Analytical Techniques to Characterise the Material, with Particular Emphasis on Alternative Methodologies

Diego Ferreño, Iñaki Gorrochategui and Federico Gutiérrez-Solana
University of Cantabria (UC) - Technological Centre of Components of Cantabria (CTC)
Spain

1. Introduction

The pressure vessel constitutes the most important structural component in a nuclear reactor from the point of view of its safety. The core of the reactor, that is, the nuclear fuel, is accommodated inside the vessel. This material is composed of fissile nuclides that undergo chain nuclear reactions of an exothermic nature, thus generating usable energy. Uranium 235 (U-235) is the only isotope in Nature which is fissile with thermal neutrons; hence, it is used as nuclear fuel in Light Water Reactors (LWRs). Two technologies of LWRs can be distinguished, Pressurized Water Reactors (PWRs) and Boiling Water Reactors (BWRs). Currently, more than 400 nuclear reactors operate in the world of which, approximately, 57% are PWR and 22% BWR. The original design lifetime for LWRs is 40 calendar years; nevertheless, the current target for most plants in many countries in Europe, Japan and the USA is to extend their operative lifetime up to 60 years.

The nuclear vessel is a virtually irreplaceable element which is subjected to operating conditions that lead to a progressive degradation in time of its constituent steel. The chain fission reactions of U-235 entail the emission of high energy neutrons that inevitably impact the internal surface of the vessel. These collisions give rise to a complex series of events in the nano and microstructural scale that, in the end, modify the mechanical properties of the steel leading to its embrittlement, that is, the decrease in its fracture toughness. This phenomenon is most intense in the so called beltline region (which is the general area of the reactor vessel near the core midplane where radiation dose rates are relatively high). The total number of neutrons per unit area that impact the internal surface of the vessel represents the neutron fluence; in practice, only that fraction of the energy spectrum corresponding to a neutron kinetic energy higher than 1 MeV is considered to be capable of triggering damage mechanisms in the vessel steel; these neutrons are referred to as fast neutrons. Typical design end of life (EOL) neutron

fluences (E>1 MeV) for BWRs are in the order of 10^{18} n/cm^2, whereas for PWRs this number is about 10^{19} n/cm^2.
In Section 2 of this chapter, the embrittlement of nuclear vessel steels is described from a purely phenomenological perspective as well as from the point of view of the legislation currently in force. The phenomenon of the ductile to brittle transition and the influence of embrittlement on it are particularly stressed. In Section 3, a brief description of the main characteristics of the nuclear power plants surveillance programmes is presented; the requirements that they envisage as well as the information that they allow to be obtained are pointed out. The physical mechanisms that take place in the nano and micro levels leading to the material embrittlement are detailed in Section 4 where a brief exposition concerning the most relevant predictive models for embrittlement is also presented. Finally, in Section 5, the promising method of the Master Curve is described; this represents an improved methodology for the description of the fracture toughness of vessel steels in the ductile to brittle transition region, susceptible to be incorporated in the current structure of the surveillance programmes.

2. The embrittlement of nuclear vessel steels and its influence on the ductile to brittle transition region: Phenomenology and regulations

2.1 The phenomenology of the embrittlement and the ductile to brittle transition region

As mentioned above, neutron irradiation reduces vessel steel toughness. To understand the concept of material fracture toughness, Fracture Mechanics theory must be referred to. In this section, the basic principles of Fracture Mechanics are briefly presented. Material toughness can be understood as its resistance to be broken when subjected to mechanical loading (forces, stresses) in the presence of cracks / flaws (that is, a sharpened discontinuity). Traditionally, RPV toughness variation has been indirectly quantified by means of Charpy impact tests. These tests were introduced by the French researcher George Charpy in 1905. The preparation, execution and interpretation of Charpy tests are currently regulated by the ASTM E 23 standard (ASTM E 23, 2001). The test consists of breaking a small dimension (10 x 10 x 55 mm^3) specimen by the impact of a pendulum released from a controlled height and measuring the height it achieves after the breaking (see Figure 1). First, a notch has to be machined in the specimen that acts as stress concentrator, forcing the fracture process. The expression "notch" refers to a blunted defect and, therefore, must not be misunderstood as a crack or flaw (described above). The main numerical result provided by the Charpy test (though not the only one) is the energy absorbed during the specimen breaking, also called resilience, which is used as a semi-quantitative estimation of the material fracture toughness.
Because of its simplicity of execution and for historical reasons, this experimental technique is still employed in several industries such as the naval, building and pressure vessel design and, in particular, in nuclear vessels. However, several limitations associated to the Charpy impact test can be mentioned. First, the fracture process occurs under highly dynamic conditions and, therefore, can hardly be extrapolated to quasi-static fracture processes (as the ones occurring in many components and, predictably, in nuclear vessels). Moreover, it has been demonstrated that the results provided by the Charpy test, such as the resilience, do not represent genuine material properties, which can be applied to components of different dimensions and loading conditions; as mentioned above, the resilience may be

considered only as a semi-quantitative estimator. Finally, the standard specimen notched geometry, which facilitates its breaking, is far from simulating an ideal crack, that is, a sharpened discontinuity embedded in the material.

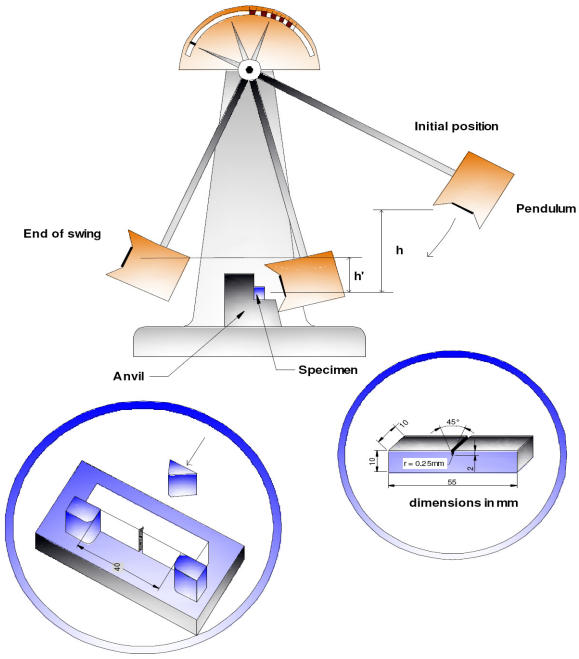


Fig. 1. Sketch of the Charpy pendulum and the standard Charpy specimen

Experience shows that the vessel steel fracture properties drastically depend on temperature. This phenomenon is clearly appreciated in the resilience curves, which are obtained from Charpy tests carried out at different temperatures. Thus, in Figure 2, it can be appreciated how the absorbed energy in the specimen breaking process varies from low values at low temperatures, in the region called brittle or Lower Shelf, to high absorbed energies at high temperatures, in the ductile region or Upper Shelf. The intermediate temperatures constitute what is called the Ductile to Brittle Transition Region, DBTR. One additional property of the DBTR is that the resilience values present an important scatter, as can also be inferred in Figure 2. In order to simplify the analysis, it is common practice to fit the data according to mathematical expressions such as the hyperbolic tangent, see Figure 2 (the free parameters A, B, C and D must be obtained through the fitting procedure, typically a least squares method). The existence of three regions and the scatter of the results in the DBTR are also observable in the fracture toughness curves. According to this, the description of the vessel steels toughness in the DBTR represents a remarkable engineering problem because, on the one hand, this property changes ostensibly with temperature and, on the other hand, it shows an important scatter.

Figure 3 shows an example of a vessel steel neutron irradiation embrittlement. In this figure, Charpy impact test results (absorbed energy vs. temperature) from unirradiated and irradiated material, respectively, are represented. As can be seen, the most remarkable effects are a shift

of the curve in the DBTR towards higher temperatures accompanied by a reduction in the absorbed energy in the Upper Shelf, that is, a generalised embrittlement process.

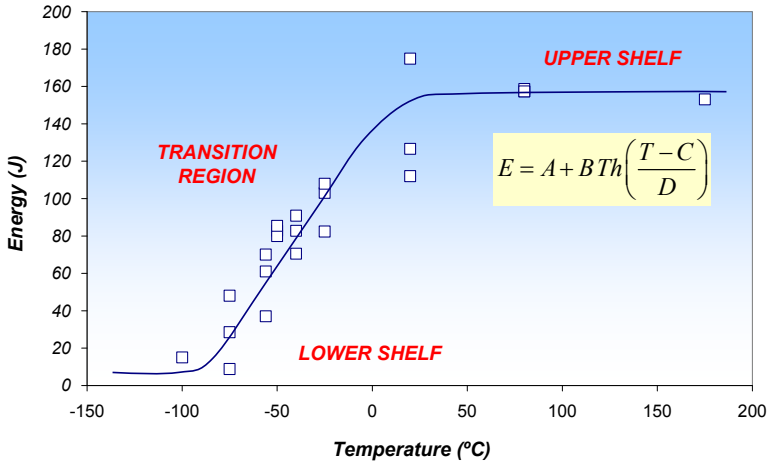


Fig. 2. Influence of temperature on the results (absorbed energy) of the Charpy test; the data were fitted to a hyperbolic tangent curve

In regard to nuclear vessel structural integrity assessment, it is of great interest to determine the neutron irradiation effect on the fracture properties in the DBTR. In the nuclear industry, it is common to simplify this phenomenon by means of the definition of standard temperatures that are considered as representative. Among these, the most used ones are T_{28J} and T_{41J}, that is, the temperatures where the Charpy impact test absorbed energies are 28 J and 41 J, respectively (obtained after fitting the results to a hyperbolic tangent curve).

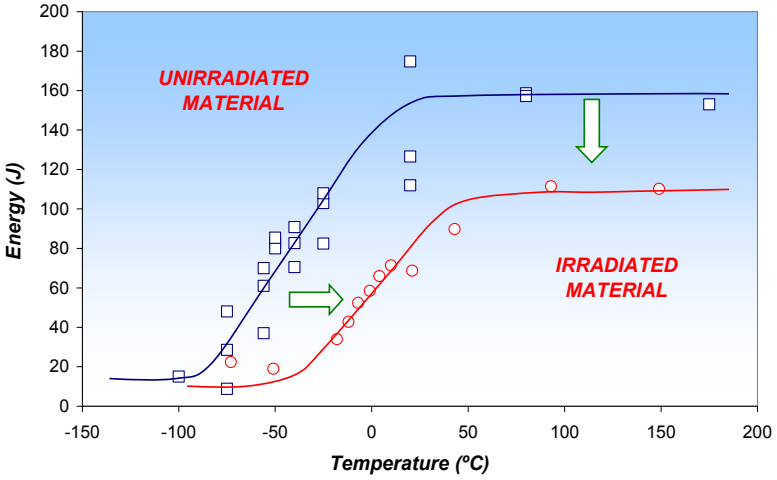


Fig. 3. Description of the influence of neutron irradiation on the Charpy curves (absorbed energy vs. temperature)

2.2 The ASME curves and the reference temperature RT_{NDT}

Since 1905, the year when George Charpy proposed his famous test, until the present, the knowledge of the mechanics that control the fracture process has experienced an important qualitative progress (a historic review can be found in the reference (Anderson, 1995)). Briefly summarising, at the end of the 50's the formulation of the so-called Linear Elastic Fracture Mechanics[1] (LEFM) discipline was available. LEFM accurately describes the fracture processes of brittle materials, that is, the materials that show a linear elastic response until failure (without the occurrence of previous relevant yielding). LEFM demonstrates that, the presence of an ideal crack in this type of materials when subjected to loading implies a stress state tending to infinite in the proximity of the crack front. On the other hand, it can also be demonstrated that the stress and strain fields, see Figure 4, in the proximity of the crack front depend exclusively on one parameter known as stress intensity factor (SIF or K_I), which is a function of the applied load (σ) and the defect size (a). It is worth mentioning that the subscript 'I' refers to the fracture mode I (tensile mode) as represented in Figure 4, which is the prevailing one in most cases of interest.

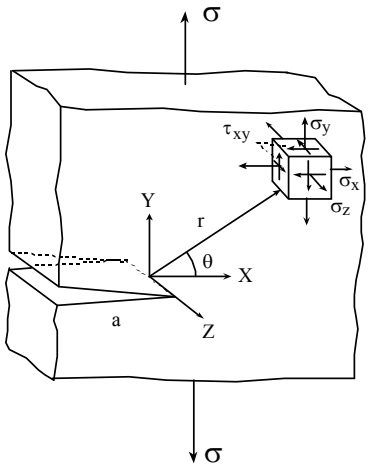


Fig. 4. Sketch of a cracked component / specimen and the stress state in the proximity of the crack front

Then, as K_I controls the stress-strain field in the fracture process zone (mode I), breaking will occur when K_I reaches a determined critical value, which will be dependent on the material. This reasoning provides a valid definition of the fracture toughness as the SIF critical value, K_{Ic}, and also the following fracture criterion (1):

$$K_I(\sigma, a) = K_{Ic}(material) \tag{1}$$

The SIF (and consequently the toughness K_{Ic}) are usually expressed in MPa $\sqrt{m}^{1/2}$. Toughness K_{Ic} is determined following the rules of the ASTM E 399 standard (ASTM E 399, 2009) or some other equivalent procedure.

[1] Thanks to the works of A.A. Griffith (Griffith, 1920), C.E. Inglis (Inglis, 1913), G.R. Irwin (Irwin, 1956, 1957) or H.M. Westergaard (Westergaard, 1939).

Unfortunately, LEFM is not valid for tough materials (such as reactor vessel steels), because these develop important yielding regions in the crack front before the occurrence of the breaking. In the 60's, several theoretical solutions describing the fracture processes in tough materials were proposed. The works of A.A. Wells (Wells, 1961) and J.R. Rice (Rice, 1968) allows the validity of LEFM to be extended, thus establishing the foundations of the so-called Elastic-Plastic Fracture Mechanics, EPFM. A new parameter, the J integral (which, like the SIF, depends on the loading state and the defect size) characterises the stress and strain fields in the crack front and, therefore, enables the fracture toughness to be defined generally, as the J integral limit value, J_c (sometimes also named J_{Ic} when it refers to the tensile fracture mode I). Also the following fracture criterion[2] can be presented (2):

$$J(\sigma, a) = J_c(material) \tag{2}$$

The J integral (and therefore the toughness J_c) are usually expressed either in $kJ \cdot m^{-2}$ or in $kPa \cdot m$.

Then, since the 70's, once the Wells and Rice postulations were recognised, a more founded theoretical basis was made available to address material fracture, either brittle (LEFM) or ductile (EPFM). However, in the 60's, the period when Generation III reactors (LWR those covered in this text among them) started operating, EPFM was not available to the power plant design engineers. Among the different possible choices to address this problem, the one adopted by the USA authorities (represented by the Nuclear Regulatory Commission, NRC) is especially noteworthy. This regulation is contained in the ASME (American Society of Mechanical Engineers) Code.

The ASME International Boiler and Pressure Vessel Code (ASME Code) establishes rules of safety governing the design, fabrication and inspection of boilers and pressure vessels, and nuclear power plant components during construction. This standard is currently in force for USA designed pressure vessels. The ASME Code is made up of 11 sections and contains over 15 divisions and subsections. The Sections II (Materials), III (Rules for Construction of Nuclear Facility Components) and XI (Rules for In-service Inspection of Nuclear Power Plant Components) are of relevance for the contents of this chapter.

The researchers who elaborated the ASME code were aware of the intrinsic limitations of the LEFM; in particular, they knew that, a tough material breaks in a brittle manner (that is, in the LEFM range), only when large specimens are used. Thus, for example, the ASTM E 399 standard (ASTM E 399, 2009) includes requirement (3) on the thickness B of a toughness specimen in order to obtain a valid result according to the LEFM requirements.

$$B \geq 2.5\left(\frac{K_{Ic}}{\sigma_Y}\right)^2 \tag{3}$$

where σ_Y represents the material yield stress.

This possibility, however, is not applicable as the best solution to carry out the follow-up of the RPV steel because the space available inside the vessel is limited. Hence, it was decided to develop a different methodology, making use of Charpy impact specimens in the surveillance programmes (see Section 3) to subsequently correlate the results with the

[2] Rigorously, this criterion describes the ductile fracture, that is, a fracture process with no stable propagation and non negligible yielding.

toughness values as defined by the LEFM principles. No doubt, this is an indirect and tortuous path but also the only one available to the ASME society researchers at that time. This decision, motivated by the theoretical limitations at the time it was made, is the cause of the conservatism of the current regulation.

Therefore, from the end of the 60's and the beginning of the 70's, an important experimental campaign was carried out, testing different vessel steels in the DBTR according to the LEFM stipulations. As explained above, in order to obtain representative results for tough steels (as those of nuclear vessels), the LEFM requires the employment of large specimens, which are difficult to be fabricated, handled and tested. The weight of some of these specimens reached a value of several tons. Due to the high cost of manufacturing, handling and testing these specimens, the reference curves were named as the "one million dollars curves". After compiling all the available data, it was appreciated that the curves obtained for the different steels showed, on the whole, a similar appearance; independently of the steel heat or composition, the most remarkable difference was its location on the temperature axis, see Figure 5. In consequence, the universal validity of the curves was established and a new parameter was defined for any sort of vessel steel, called reference temperature RT_{NDT}, which allows its toughness-temperature curve to be located in the abscissa axis. Thus, representing in the abscissa axis the corrected scale ($T - RT_{NDT}$), see Figure 5, all the different curves overlap in one single curve which is assumed to be the universal curve for the material behaviour.

In this sense, the toughness behaviour of the vessel material in the DBTR before irradiation (unirradiated material) is described by the reference temperature $RT_{NDT(U)}$ which is obtained from a combination of the results of Charpy (ASTM E 23, 2001) impact and Pellini drop weight tests (ASTM E 208, 1995). Thus, this procedure implies correlating the fracture material behaviour from dynamic tests with uncracked specimens (only notched). This is not rigorously justified, either theoretically, or experimentally. RT_{NDT} is used to index two generic curves, developed in 1973, provided by the ASME Code relating toughness vs. temperature: the K_{Ic} curve describes the lower envelope to a large set of K_{Ic} data while the K_{IR} is a lower envelope to a combined set of K_{Ic}, K_{Id} (dynamic test) and K_{Ia} (crack arrest test) data, being therefore more conservative than the former. Three important features can be appreciated: first, it is assumed that the ASME curves are representative of any vessel steel; second, in both curves, LEFM is considered; finally, the large scatter in the DBT region is removed by taking into account lower envelopes. Consequently, the method provides high conservatism in most cases.

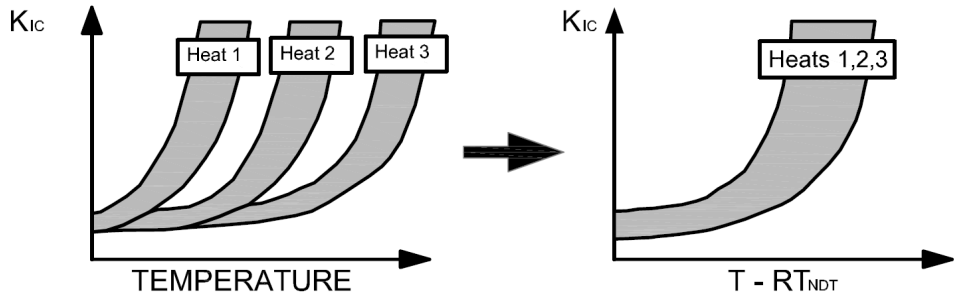


Fig. 5. Description of the process to obtain the ASME curves; definition of the reference temperature RT_{NDT}.

Figure 6 represents the ASME curves along with the experimental population which allowed their definition. The following equations (4, 5) are the mathematical expressions of the two curves:

$$K_{Ic}\left(MPa{\cdot}m^{1/2}\right) = 36.45 + 22.766{\cdot}e^{0.036{\cdot}\left[T(^{o}C)-RT_{NDT}(^{o}C)\right]} \tag{4}$$

$$K_{IR}\left(MPa{\cdot}m^{1/2}\right) = 29.40 + 13.776{\cdot}e^{0.0261{\cdot}\left[T(^{o}C)-RT_{NDT}(^{o}C)\right]} \tag{5}$$

This procedure must be considered as a compromise solution, that is, an engineering tool with relevant limitations and inconsistencies. As will be mentioned in Section 5 of this Chapter, currently an alternative methodology, the Master Curve method, is available. This approach, besides offering a higher simplicity from the methodological and experimental point of view, also provides a complete robustness from the theoretical perspective. The Master Curve procedure enables a realistic follow-up of the surveillance programme vessel steel toughness evolution (see Section 3), avoiding the use of questionable correlations which unnecessarily penalise the representativeness of the subsequent structural calculations.

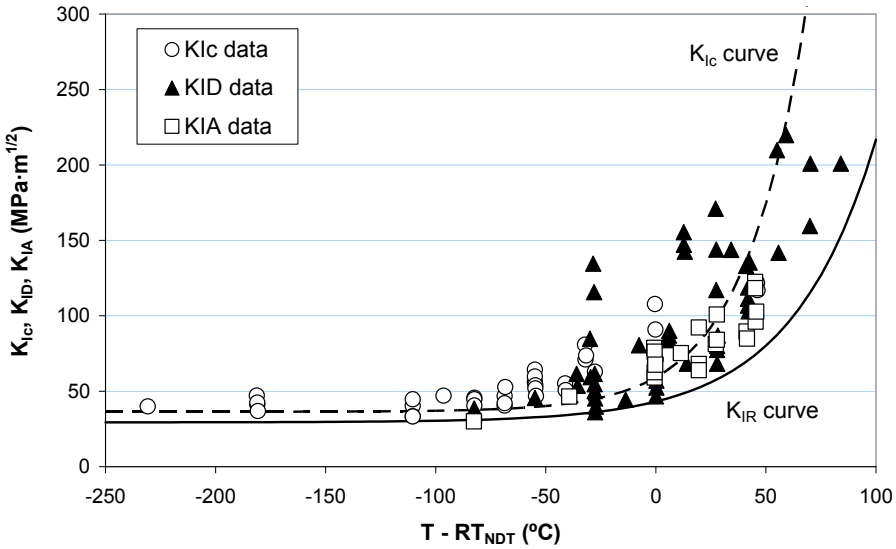


Fig. 6. ASME curves with the experimental data set used for their definition

2.3 Influence of the neutron irradiation on the reference temperature RT_{NDT}

Although the physical mechanisms which lead to the vessel steel embrittlement will be extensively described in Section 4, it is considered convenient to include here some points in order to facilitate the comprehension of Section 3. The reduction of the material toughness due to neutron irradiation in the DBTR is currently estimated through semi-empirical methods based on the shift experienced by the Charpy impact curves obtained from the

surveillance capsule specimens (see Section 3). As stated in 10CFR50 (10CFR50, 1986), the effect of neutron fluence on the behaviour of the material is predicted by the Regulatory Guide 1.99 Rev. 2 (RG 1.99 (2), 1988) which provides equation (6) for the calculation of the evolution of RT_{NDT}:

$$RT_{NDT} = RT_{NDT(U)} + \Delta RT_{NDT} + M \tag{6}$$

where ΔRT_{NDT} represents the shift in the reference temperature due to irradiation which is assumed to be equal to the shift of the Charpy transition curve indexed at 41J; thus, $\Delta RT_{NDT} = \Delta T_{41J}$. The third term, M, is the margin that is to be added to obtain a conservative estimation. The procedure in (RG1.99 (2), 1988) allows ΔRT_{NDT} to be obtained even when no credible surveillance data are available by means of an equation based on the chemistry of the steel and the neutron fluence received (this issue is extensively described in Section 4).

3. Surveillance programmes of nuclear power plants

For decades, the nuclear industry has known of the above described problems and, in consequence, has acquired the appropriate tools to be able to evaluate in advance the degradation of the vessel steel properties. Surveillance programmes are the tools currently employed to perform the follow-up of the evolution of the RPV steel properties throughout the operating lifetime of the reactor. These programmes consist of placing inside the reactor, from the beginning of plant operation, surveillance capsules containing specimens; these are made of a steel (base material, weld metal or heat affected zone material – see comments below regarding to this point) identical to the one which constitutes the vessel – along with flux wires necessary to estimate the neutron fluence and temperature gauges.

Surveillance capsules must be located within the reactor vessel, in the beltline, so that the specimen irradiation history duplicates as closely as possible, within the physical constraints of the system, the neutron spectrum, temperature history and maximum neutron fluence experienced by the reactor vessel. The fabrication history (austenitizing, quenching and tempering, and post-weld heat treatment) of the test materials will be fully representative of the fabrication history of the materials in the beltline of the reactor vessel. Because the capsules are located closer to the core than the inner vessel wall, the specimens suffer accelerated neutron fluence doses and, therefore, provide advanced information about the evolution of the embrittlement process. This phenomenon is described through the surveillance capsule lead factor (i.e., the ratio of the neutron fluence rate at the specimens in a surveillance capsule to the neutron fluence rate at the inner vessel wall, E > 1 MeV). It is recommended that the surveillance capsule lead factor be greater than one and less than or equal to three. This range of lead factors has been selected to minimise the calculation uncertainties in extrapolating the surveillance measurements from the specimens to the reactor vessel wall.

Periodically, taking advantage of the plant refuelling outages, specimen removal is carried out (along with flux wires and temperature gauges) from inside the capsules. These specimens are subsequently tested, thus providing realistic information about the evolution of the material properties throughout the plant operating period. This is a method for assessing the material degradation caused by neutron irradiation. The issues related to the surveillance programme design along with the rules for the interpretation of the results of such programmes are

regulated in several ASTM standards (ASTM E 185, 2002; ASTM E 2215, 2002; ASTM E 853, 2001). Next, the minimum requirements considered in these standards for the design of a surveillance programme for monitoring the radiation-induced changes in the mechanical properties of ferritic materials in the beltline region are summarised; it is worth noting, for the sake of clarity, that the following list of characteristics is mainly focused on those aspects related with the material's behaviour; other issues, as the measurement of neutron fluence through dosimetry or the measurement of temperatures, are not considered of relevance for the purposes of the present contribution.

- It is recommended to obtain the baseline (i.e., unirradiated) data during the design of the surveillance programme. Alternatively, the specimens may be placed in secured storage for testing at the time the first surveillance capsule is withdrawn for evaluation.
- A minimum test programme shall include material taken from the following locations: (1) base metal from each plate or forging used in the beltline, and (2) each weld metal made with the same heat of weld wire and lot of flux and by the same welding procedure as that used for each of the beltline welds. The base metal used to fabricate the weld shall be one of the base metals included in the surveillance programme.
- Charpy V-notch specimens (corresponding to the Type A specimen described in Test Method E 23 ASTM E 23, 2001)) shall be used. Tension specimens of the type, size, and shape described in Test Method E 8 (ASTM E 8, 2009) are recommended. Fracture toughness test specimens shall be consistent with the guidelines provided in Test Methods E 1820 (ASTM E 1820, 2001) and E 1921 (ASTM E 1921, 2009).
- A minimum of 15 Charpy specimens shall be tested to establish a full transition temperature and upper- shelf curve for each material. It is recommended that at least six tension test specimens be provided to establish the unirradiated tensile properties for both the base metal and the weld metal. A minimum of two specimens at room temperature and two specimens at reactor vessel beltline operating temperature should be tested. The remainder of the specimens may be tested at intermediate temperatures as needed to define the effects of temperature on the tensile properties. It is recommended as well that a minimum of 8 fracture toughness specimens be tested to establish the reference temperature, T_0, through Test Method E 1921 (ASTM E 1921, 2009) (see Section 5 for details about this issue) for the limiting beltline material and/or other fracture toughness tests be performed following Test Method E 1820 (ASTM E 1820, 2001).
- The first step in the implementation of a surveillance programme is the estimation of the peak EOL vessel inside surface fluence (EOL ID) and the peak EOL vessel 1/4-T (T being the thickness of the vessel) fluence[3] (identified as EOL 1/4-T). This data allows the transition temperature shift, ΔRT_{NDT}, to be estimated. The reference temperature shift can be determined from the relationship found in Guide E 900 (ASTM E 900, 2002) (for details and alternative procedures, see Section 4 in this chapter).

[3] The neutron fluence at any depth in the vessel wall attenuates as a consequence of the s cattering processes that occur between the neutrons and the crystal structure of the steel; these scattering processes are, in turn, responsible of the damage mechanisms leading to material's embrittlement. There are different models to predict this phenomenon, see, for example (RG 1.99 (2), 1988; ASTM E 900, 2002).

- A sufficient number of surveillance capsules shall be provided to monitor the effects of neutron irradiation on the reactor vessel throughout its operating lifetime. The basis for the number of capsules to be installed at the beginning of life is the predicted transition temperature shift, ΔRT_{NDT}.
- The minimum number of tests on irradiated specimens (base and weld materials) for each irradiation exposure set (capsule) shall be as follows: 15 Charpy specimens, 3 tension specimens and 8 fracture specimens (only fracture toughness specimens from the limiting material are required. It is suggested that a greater quantity of specimens be included in the irradiation capsules whenever possible).
- The capsule withdrawal schedule should permit periodic monitoring of long-time irradiation effects. The suggested withdrawal schedule is provided in Table 1. At a minimum, the surveillance programme should consist of two priority 1 capsules (see Table 1) including samples of all irradiated materials (if there is no weld in the beltline, then there is no requirement to include weld material in the surveillance programme. Moreover, experience has shown that it is no longer necessary to include the heat-affected zone material in the surveillance programme). The specimens in these capsules must be tested with target fluences corresponding to EOL-ID and EOL 1/4-T, respectively. If the estimated EOL ΔRT_{NDT} is less than 56°C (100°F), the vessel embrittlement may be adequately monitored by testing only these two priority 1 capsules. The priority 2 capsule is required for plants that exceed this limit; this capsule should be scheduled for withdrawal early in the vessel life to verify the initial predictions of the surveillance material response to the actual radiation environment. Past practice for PWRs has shown that a target fluence of $5.3\ 10^{18}$ n/cm^2 (E>1 MeV) is adequate; a lower target fluence is appropriate for BWRs. The priority 3 capsule is included to better define the embrittlement trend in materials with large projected EOL ID ΔRT_{NDT} values greater than 111°C (200°F).

Sequence	**Target fluence**	**Priority**
First	PWRs: $5 \cdot 10^{18}$ n/cm^2 (E>1 MeV)	2 (required if ΔRT_{NDT} > 56°C)
Second	EOL ¼-T	1 (required for all materials)
Third	EOL ID	1 (required for all materials)
Fourth	(EOL ¼-T – 1st Capsule)/2	3 (required if ΔRT_{NDT} > 111°C)
Subsequent	Supplemental Evaluations	Not required

Table 1. Suggested Withdrawal Schedule

- Additional capsules may be needed to monitor the effect of a major core change or annealing of the vessel, or to provide supplementary toughness data for evaluating a flaw in the beltline or for monitoring the vessel over the extended lifetime. Supplementary testing may also be based on reconstitution of previously tested specimens following Guide E 1253 (ASTM E 1253, 1999).

4. The physical mechanisms of embrittlement, predictive models

The recognition of the importance of embrittlement processes led to the establishment of surveillance programmes as a strategy for monitoring the material properties of the vessel.

The limited data available from each specific monitoring programme is insufficient to predict the shift in the transition temperature. To get round this problem, the most widely used strategy is to turn to predictive formulas based on the statistical analysis of large populations of data from the monitoring programmes of several plants (Eason et al., 1998). The physical models developed in recent times include temperature, T_i, fluence and the neutron flux with energies above 1 MeV, Φ and φ, respectively, the contents of copper, nickel and phosphorus, and the form presented by the steel (plate, forging or welding). Recent studies have revealed minor influences associated with the presence of manganese and the heat treatment applied (Odette & Lucas, 1998). The phenomenon of the recovery of properties after an annealing treatment along with the subsequent embrittlement are scarcely documented at present and require further study. Given the large number of variables involved in the process, it is not feasible to elaborate predictive expressions of a purely empirical nature, particularly when these are to be extrapolated to high neutron doses. Fortunately, knowledge of the physical processes involved in the embrittlement enables reliable expressions to be proposed for estimating the shift in the transition temperature in terms of these variables.

It should be noted that all of the predictive equations for embrittlement used in the nuclear sector and gathered in the codes and regulations of each country are based on the results of the Charpy impact test. Most of these expressions were obtained for manganese-molybdenum-nickel steels, (typically used in the United States, Western Europe, Japan and Korea) or chromium-molybdenum-vanadium and chromium-molybdenum-nickel-vanadium steels (used in the WWER reactors of Eastern Europe). These models express the embrittlement as a function of the presence of chemical elements (such as copper, phosphorus and nickel) and the value of fast neutron fluence. Sometimes, the irradiation temperature and the nature of the steel (weld or base material) are taken into account. It is also worth mentioning that, for most of the regulations, the fast neutron fluence is defined as the portion of the neutron spectrum with E> 1 MeV, while in the Russian WWER reactors, the fraction with E> 0.5 MeV is used.

4.1 Physical mechanisms of embrittlement

The mechanism of embrittlement is primarily based on the hardening produced by nanometric features that develop as a consequence of the impact of the neutrons on the crystal structure. The main embrittlement processes include (Odette & Lucas, 2001):

- Generation of lattice defects in displacement cascades by high-energy recoil atoms from neutron scattering and reactions. These primary defects are in the form of single and small clusters of vacancies and self-interstitials (Frenkel defects).
- Diffusion of primary defects also leading to enhanced solute diffusion and formation of nanoscale defect-solute cluster complexes, solute clusters and distinct phases, mainly copper-rich precipitates (CRPs).
- The presence of these nanofeatures can arrest the advance of the dislocations.
- As a consequence, this hardening process leads to the shift of the transition temperature detected experimentally, thus facilitating the material fracture through cleavages.

Obtaining a predictive expression to determine the transition temperature shift can be achieved by developing physical submodels of the above mentioned processes which depend on the metallurgical (Cu, Ni, P ,…) and irradiation (Φ, φ, T_i) variables (Eason et al., 1998; Odette & Lucas, 1998). The current understanding of the mechanisms of damage

production is essentially based on molecular dynamics (Stoller et al., 1997) and Monte Carlo numerical simulations (Wirth & Odette, 1999), combined with indirect experimental measurements. Thermodynamic-kinetic models are used to track the transport and fate of irradiation defects and solutes and to predict the number, size distribution and composition of the evolving nanofeatures. The characterization methods include: X-ray and neutron scattering, various types of electron microscopy, three-dimensional atom probe-field ion microscopy and positron annihilation spectroscopy. These researches allowed at establishing the following sequence (Odette & Lucas, 2001):

Neutrons create vacancies and interstitials, separated by some distance, by displacing atoms from their normal crystal lattice sites. The displacements are produced in cascades resulting from highly energetic primary recoiling atoms (PRA) generated by neutron scattering and reactions. The interaction of a high-energy neutron with an atomic nucleus results in significant energy transfer (R). For example, a 1 MeV neutron transfers up to about 70 keV to an iron PRA (Odette & Lucas, 2001). Part of the recoil energy is lost in the electronic cloud; hence, a lower kinetic energy is available for atomic collisions, $R_d < R$. The PRA kinetic energy is transferred by secondary, tertiary and n-subsequent generations of collision displacements, producing 2^n recoiling atoms at lower energies ($\sim R_d/2^n$). The process terminates when the kinetic energy of the nth-generation of recoils falls below that needed to cause additional displacements. On average, a PRA creates $R_d/2D$ displacements, where $D\sim0.05$ keV. Thus, a typical value R_d = 20 keV is able to produce a cascade of ~200 displacements. Closely spaced interstitials and vacancies quickly recombine and only about one third of the initial displacements survive. Typically, this leaves a vacancy-rich cascade core, surrounded by a shell of interstitials.

Most of interstitials quickly cluster to form small, disc-shaped features that are identical to small dislocation loops. Along with the interstitials, these loops are very mobile. Diffusion of interstitials and loops within the cascade region causes additional recombination prior to their rapid long-range migration (unless they are strongly trapped by other defects or solutes). Although they are less mobile than the interstitials, vacancies also eventually diffuse. Through a series of local jumps, the vacancies and solutes in the cascade quickly begin to evolve to lower energy configurations, forming small, three-dimensional clusters, while others leave the cascade region. The small clusters are unstable and can dissolve by vacancy emission. However, the small clusters also rapidly diffuse and coalesce with each other, forming larger nanovoids, which persist for much longer times. Solute atoms bind to the vacancies and segregate to clusters. The vacancy emission rate is lower from vacancy-solute cluster complexes. Small solute clusters remain after all the vacancy clusters have been finally dissolved.

The neutron dose or damage exposure can be expressed in terms of displacements-per-atom (dpa) which partially accounts for the effect of the neutron energy spectrum. For a typical RPV neutron spectrum, an end of life $\Phi = 3 \cdot 10^{19}$ n·cm^{-2} (E>1 MeV) produces about 0.1 dpa. However, most of the vacancies and interstitials eventually migrate and annihilate at sinks long distances from the cascade region. Thus long-range diffusion results in additional nanostructural evolution.

As a consequence of the above mentioned mechanisms three broad categories of nanofeatures can be distinguished:

- Copper rich or manganese-nickel rich precipitates (CRPs/MNPs).

- Unstable matrix defects (UMD) that form in cascades even in steels with low or no copper, but that anneal rapidly compared to typical low Φ irradiation times.
- Stable matrix features (SMF) that persist or grow under irradiation even in steels with low or no copper.
- Grain Boundary Segregations (GBSs) of embrittling elements as phosphorus (Server et al., 2001).

Whereas the contributions of SMDs and CRPs imply the creation of obstacles blocking the displacement of the dislocations, the GBSs can reduce the intergranular cohesion thus modifying the mechanisms and properties of fracture. As a consequence, a change from transgranular to intergranular fracture can ocurr. This phenomenon has not been observed in USA vessels but in Brtitish (Server et al., 2001).

4.2 Standard models for predicting the material embrittlement

This section includes a brief revision of the models proposed by the legislation currently in force that allows the embrittlement of the steel of the vessel to be determined throughout its operative lifetime. The vessels designed or manufactured in the USA are subjected to the regulations of that country. Specifically, the code of federal regulations 10CFR50 (10CFR50, 1986) is the reference document. According to this law, the Regulatory Guide 1.99 Rev. 2 (RG1.99 (2), 1988) establishes the procedure to determine the material embrittlement (see equation (6)); on the other hand, it points out that the plant surveillance programme must be prepared in agreement to that stipulated in the standard (ASTM E 185, 2002) which, in turn, refers to the standard (ASTM E 900, 2002) for the prediction of the material embrittlement. For these reasons, Section 4.2.1 is devoted to the first of the procedures mentioned (RG1.99 (2), 1988) while the last one is described in Section 4.2.2. Finally, it is worth noting that alternative models, which can be found in the scientific literature, have been developed: see for example (Odette and Lucas, 1998, 2001; Eason, Wright and Odette, 1998; Yamashita, Iwasaki, Dozaki, 2010); however, these models are considered to be beyond the scope of the present chapter.

4.2.1 The Regulatory Guide Rev. 2

The U.S. Regulatory Guide 1.99, Revision 2 (RG 1.99 (2), 1988) provides the current generic basis to evaluate ΔRT_{NDT} (which is considered to be equivalent to ΔT_{41J}) in terms of the copper and nickel contents of the steel and fast neutron fluence (E > 1 MeV) for weld and plate product forms. The model equations for (RG 1.99 (2), 1988) were statistically fitted to a small surveillance database (177 data points) on steels irradiated in surveillance capsules at flux levels somewhat higher than at the vessel wall. The (RG 1.99 (2), 1988) model was derived prior to 1985 and reflects the emerging, but far from complete, physical understanding of embrittlement mechanisms.

The actual (or adjusted) reference temperature RT_{NDT} for each material in the beltline region is given by expression (6). $RT_{NDT(U)}$ is the reference temperature for the unirradiated material as defined in (ASME Code). If measured values of initial RT_{NDT} for the material in question are not available, generic mean values for that class of material may be used if there are sufficient test results to establish a mean and standard deviation for the class. In the case of Linde 80 welds, a generic value $RT_{NDT(U)}=0^oF$ is accepted whereas for Linde 0091, 1092, 124 and ARCOS B-5, $RT_{NDT(U)}=-56^oF$. As stated in (RG 1.99 (2), 1988), ΔRT_{NDT} must be obtained by applying equation (7); the basis for equation (7) is contained in publications

(Guthrie, 1984) and (Odette et al., 1984). Both contributions used surveillance data from commercial power reactors. The bases for their regression correlations were different in that Odette made greater use of physical models of radiation embrittlement. The two papers contain similar recommendations: (1) separate correlation functions should be used for weld and base metal, (2) the function should be the product of a chemistry factor and a fluence factor, (3) the parameters in the chemistry factor should be the elements copper and nickel, and (4) the fluence factor should provide a trend curve slope of about 0.25 to 0.30 on log-log paper at 10^{19} n/cm^2 (E > 1 MeV), steeper at low fluences and flatter at high fluences.

$$\Delta RT_{NDT} = (CF) \cdot f^{(0.28 - 0.10 \log f)} \tag{7}$$

CF (°F) in equation (7) is the chemistry factor, a function of copper and nickel content. CF is tabulated in (RG 1.99 (2), 1988) for welds and base metal (plates and forgings). Linear interpolation is permitted. If there is no information available, 0.35% copper and 1.0% nickel should be assumed.
The neutron fluence at any depth x in the vessel wall, f(x) (10^{19} n/cm^2, E > 1 MeV), is determined following equation (8):

$$f(x) = f_{SURF} \cdot e^{-0.24 \cdot x} \tag{8}$$

where f_{surf} (10^{19} n/cm^2, E > 1 MeV) is the calculated value of the neutron fluence at the inner surface of the vessel, and x (in inches) is the depth in the vessel wall measured from the vessel inner surface. Alternatively, if dpa calculations are made as part of the fluence analysis, the ratio of dpa at the depth in question to dpa at the inner surface may be substituted for the exponential attenuation factor in equation (8).
The third term in (6), M, is the quantity, °F, that is to be added to obtain conservative, upper-bound values of the adjusted reference temperature. M is obtained through equation (9).

$$M = 2 \cdot \sqrt{\sigma_U^2 + \sigma_\Delta^2} \tag{9}$$

Here, σ_U is the standard deviation for the initial RT_{NDT}. If a measured value of initial RT_{NDT} for the material in question is available, σ_U is to be estimated from the precision of the test method. If not, generic mean values for that class of material are used. The standard deviation in ΔRT_{NDT}, σ_Δ, is 28°F for welds and 17°F for base metal, except that σ_Δ need not exceed 0.50 times the mean value of ΔRT_{NDT}.
Finally, when two or more credible surveillance data sets become available from the reactor in question, they may be used to determine the adjusted reference temperature of the beltline materials. In this case, if there is clear evidence that the copper or nickel content of the surveillance weld differs from that of the vessel weld, the measured values of ΔRT_{NDT} should be adjusted by multiplying them by the ratio of the chemistry factor for the vessel weld, $(CF)_V$, to that of the surveillance capsule weld, $(CF)_C$, see equation (10).

$$(\Delta RT_{NDT})_V = (\Delta RT_{NDT})_C \cdot \frac{(CF)_V}{(CF)_C} \tag{10}$$

Second, the surveillance data should be fitted using equation (7) to obtain the relationship of ΔRT_{NDT} to fluence. To do so, calculate the chemistry factor, CF, for the best fit by

multiplying each adjusted ΔRT_{NDT} by its corresponding fluence factor, summing the products, and dividing by the sum of the squares of the fluence factors. The resulting value of CF will give the relationship of ΔRT_{NDT} to fluence that fits the plant surveillance data in such a way as to minimise the sum of the squares of the errors.

4.2.2 Standard ASTM E 900 - 02

The purpose of ASTM Standard E 900-02 (ASTM E 900, 2002) is to establish an improved correlation in respect of that proposed by the Regulatory Guide (RG1.99 (2), 1988) which allowed the transition temperature shift, ΔT_{41J}, to be obtained as a function of the neutron fluence. The data base used in this case consisted of 600 data; thus, it is much more robust than that of the Regulatory Guide where only 177 data were available. Here, the expressions are not purely phenomenological but physically guided, taking into consideration when possible the nanofeatures participating in the process (described above). Another interesting feature is that the ASTM E 900-02 formulation is based on robust statistical techniques for the treatment of the large data set (non-linear, least-square regression analysis). It is worth noting that the origin of this procedure relies on the expressions proposed by (Eason et al., 1998). The procedure distinguishes between the three mechanisms described in Section 4.1:

- Stable matrix damage (SMD) associated with the presence of point defects and loop dislocations.
- Copper rich precipitates (CRPs).
- Grain Boundary Segregations (GBSs) of embrittling elements as phosphorus.

As there is no evidence of the influence of this last mechanism on USA vessels, the form of the correlation involves only the two major embrittlement terms: the SMD and the CRP; nevertheless, the influence of phosphorus is indirectly present through the SMD. The mean value of the transition temperature shift is calculated as follows (11):

$$Shift = SMD + CRP \tag{11}$$

The formulas for both terms must take into consideration the empirical reality; specifically, the CRP mechanism saturates with fluence while, on the contrary, the SMD damage process grows monotonically, apparently without limit.

Expression (12) represents the transition temperature shift, in ^{o}F, due to the SMD

$$SMD = 6.70 \cdot 10^{-18} \cdot e^{\frac{20730}{T_i + 460}} \cdot (\Phi)^{0.5076} \tag{12}$$

The main characteristics as well as the meaning of the variables are explained below:

- The influence of irradiation temperature, T_i, in ^{o}F, in the range 275-295 ^{o}C (527 – 563 ^{o}F, respectively) is modelled by means of an exponential function.
- The effect of the irradiation, which is expressed in the neutron fluence, Φ (n·cm^{-2}, E > 1 MeV), increases indefinitely, without saturation.
- There is no explicit dependence on the neutron flux, φ.

The second term in (11) which represents the shift in the transition temperature due to the CRP mechanism, in ^{o}F, responds to formula (13), together with expressions (14) and (15):

$$CRP = B \cdot \left(1 + 2.106 \cdot Ni^{1.173}\right) \cdot F(Cu) \cdot G(\Phi) \tag{13}$$

$$F(Cu) = (Cu - 0.072)^{0.577} \tag{14}$$

where Cu≤0.305 in general and Cu≤0.25 for Linde 80 and 0091.

$$G(\Phi) = \frac{1}{2} + \frac{1}{2} \cdot Tanh\left[\frac{\log(\Phi) - 18.24}{1.052}\right] \tag{15}$$

The relevant features in (13-15) and the meaning of the variables is explained in the following:

- The coefficient B takes different values depending on the material (B=234, for welds; B=128 for forgings; B=208 for combustion engineering plates; B=156 for other plates).
- There is a strong influence of the nickel content, Ni (expressed in % wt.).
- The influence of Cu is taken into consideration through F(Cu) (14); this effect occurs only for Cu>0.072% and saturates for Cu>0.305% (0.25% for Linde 80 and 0091)
- The irradiation temperature is not considered to play a role.
- The neutron fluence Φ influence is represented through the term G(Φ) where the saturation is modelled as a hyperbolic tangent function.

Moreover, a term corresponding to the standard deviation σ_{TTS}=22ºF must be considered. It takes into consideration the uncertainties in the input data and in the preparation of the model.

5. The Master Curve model for the description of the fracture toughness in the DBT region

5.1 Description of the model

Advances in fracture mechanics technology have made it possible to improve the semi-empirical indirect methodology described above, currently in force to describe the fracture toughness in the DBT region in different aspects. First, the development of EPFM allows fracture toughness values to be determined using much smaller specimens and utilising J integral techniques, that is, measuring values of K_{Jc} instead of K_{Ic}. Moreover, the analytical techniques for structural integrity assessment can now be expressed in terms of EPFM. The first issue, which is considered relevant for the contents of this chapter, is described in the present section; indeed, the scope is explicitly focused on the Master Curve (MC) approach to describe the fracture toughness of vessel steels in the DBT region.

The MC model, originally proposed by Wallin (Wallin, 1984; Wallin et al. 1984; Wallin, 1989; Wallin, 1995), provides a reliable tool based on a direct characterisation of the fracture toughness in the DBT region. This approach is a consequence of the developments in EPFM together with an increased understanding of the micro-mechanisms of cleavage fracture. Valiente et al. (Valiente et al., 2005) have briefly but comprehensively reviewed the previous contributions made to understand cleavage in a ferritic matrix that leads to the MC approach. The basic MC method for analysis of brittle fracture test results is defined in the standard ASTM E 1921 (ASTM E 1921, 2009). The mathematical and empirical details of the procedure are available in (Merkle et al., 1998; Ferreño, 2008; Ferreño et al. 2009). The main features and advantages of the method are hereafter summarised:

- MC assumes that cleavage fracture in non austenitic steels is triggered by the presence of particles close to the crack tip. Therefore, fracture is mainly an initiation dependent process. As a consequence, fracture is governed by weakest link statistics which follows a three parameter Weibull distribution. For small-scale yielding conditions, therefore using EPFM, the cumulative failure probability, P_f, is given by equation (16):

$$P_f = 1 - e^{-\frac{B}{B_0}\cdot\left(\frac{K_{Jc}-K_{\min}}{K_0-K_{\min}}\right)^m} \tag{16}$$

where K_{Jc} is the fracture toughness for the selected failure probability, P_f , K_0 is a characteristic fracture toughness corresponding to 63.2% cumulative failure probability, B is the specimen thickness and B_0 a reference specimen thickness, B_0 = 25.4 mm. The experimental data allows the Weibull exponent, m=4, to be fixed (Merkle et al., 1998) and the minimum value of fracture toughness for the probability density function, K_{min} = 20 MPa·m$^{1/2}$. Therefore, only K_0 must be estimated from the empirical available data.

- The dependence between K_0 (in MPa·m$^{1/2}$) and temperature (^{o}C) for cleavage fracture toughness is assumed to be (17):

$$K_0 = 31 + 77\cdot e^{0.019\cdot(T-T_0)} \tag{17}$$

where T_0 is the so-called MC reference temperature; it corresponds to the temperature where the median fracture toughness for a 25 mm thickness specimen (1T, according to ASTM terminology) has the value 100 MPa·m$^{1/2}$.

- One of the main advantages of the method is that it allows data from different size specimens to be compared. As thickness increases, the toughness is reduced, due to the higher probability of finding a critical particle for the applied load. The ASTM standard (ASTM E 1921, 2009) provides expressions to relate the fracture toughness for specimens of different thicknesses. Equation (16) can be re-written considering the same failure cumulative probability, P_f, for two specimens of different thickness, namely B_1 and B_2, thus leading to expression (18):

$$K_{J_c,2} = K_{\min} + \left(K_{J_c,1} - K_{\min}\right)\cdot\left(\frac{B_2}{B_1}\right)^{\frac{1}{4}} \tag{18}$$

- The distribution fitting procedure involves finding the optimum value of T_0 for a particular set of data. For this task, all data are thickness adjusted to the reference specimen thickness B_0 = 25.4 mm using equation (18). The procedure can be applied either to a single test temperature or to a transition curve data, T_i being the generic temperature of the different tests. In the latter approach (the former is just a particular case) T_0 is estimated from the size adjusted K_{JC} data ($K_{JC,1T}$) using a multi-temperature maximum likelihood expression (see equation (19)). To estimate the reference temperature, T_0, a previous censoring of the non size-adjusted data must be applied. Fracture toughness data that are greater than the validity limit given by equation (20), as defined in (ASTM E 1921, 2009), are reduced to the validity limit, $K_{Jc(lim)}$ and treated

as censored values in the subsequent estimation stage (δ_i =0 in expression (19)). This condition is imposed to guarantee high constraint conditions in the crack front during the fracture process.

$$\sum_{i=1}^{N}\delta_i\cdot\frac{e^{0.019\cdot(T_i-T_0)}}{11+77\cdot e^{0.019\cdot(T_i-T_o)}}-\sum_{i=1}^{N}\frac{\left(K_{Jc,i}-20\right)^4\cdot e^{0.019\cdot(T_i-T_0)}}{\left[11+77\cdot e^{0.019\cdot(T_i-T_0)}\right]^5}=0 \tag{19}$$

$$K_{Jc(\lim)}=\sqrt{\frac{E\cdot\sigma_Y\cdot b_0}{30\cdot\left(1-\nu^2\right)}} \tag{20}$$

In equation (20) σ_y is the yield strength at test temperature, E is the Young's modulus, b_0 the initial ligament and ν is the Poisson's ratio. It must be stressed that factor 30 in equation (20) is currently under discussion (Merkle et al., 1998) and that, for instance, the ASTM E1820-01 Standard (ASTM E 1820, 2001) imposes a more demanding limit with a factor 50 or 100 depending on the nature of the steel.

- The standard deviation in the estimate of T_0, expressed in ^oC, is given by (21):

$$\sigma_{T_0}=\frac{\beta}{\sqrt{r}} \tag{21}$$

where r represents the total number of valid specimens (not censored results) used to establish T_0. The values of the factor β are provided in (ASTM E 1921, 2009).

- The statistical analysis can be reliably performed even with a small number of fracture toughness tests (usually between 6 and 10 specimens). Moreover, as an EPFM approach is used, the specimen size requirement, equation (20), is much less demanding than that of the LEFM (ASTM E 399, 2009). These remarks are of great relevance in nuclear reactor surveillance programmes where the amount of material available is usually very limited and consists of small size samples (Charpy specimens).
- By rearranging equations (16) and (17) it is possible to obtain expression (22) which provides an estimate of K_{Jc} for a given cumulative failure probability, P_f, once T_0 has been determined. In this way, the confidence bounds of the distribution (usually taking P_f = 0.01 or 0.05 for the lower bound and 0.95 or 0.99 for the upper bound) can be obtained. As a particular case, the expression for the median fracture toughness (P_f = 0.5) (see Equation (23)) is determined.

$$K_{J_c,P_f}=K_{\min}+\left[-\ln\left(1-P_f\right)\right]^{0.25}\cdot\left[11+77\cdot e^{0.019\cdot(T-T_0)}\right] \tag{22}$$

$$K_{J_c(med)}=30+70\cdot e^{0.019(T-T_0)} \tag{23}$$

- Finally, any test that does not fulfil the requirement for crack front straightness or that terminates in cleavage after more than a limit of slow-stable crack growth will also be regarded as invalid.

5.2 Open issues concerning the Master Curve approach

The MC has become a mature tool for characterising the fracture toughness of ferritic steels in the DBT region. Considerable empirical evidence provides testament to the robustness of the MC procedure. One of the main advantages of this procedure relies on the possibility of assessing the state of a RPV vessel by direct measurement of fracture toughness rather than through the use of the currently accepted correlative approaches, based on Charpy tests. The procedure currently accepted to assess the steel neutron embrittlement partially incorporates the MC reference temperature concept, T_0; in this sense, to enable the use of the MC methodology without completely modifying the structure of the ASME code the approach stated in code cases N-629 (ASME CC 629, 1999) and N-631 (ASME CC 631, 1999) was adopted. It consists of defining a new index temperature, RT_{T0}, for the K_{Ic} and K_{IR} ASME curves (4, 5), as an alternative to RT_{NDT}. The definition of RT_{T0} is given in equation (24). This value of RT_{T0} is set, see (VanDerSluys et al., 2001), by imposing that the ASME K_{Ic} curve indexed with RT_{T0} in place of RT_{NDT} will bound the majority of the actual material fracture toughness data. In this sense, RT_{T0} was set such that the corresponding ASME K_{Ic} curve falls below the MC 95% confidence bound for at least 95% of the data generated with 1T specimens.

$$RT_{T_0} = T_0 + 19.4\ ^{o}C \tag{24}$$

Evidently, this approach, currently in force, is merely a compromise solution that attempts to fit the new concepts into the old structure. Apart from this practical aspect, several other open issues remain concerning the application of the MC as well as theoretical aspects. The following issues must be emphasised:

- There is no experimental data that allows the MC to be used in special applications such as irradiated materials with high neutron fluence, materials susceptible to intergranular fracture or materials showing exceptional lower-shelf or transition behaviour. Indeed, the main feature of the MC method consisting of assuming that the dependence of the fracture toughness of a material on temperature in the transition range is not sensitive to characteristics such as the mechanical properties and the microstructure is purely speculative for the cases mentioned above.
- The published literature shows that the PCCv (Pre Cracked Charpy-V Notched) specimen analysed using the ASTM Standard E 1921 (ASTM E 1921, 2009) generally shows a reference temperature ~10 °C lower than the CT (compact tension) specimen. Compared with the inherent scatter in the transition temperature, this difference is small. However, it has been observed in many materials. Although different hypotheses were proposed a decade ago in order to explain this fact, the current consensus in the scientific community is that this difference is motivated by the different level of constraint in single edge notch bend and CT geometries.
- The above issue is a particular case of the general question of how crack-tip constraint effects (stress tri-axiality in the vicinity of the crack tip) can be described. In fracture mechanics, it is well known that crack-tip constraint can be influenced by loading (out of plane or multi-axial loading) or by the crack shape and crack depth to ligament ratio. Nevertheless, to date, there is no agreement about how to manage crack tip constraint in the practical application of structures and components containing postulated or real cracks and made of ferritic steel.

- The MC approach procedure standardised in ASTM E1921 (ASTM E 1921, 2009) is defined for quasi-static loading conditions. However, the extension of the MC method to dynamic testing is still under discussion although a great effort has been made over the last decade to qualify the method for dynamic loading conditions and to use it for structural purposes.

This list of questions currently under discussion reveals that, although the MC methodology is increasingly being recognised as an attractive alternative for describing the fracture toughness of ferritic steels in the DBT region, further research needs to be done in order to properly deal with the open issues mentioned above.

6. Conclusions

The purpose of this chapter was to provide the readers with an introductory text, self-contained insofar as possible, concerning the current state of the art in the process of embrittlement that takes place in nuclear vessel steels, paying particular attention to the ductile to brittle transition region. It was the purpose of the authors to introduce the topics in a logical sequence in an attempt to explain the scientific and historical arguments that justify the different methods and tools currently available. A phenomenological and scientific description of the causes and consequences of material embrittlement was presented. An explanatory description of the characterisation tools that are available for the nuclear facilities -implemented in their surveillance programmes- to determine the evolution of the fracture toughness of the vessel steel throughout the operative lifetime of the plant, emphasising their advantages and limitations, was also included. This leads, in a natural way, to the Master Curve methodology, as an alternative procedure for obtaining, in the context of Elastic-Plastic Fracture Mechanics, the material fracture toughness; as stressed in the text, this procedure offers many advantages and few limitations, which is why it is widely used at present in a great number of ambitious scientific research projects. It is the opinion of the authors that all of the evidence available points to the fact that the Master Curve approach is set to become an indispensable ingredient in the future of surveillance programmes.

7. References

ASTM E 23–01. Standard Test Methods for Notched Bar Impact Testing of Metallic Materials.

Anderson, T.L. (1995). *Fracture Mechanics. Fundamentals and Applications* (2nd Ed.), CRC Press, ISBN-13: 9780849342608, USA.

Griffith, A.A. (1920). The Phenomena of Rupture and Flow in Solids, *Philosophical Transactions*, Series A, Vol. 221, pp. 163-198.

Inglis, C.E. (1913). Stresses in a Plate Due to the Presence of Cracks and Sharp Corners, *Transactions of the Institute of Naval Architects,* Vol. 55, pp. 219-241.

Irwin, G.R. (1956). Onset of Fast Crack Propagation in High Strength Steel and Aluminium Alloys, *Sagamore Research Conference Proceedings*, Vol. 2, pp. 289-305.

Irwin, G.R. (1957). Analysis of Stresses and Strains near the End of a Crack Traversing a Plate. *Journal of Applied Mechanics*, Vol. 24, pp. 361-364.

Westergaard, H.M. (1939). Bearing Pressures and Cracks. *Journal of Applied Mechanics*, Vol. 6, pp. 49-53.

ASTM E399 - 09e2. Standard Test Method for Linear-Elastic Plane-Strain Fracture Toughness K_{Ic} of Metallic Materials.

Wells, A.A. (1961). Unstable Crack Propagation in Metals: Cleavage and Fast Fracture, *Proceedings of the Crack Propagation Symposium*, Vol. I, Paper 84.

Rice J.R. (1968). A Path Independent Integral and the Approximate Analysis of Strains Concentration by Notches and Cracks. *Journal of Applied Mechanics*, Vol. 35, pp. 379-386.

ASME Boiler and Pressure Vessel Code. American Society of Mechanical Engineers, New York.

ASTM E 208 – 95a. Standard Test Method for Conducting Drop-Weight Test to Determine Nil-Ductility Transition Temperature of Ferritic Steels.

Rules and Regulations Title 10 Code of Federal Regulations (CFR) Part 50.61, Appendix G (1986). Fracture Toughness Requirements for Protection Against Pressurized Thermal Shock Events, Washington, D.C., U.S. Government Printing Office, U.S. Nuclear Regulatory Commission.

Regulatory Guide 1.99-Rev 2 (1988). Radiation Embrittlement of Reactor Pressure Vessel Materials, U.S. Government Printing Office, U.S. Nuclear Regulatory Commission, NRC, Washington D.C..

ASTM E 185 – 02. Standard Practice for Design of Surveillance Programs for Light-Water Moderated Nuclear Power Reactor Vessels.

ASTM E 2215 – 02. Standard Practice for Evaluation of Surveillance Capsules from Light-Water Moderated Nuclear Power Reactor Vessels.

ASTM E 853 – 01. Standard Practice for Analysis and Interpretation of Light-Water Reactor Surveillance Results, E706(IA).

ASTM E8 / E8M - 09 Standard Test Methods for Tension Testing of Metallic Materials

ASTM E1820-01. Standard Test Method for Measurement of Fracture Toughness.

ASTM E1921-09-a. Test Method for the Determination of Reference Temperature T0 for Ferritic Steels in the Transition Range.

ASTM E 900-02. Standard Guide for Predicting Radiation-Induced Transition Temperature Shift in Reactor Vessel Materials, E706 (IIF).

ASTM 1253-99. Standard Guide for Reconstitution of Irradiated Charpy Specimens.

Eason, E.D., Wright, J.E., Odette, G.R. (1998). *Improved Embrittlement Correlations for Reactor Pressure Vessel Steels*. NUREG/CR-6551, MCS 970501, U.S. Government Printing Office, U.S. Nuclear Regulatory Commission, NRC, Washington D.C..

Odette, G.R., Lucas, G.E. (1998). *Recent Progress in Understanding Reactor Pressure Vessel Embrittlement*. Radiation Effects and Defects in Solids, No. 144, pp. 189-231.

Odette, G.R., Lucas, G.E. (2001). *Embrittlement of Nuclear Reactor Pressure Vessels*. JOM, Vol. 53, No. 7, pp. 18-22.

Stoller, R.E., Odette, G.R., Wirth, B.D. (1997). *Primary Damage Formation in BCC Iron*. Journal of Nuclear Materials, No. 251, pp 49-60.

Wirth, B.D., Odette, G.R. (1999). *Kinetic Lattice Monte Carlo Simulations of Cascade Aging in Dilute Iron-Copper Alloys*. Microstructural Processes in Irradiated Materials, MRS Symp. Proc. 540, ed. J. Zinkle el al. , pp. 637-642.

Server, W., English, C., Naiman, D., Rosinski, S. (2001). *Charpy Embrittlement Correlations-Status of Combined Mechanistic and Statistical Bases for U.S. RPV Steels (MRP-45),* PWR Materials Reliability Programme (PWRMRP), 1000705, EPRI Final Report.

Yamashita, N., Iwasaki, M., Dozaki, K. (2010). *Industry Practice for the Neutron Irradiation Embrittlement of Reactor Pressure Vessels in Japan*. J. Eng. Gas Turbines Power, Vol. 132, Issue 10.

Guthrie, G.L. (1984). *Charpy Trend Curves Based on 177 PWR Data Points*. LWR Pressure Vessel Surveillance Dosimetry Improved Program. NUREG/CR-3391.

Odette, G.R. et al. (1984). *Physically Based Regression Correlations of Embrittlement Data from Reactor Pressure Vessel Surveillance Programs*. Electric Power Research Institute, NP-3319.

Wallin, K. (1984). The scatter in K_{IC}-results. *Engineering Fracture Mechanics*, Vol. 19, Issue 6, pp. 1085-1093.

Wallin K, Saario T & Törrönen K (1984). Statistical model for carbide induced brittle fracture in steel. *Metal Sci* , Vol. 18, pp. 13–16.

Wallin K. (1989). A simple theoretical Charpy V–K_{Ic} correlation for irradiation embrittlement. *Innovative approaches to irradiation damage and fracture analysis, PVP 170.* ASME. pp. 93–100.

Wallin, K. (1995). The size effect in K_{IC} results. *Engineering Fracture Mechanics*, Vol. 22, Issue 1, pp. 149-163.

Valiente, A. Ruiz J. & Elices, M. (2005). A probabilistic model for the pearlite-induced cleavage of a plain carbon structural steel. *Engineering Fracture Mechanics*, Vol. 72, Issue 5, pp. 709-728.

Merkle J., Wallin K. & McCabe D. (1998). Technical basis for an ASTM Standard on determining the reference temperature, T_0, for ferritic steels in the transition range. Washington, DC: Oak Ridge National Laboratory. NUREG/CR-5504, ORNL/TM-13631.

Ferreño, D (2008). Integridad estructural de vasijas nucleares en base a la curva patrón obtenida mediante probetas reconstruidas. PhD thesis, University of Cantabria, Spain.

Ferreño D., Scibetta M., Gorrochategui I., Lacalle R., van Walle, E. & Gutiérrez- Solana, F. (2009). *Engineering Fracture Mechanics*, Vol. 76, Issue 16, pp. 2495-2511.

ASME Boiler and Pressure Vessel Code–Code Case N-629. Use of fracture toughness test data to establish reference temperature for pressure retaining materials, Section XI, Division 1, 1999.

ASME Boiler and Pressure Vessel Code–Code Case N-631. Use of fracture toughness test data to establish reference temperature for pressure retaining materials other than bolting for class 1 vessels, Section III, Division 1, 1999.

VanDerSluys, W.A., Hoffmann, C.L., Yoon, K.K., Server, W.L., Lott, R.G., Rosinski, S., Kirk, M.T., Byrne, S., Kim, C.C. (2001). Fracture toughness master curve development: Application of master curve fracture toughness methodology for ferritic steels, Bulletin 458, Welding Research Council, WRC, New York.

13

Corrosion Monitoring of the Steam Generators of V-th and VI-th Energy Blocks of Nuclear Power Plant "Kozloduy"

Nikolai Boshkov, Georgi Raichevski, Katja Minkova and Penjo Penev
Institute of Physical Chemistry, Bulgarian Academy of Sciences
Nuclear Power Plant "Kozloduy"
Bulgaria

1. Introduction

Corrosion resistance and erosion wear of the metal equipment (body, pipe bundles, steam generators etc.) are of great importance for the operating time limit of the nuclear power plants (NPP) bearing in mind their extremely heavy technological and water-chemical regime (WCR). The horizontally placed steam generators (SGs) used in the Energy Blocks of NPP "Kozloduy" work at exploitation conditions leading to corrosion and erosion provoked processes (Andreeva M. et. al., 2008). Similar is the situation in the greater part of the modern power plants (R.W. Staehle & J.A. Gorman, 2003; R.W. Staehle & J.A. Gorman, 2004; R.W. Staehle & J.A. Gorman, 2004). This fact predetermines the necessity of very careful daily and also periodical monitoring concerning the real operating state of the equipment during its exploitation time (Sviridenko, I., 2008; Hadavi, S.M.H., 2008; Viehrig, H.-W., et. al., 2006; Slugen, V., et. al., 2005). The evaluation and explanation of the obtained results for the electrochemical-corrosion parameters will ensure a realistic base for a precise prognosis and trouble-free operation of the installations (J. Congleton et all, 1985; Raichevski G. et al.,. 2007; Hojna A. et. al, 2007; Kaczorowski D. et. al., 2006).

Markedly dangerous negative influences leading to technical failures and damages have the conglomerated corrosion products (CPs) the latter appearing as a result of the WCR mainly in the second contour of NPP. Considerable part of them enters the condense-feeding part of SGs and deposit as loose layers on the pipe surface or falls as middling slime on the bottom changing in such a way the overall electrochemical and hydrodynamic situation in the internal volume during the operation procedures.

In general, CPs presenting in the steam generators consist of iron and copper oxides and hydroxides. If their adhesion to the surface is enough strong (for example, like in the case of magnetite – Fe_3O_4) it can be expected that a passive layer will appear on the surface of the equipment and in such a way this part of the installation will be additionally protected. If CPs are crumbly the operative situation in the SGs will be worsen since these deposits often lead to accelerated formation and development of local corrosion phenomena – holes, pits, cracks, including appearance of stress corrosion cracking process etc., the latter being one of the most dangerous events due to its practical invisibility up to the moment of the sudden physical-mechanical destroying (Zubchenko, A.S., et al., 2004; Slugen, V., et al., 2002; Lunin, G.L., et al., 1997; N.D. Budiansky, et al., 2005).

The presence of sulphate, copper and especially chloride ions additionally contributes to the corrosion damages resulting in local ruptures with different sizes. When CPs have loose structure and high porosity the concentration of Cl^- ions in their volume increases due to the flowing capillary processes. This effect is stronger expressed in the contact places of the pipes with the distance grid as well as at the lowest placed rows of the pipe bundles in SGs.
In general, the aim of the present investigation is to determine and evaluate the role of different aggressive components presenting in the technological medium on the corrosion behavior of austenitic stainless steel of the type 18Cr10NiTi which material is used to produce the pipe bundles in the SGs of the second contour of Energy blocks No. V and No. VI of NPP "Kozloduy". Some of the results obtained by polarization technique are compared with the experimental data of low-alloyed steel.

2. Experimental

2.1 Samples

Two different steel sample types with dimensions 1 x 1 cm (working area of 2 cm^2) are used:

- High-alloyed (HAS) austenitic stainless steel 18Cr10NiTi which is applied as a main construction material for the pipe bundles in the SGs of NPP "Kozloduy" (composition: Cr – 18 wt.%; Ni – 10 wt.%; Ti – about 1 wt.%; balance – Fe).
- Low-alloyed (LAS) steel 38GN2MFA (composition: Cr, Ni, Mn and Si – about 6 wt.% total; balance - Fe) used in general as a construction steel and as a comparative sample.

2.2 Corrosion medium

The corrosion tests are carried out in a model corrosion medium (MCM) consisting of trebly distilled demineralised water (specific electric conductivity ≤ 1 μS/cm), hydrazine (25 μg/L) and in definite cases selected anions or cations (Cl^-, SO_4^{2-}, Na^+, Cu^+, Cu^{2+} - both latter presented in the text as Cu_{total}). The pH value is about 9 aiming the modelling of conditions maximal close to the exploitation ones. Special attention is paid to the effectiveness of the added compound that inhibits the development of the damaging processes on the surface and in the depth – monoethanolamine (MEA) at concentration of 2 mg/L.

2.3 Sample characterization

2.3.1 Potentiodynamic (PD) and potentiostatic (PS) polarization curves

Potentiodynamic (PD) polarization curves (scan rate of 1 mV/s.) are carried out at two selected operating temperatures (35 °C and 95 °C, respectively) in a three-electrode experimental glass cell using "VersaStat" (AMETEK PAR) device. The cell has a volume of 300 cm^3 and Luggin capillary for minimizing the Ohmic resistance. Platinum plate serves as a counter electrode and saturated calomel electrode (SCE) - as a reference electrode.
The investigations of the local corrosion phenomena are realized in a special constructed test electrode the latter allowing simulation of the processes taking place in cracks with different widths. This electrode consists of ceramic body, separators (mica lamellas), titanium linings etc. giving the opportunity to work also at high temperatures and pressures aiming the modelling (approaching) of the real operating conditions in the SGs.
In order to receive also several results at conditions maximal close to the exploitation ones an autoclave system is used for some of the electrochemical and corrosion measurements. The autoclave volume is approximately 1,5L and its maximum capacity is to work at temperatures of 350 °C and pressures of 150 atm. Aiming better safety the experiments are

carried out in the interval of about 220 - 230 °C and pressure of 150 atm. Three-electrode electrochemical Teflon cell with working volume of 300 cm^3 and Ag/AgCl reference electrode assembly is mounted in the autoclave unit. The polarization studies in that case are carried out in a potentiostatic way contrary to the anodic and cathodic polarization curves realized at the abovementioned temperatures and at atmospheric pressure.

2.3.2 Scanning electron microscopy

SEM images and EDX analysis complete the results of the investigation giving more information about the surface morphology of the tested samples and of the collected CPs. A scanning electron microscope JEOL JSM-5300, Japan (equipped with EDX device) is used for registration of the morphology changes on the sample surfaces of HAS after the treatment in the medium. This device is used also for quantitative analysis of the element composition in selected special places – pits, cracks etc. The results obtained by this method ensure the possibility to present more realistic prognosis about the future exploitation of the SGs.

2.3.3 Moessbauer spectroscopy

Moessbauer spectroscopy completes phase analysis of the chemical compounds for characterization of the type and quantity of the iron compounds presenting on the sample. This method is based on the high sensitivity of Fe in relation to the resonance adsorption toward the beams of the radiation source applied. In the present investigations the absorption variant of the method is used which allows to receive information from the whole thickness of the CPs. The separation of the phases is realized with high sensitivity accepting that the whole amount of all compounds presenting is 100%. Actually, Moessbayer spectra are carried out at room temperature using electromechanical spectrometer working at conditions of a constant acceleration with a $^{57}Co/Cr$ source (activity ≅ 60 mCi) and α-Fe probe as a standard. Experimentally obtained spectrum is additionally calculated using a special program based on the method of the least squares. The content of iron ions in every component is determined under the conditions of the presumption of identical Moessbayer-Lamb factor.

2.3.4 X-ray photoelectron spectroscopy

XPS analysis is carried out with ESCALAB MkII (VG Scientific) electron spectrometer characterizing with base pressure in the chamber up to 5×10^{-10} mbar (during the test 1×10^{-8} mbar) with MgKα X-ray source (excitation energy hν = 1253,6 eV).
Due to the weak signals for Ni2p and Cr2p as well as the presence of a broadly expressed Fe2p line an accelerating energy value of 50 eV and a sighthole with 6 mm width (entry/exit) is used. Instrumental resolution measured as full width at a half maximum (FWHM) of the photo electron peak of $Ag3d_{5/2}$ is lower than 1 eV. The binding energy scale is corrected to the maximal C1s peak at 285 eV for electrostatic charge. The spectra obtained are additionally analyzed using symmetric fitting Gauss-Lorenz curve after eliminating the background accordingly to the requirements of the method of Shirley.

3. Results and discussion

3.1 Potentiodynamic polarization investigations

3.1.1 Potentiodynamic investigations of low-alloyed steel (LAS)

The results for the corrosion behavior of LAS at 35 °C are demonstrated in Figure 1 and are used to characterize this steel and as a benchmark for comparison with high-alloyed steel.

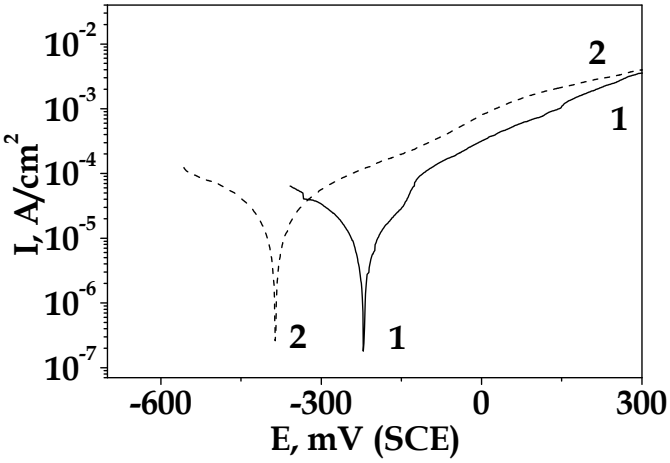


Fig. 1. PD polarization curves of LAS at 35 °C in:
1 – MCM; 2 – MCM with addition of 10 μg/L Cl^-.

It can be registered that even at low concentration of the added corrosion activators - Cl^- ions (10 μg/L), curve 2 - their unfavourable influence is well expressed. As seen from the results the corrosion potential value (E_{corr}) of the tested steel sample is strongly shifted in negative direction in that case with about 170 mV compared to the same parameter in the initial MCM (curve 1). Accordingly to this observation and as could be expected, the corrosion current density (i_{corr}) and the rate of the active anodic dissolution in the presence of chloride ions are about 2-3 times higher leading to accelerated damaging processes (greater anodic dissolution) under the conditions of external polarization.

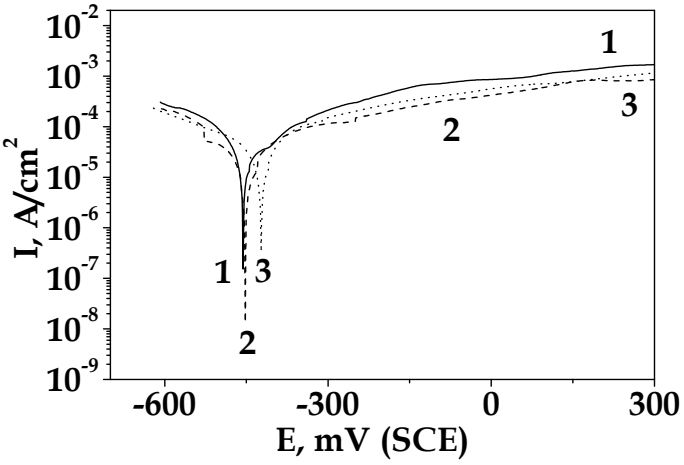


Fig. 2. PD polarization curves of LAS at 35 °C in:
1 – MCM with 300 μg/L Cl^-; 2 – MCM with 300 μg/L Cl^- and 2 mg/L MEA;
3 – MCM with 1 mg/L Cl^- and 2 mg/L MEA.

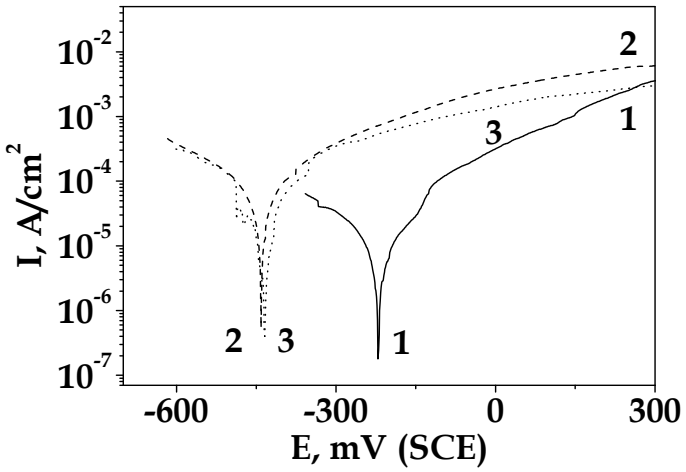


Fig. 3. PD polarization curves of LAS at 35 °C in:
1 – MCM; 2 – MCM with 300 µg/L SO_4^{2-};
3 – MCM with 300 µg/L SO_4^{2-} and 2 mg/L MEA.

Further increasing of the Cl^- ions concentration up to 300 µg/L (Figure 2, curve 1) results in another potential shift in the negative direction with about 200 - 230 mV compared to the case in MCM without addition of activators (Figure 1, curve 1) and with about 70 mV compared to case with the lower Cl- concentration (Figure 1, curve 2). Corrosion current density value also increases especially in the potential range between -300 and 0 mV (SCE). The addition of MEA even at not very high concentration (that could be applied in the practice) plays definitely a positive role – curve 2 – leading to lower corrosion and anodic current density values and fixing the further shifting of the corrosion potential. Even at very high concentration of corrosion activators (1 mg/L chloride ions – Figure 2, curve 3) the presence of the added corrosion inhibitor MEA decrease partially the cathodic (oxygen reduction) and also the anodic (metal dissolution) reactions demonstrating lower corrosion current - i_{corr} – values and ensuring better protection against corrosion.
Similar effect of MEA is observed also in the case when sulphate ions are used as corrosion activators in the same model medium and at the same experimental conditions (Fig. 3). The corrosion potential is shifted with about 220-230 mV in negative direction (curve 2) and i_{corr} increases about 4 – 5 times. After the addition of MEA the anodic dissolution rate is slowed down (curve 3) which demonstrates the positive influence of this compound. That result leads also to the conclusion that the protecting mechanism of MEA is most probably similar in aggressive media with different corrosion activators.
It is very important for the practice to register and estimate the effects also in the case when some of the most dangerous corrosion activators (Cl^-, SO_4^{2-} and Cu_{total}) simultaneously present in the medium at their extreme high concentrations.
The results obtained for LAS treated in this type of model solution are demonstrated in Fig. 4. It is evident that E_{corr} is strongly shifted in negative direction again and the corrosion current density values are very high which can be expected at these conditions. Also at these

extremely aggressive medium the addition of MEA decreases the i_{corr} and the dissolution in the whole anodic region. The reason for this positive result is that the influence of MEA simultaneously slow down the cathodic (reduction of depolarizer) and the anodic (dissolution of the metal) reaction rates. It can be supposed from these observations that the used compound MEA is an inhibitor of a "mixed" type.

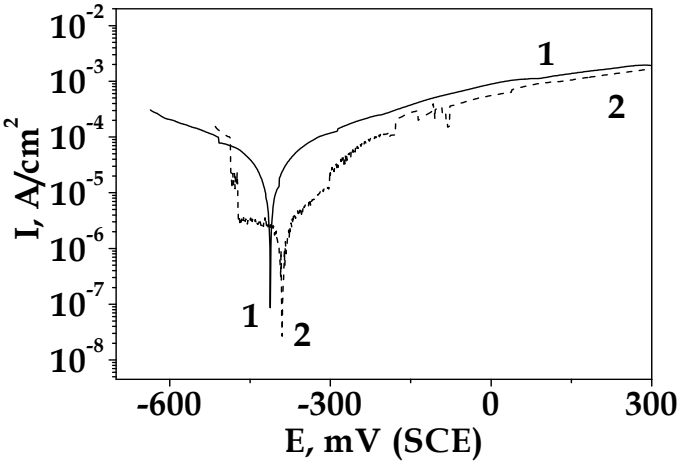


Fig. 4. PD polarization curves of LAS at 35 °C in:
1 – MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-} and 5 µg/L Cu_{total};
2 – MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-}, 5 µg/L Cu_{total} and 2 mg/L MEA.

The potentiostatic studies in the autoclave realized in MCM containing chloride ions show a slight shift of the corrosion potential in cathodic direction and about 3,5 times increase of the corrosion current density value for this steel type.

3.1.2 Potentiodynamic investigations of high-alloyed steel (HAS)

The results obtained for HAS which is practically used in the steam generators are demonstrated in Figure 5. It can be observed that low Cl^- ions concentration in the model medium (10 µg/L) – curve 2 - increases significantly (about 4-5 times) the corrosion rate. Additionally, well expressed shift of E_{corr} in negative direction of about 300 mV can be registered. Nevertheless, compared to the case of LAS (see Figures 1 and 3, curves 1) i_{corr} is about 5 – 6 times lower at these conditions. The registered differences are attributed to the better protective influence of the alloying components presenting in HAS. The addition of MEA (curve 3) expresses strong favourable effect leading to the shifting of E_{corr} in positive direction with about 350 mV while i_{corr} decreases about 4 times compared to the case of MCM with addition of Cl^-, but without MEA (compare curves 2 and 3).
Similar is the influence of MEA also in the case when SO_4^{2-} ions present as activators at equal concentrations in the model medium (Figure 6). As can be registered sulphate ions extremely increase the corrosion current density but the addition of MEA leads to inhibiting of the corrosion and anodic processes.

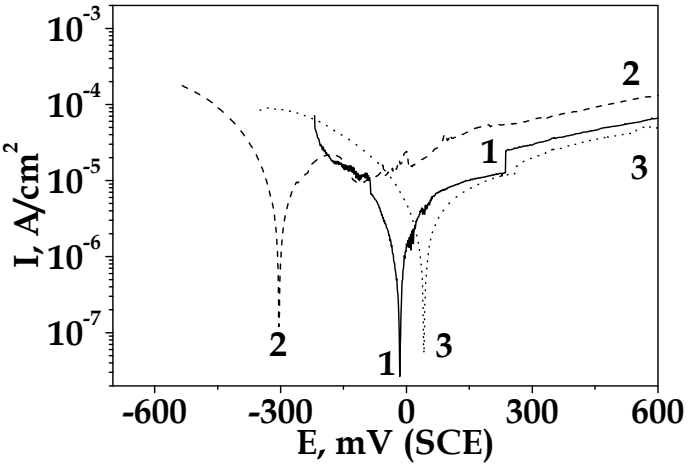


Fig. 5. PD polarization curves of HAS at 35 °C in:
1 – MCM; 2 – MCM with 10 μg/L Cl^-;
3 – MCM with 10 μg/L Cl^- and 2 mg/L MEA.

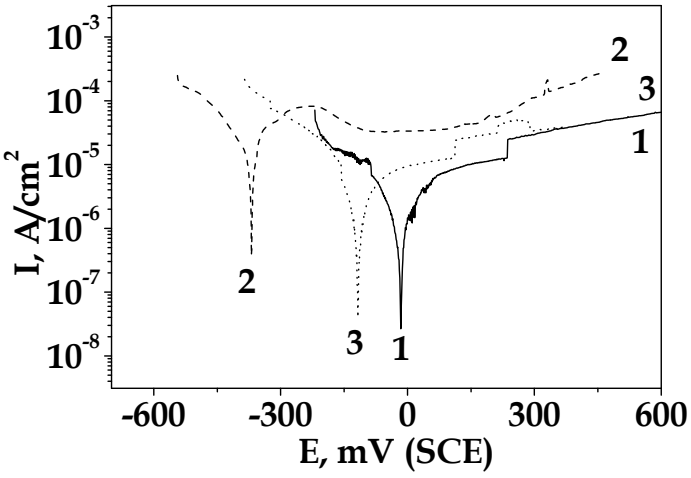


Fig. 6. PD polarization curves of HAS at 35 °C in:
1 – MCM; 2 – MCM with 300 μg/L SO_4^{2-};
3 – MCM with 300 μg/L SO_4^{2-} and 2 mg/L MEA.

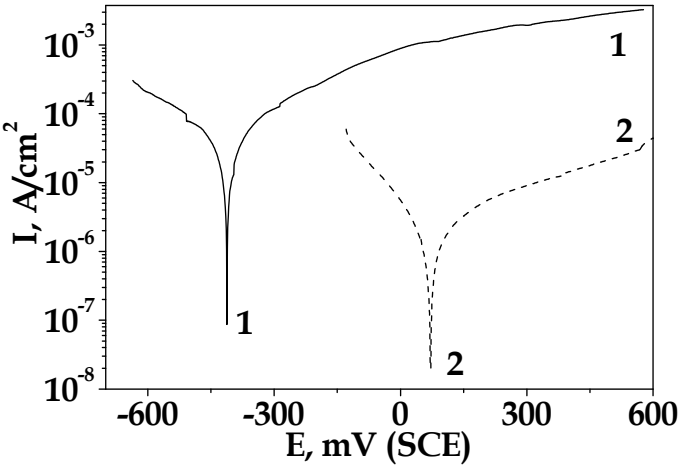


Fig. 7. PD polarization curves of HAS at 35 °C in:
1 – MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-} and 5 µg/L Cu_{total};
2 – MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-}, 5 µg/L Cu_{total} and 2 mg/L MEA.

In the case, when the combination of the most dangerous activators (Cl^-, SO_4^{2-} and Cu_{total}) is applied in MCM the corrosion activity is also very strong demonstrated - Figure 7. It is evident from the figure that the corrosion rate in that case is much greater and the corrosion potential is additionally shifted to the negative direction up to the value of about -410 mV.
Nevertheless, the inhibiting effect of MEA is also very strong (curve 2) - the corrosion potential is placed in positive direction, i_{corr} value is more than 10 times lower and the anodic dissolution seems to be partially hampered. The latter can be explained with the process of the effective inhibitor adsorption exactly on the active anodic zones on the metal surface. Finally, it could be prognosticated that MEA will slow down the corrosion processes also at exploitation conditions in the case when the three activators simultaneously present in their extreme high for the practice concentrations.
The potentiostatic investigations in the autoclave show more clearly expressed shift of the corrosion potential in cathodic direction with about 100 mV and about 4 times increase of the corrosion current density value for this steel. Despite of these results the new data obtained from this test are much lower compared to the case of LAS.

3.1.3 Potentiodynamic investigations of HAS in cracks

As a result of the operating conditions and WCR the equipment in NPP and in particular the SGs can be effective damaged by large number of cracks, holes, slits etc. This is a specific type of local corrosion and these dangerous places strongly differ in their morphology and element content compared to the whole surface of the unit. For example, in the case of newly appeared cracks the electrochemical potentials in and outside can be extremely different which will lead to appearance of galvanic macro-elements with own electromotive force. It is very important to know that local corrosion phenomena can appear not only in the already existing constructive slits. During the exploitation new cracks can occur especially in the cases where different materials or media are in contact or as a result of simultaneous action of electrochemical and mechanical forces.

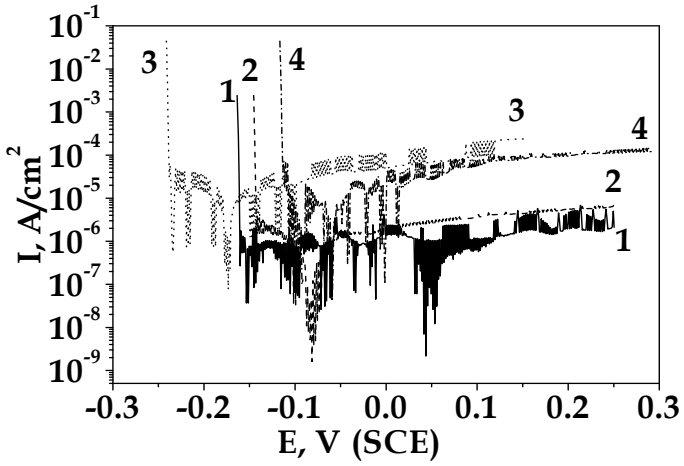


Fig. 8. PD curves of HAS in MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-} and 5 µg/L Cu_{total}:
1 – in crack 0,5 mm at 35 °C; 3 - in crack 0,1 mm at 35 °C;
2 - in crack 0,5 mm at 95 °C; 4 - in crack 0,1 mm at 95 °C.

It such a case it is very important to know the current density value into the crack that can be mould out of a specially constructed for this case electrode. It allows simulating the corrosion processes in cracks with different preliminary fixed widths. The present investigations are realized using two artificially constructed widths (sizes of 0,1 and 0,5 mm respectively). For practical purpose the special electrode has being filled in with the experimental solutions (with or without addition of corrosion activators) directly before the beginning of the test aimed to eliminate possible diffusion limitations.

The polarization investigations presented here (Fig. 8 and Fig. 9) are carried out in MCM containing the extreme high concentrations of the corrosion activators – 1 mg/l Cl^-, 1 mg/l SO_4^{2-} and 5 µg/l Cu_{total} – with and without addition of MEA, respectively. Comparing the characteristic parameter E_{corr} following can be summarized – Fig. 8:

- at 35 °C E_{corr} for the narrower crack (0,1 mm – curve 3) is strongly replaced in negative direction with about 100 mV compared to the wider one (0,5 mm - curve 1) as a result of the reasons described above;
- at 95 °C E_{corr} of both samples are relative closer – compare curves 2 and 4.

The oscillations of the curves are most probably on account of the inhibition of the processes of access and taking away the oxygen into the depth of the crack. In the anodic region the current density values are higher in the case of the narrower crack which means that the processes of dissolution are strongly expressed.

The addition of MEA in concentration of 2 mg/l shows favourable effect – E_{corr} shifts in positive direction for all investigated samples (compare Figures 8 and 9). Additionally, the observed i_{corr} values are also lower at both temperatures and crack widths – Fig. 9 – which is a sign for inhibiting of the corrosion processes. The electrochemical behavior of the steel in the narrow cracks and holes characterizes with some peculiarities mainly as a result of the transportation difficulties – inhibited access of the corrosive and passivating agents and difficult take away of CPs from these places. In neutral medium Fe shows increased rate of the anodic process and decreased – of the cathodic one. The lower oxygen concentration in

the cracks leads to more negative potentials of iron ionization with predominantly formation of bi-valence ions and their compounds the latter in general possessing insufficient protective properties.

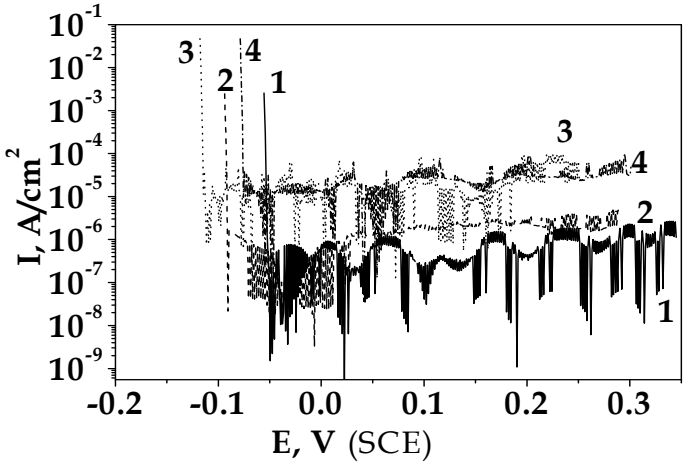


Fig. 9. PD curves of HAS in MCM with 1 mg/L Cl^-, 1 mg/L SO_4^{2-}, 5 μg/L Cu_{total} and 2 mg/L MEA:
1 – in crack 0,5 mm at 35 °C; 3 - in crack 0,1 mm at 35 °C;
2 - in crack 0,5 mm at 95 °C; 4 - in crack 0,1 mm at 95 °C.

The passive state of the metals owed mainly to the presence of oxygen may be affected by inhibited access of this component leading to depassivating and accelerating of the anodic reaction. Additionally, the cathodic reduction of the depolarizer becomes slowly, E_{corr} and PD curves shift to negative direction. The reason for this effect is the fact that the electrolyte situated in the cracks does not intermix convectively any more and the oxygen transportation occurs only by molecular diffusion. The inhibited refreshing of the electrolyte leads to relative fast change of the pH value of the medium in these places compared to the surface outside the cracks. The composition of the electrolyte in these close volumes also changes during the exploitation process and as result oxides, hydroxides etc. accumulate on the surface. Their concentration depends in general on the product of solubility value of the CPs. The main changes of the corrosion medium in the cracks appear as a result of the macro-element activity between these places and the "bare surface". The metal inside the cracks works as an anode and after incubation period during which the current of the galvanic element is very low the "inner" electrolyte becomes more acidic that leads to beginning of depassivation process.
After prolonged exploitation of the equipment the corrosion damages in the cracks become extremely dangerous due to the fact that the appearance of the passive state of the steels is ensured mainly from the oxygen in the medium. The permanent lack of this element into the deep volume of the cracks leads to depassivating of different places and acceleration of local corrosion processes. These results confirm that the modelling (although approximately) of the exploitation conditions is useful and necessary in order to receive experimental data for dangerous corrosion phenomena and their detrimental influence as well as to give the opportunity to prognosticate them.

3.2 SEM and EDX investigations

3.2.1 SEM investigations

Steel samples are initially placed in the special electrode for investigation of the corrosion processes in cracks with different widths. Thereafter the same electrode is putted in an autoclave and leaved at E_{corr} in MCM containing the extreme high concentrations of chloride, sulphate and copper ions. One 24 hours-cycle test consists of 4 hours stay at the operating conditions (selected temperature and pressure) followed by 20 hours for equalizing with the room temperature (slow cooling process). The SEM images of HAS samples are presented in Fig. 10(A,B). The results are obtained after preliminary preparation of the corrosion treated steel samples – rinsing with distilled water and drying with hot air.
Figure 10A shows typical corrosion damages of HAS samples placed in the electrode with 0,5 mm width after 3 cycles of the corrosion treatment described above. It can be registered that after the treatment a damaged place appears on the surface. Much hardly attacked seems to be the sample simulating the crack with 0,1 mm width – Fig. 10B. The reasons for this result are most probably the lower oxygen concentration in the solution. Additional risk is the accumulation of loose corrosion products in which mass the activators concentration will be greater.

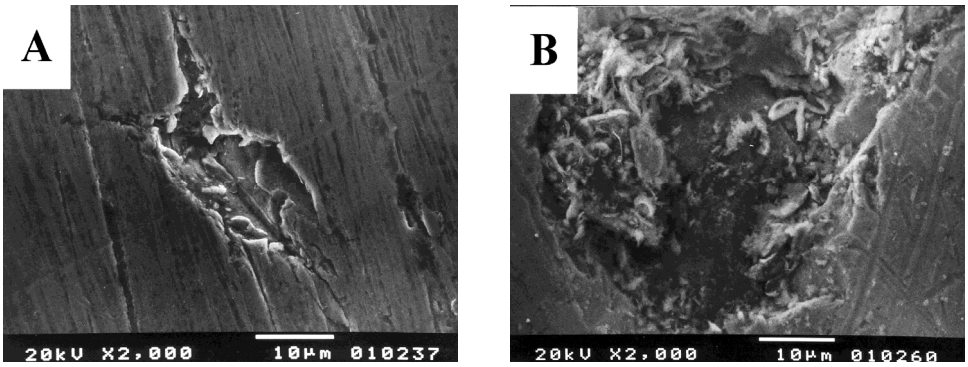


Fig. 10. SEM images of steel morphology after corrosion treatment of HAS samples:
A – in simulated crack with 0,5 mm width; B – in simulated crack with 0,1 mm width.

3.2.2 EDX investigations

The method gives the opportunity to determine the element composition in selected places on the steel surface characterized with different phenomena and processes - local holes, cracks etc. appeared after ruptures of the passive film. It is well known that after a definite time period the passive film in or over the crack or pitting can be totally destroyed and as a result open pits take shape. The average results from the analysis inside selected separate pits registered on the surface of the sample placed in the special electrode for investigation of the corrosion processes in the cracks and tested in MCM are presented in Table 1.
The main differences in the composition of the pits registered with and without MEA refer to the alloying components – Cr and Ni – both latter demonstrating higher percentages in the case with MEA. This is a clear positive effect since higher Ni and Cr amounts inhibit the penetration of the corrosion processes into the depth. The higher amount of iron detected when MEA present in MCM means that the dissolution of this metal is slowly.

Element	Content (in wt.%) after treatment in MCM	
	without MEA	with MEA
Fe	63,6	66,3
Cr	15,7	18,7
C	9,9	3,8
Ni	8,5	10,0
Si	1,9	0,8
Ti	0,4	0,4

Table 1. EDX analysis in selected pits and influence of MEA

3.2.3 SEM investigations of CPs

As already presented above the presence of MEA in the exploitation (or model) medium plays a positive role concerning the corrosion processes leading to their inhibiting. This compound can influence the appearance and development of the corrosion damages especially in places with limited oxygen access. The looseness, adhesion and the average size of the CPs particles are also of great importance in that case.

In the steam generators the corrosion products deposit on the body, pipe bundles, distance grids, bottom etc. and have in general sizes between 1 and 10 μm. On some places a compact film with good adhesion and low solubility consisting mainly of magnetite (Fe_3O_4) can appear. Additionally, a loose mass of new porous and with bad adhesion layers accumulate on it. As a result a strong increase of the local concentrations of Cl^-, SO_4^{2-} and Cu_{total} ions occur which leads to depassivation and appearance of pits on several places. In that sense some changes in the WCR leading to CPs with high dispersion are desirable aiming the easier removing of the particles from that zones.

The addition of MEA results in obtaining of CPs with high dispersion having average particle sizes less than 1 μm. Positive sign is also the fact that MEA do not lead to changes in the element and phase composition of CP. These effects are presented in Fig. 11.

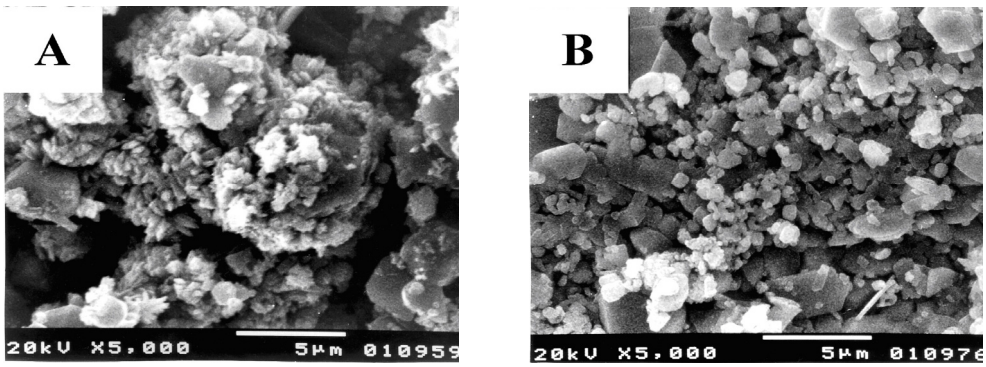


Fig. 11. SEM images of corrosion products:
A – without MEA in the medium; B – in the presence of 2 mg/l MEA.

3.3 Moessbauer investigations

The Moessbauer analysis (Fig. 12) of the surface film of corrosion products collected from characteristic places from the inner volume of the SGs during the shut-down procedure

show that the film consists of magnetite - Fe_3O_4 - and alpha-hematite - α – Fe_2O_3. As well known, these compounds are low soluble and highly resistive oxides, even in aggressive corrosion medium. This fact suggests that the corrosion products on the bundle surfaces may form a layer with high protective properties the latter impeding the penetration of the destructive corrosion processes in the depth of the metal tubes. The obtained average results are presented in Table 2.

Components	Weight %
Sxt 1 - Hematite, α – Fe_2O_3	62
Sxt 2 - Magnetite – tetra, Fe_3O_4	15
Sxt 3 - Magnetite - octa, Fe_3O_4	23

Table 2. Results from Moessbauer investigations

The experimental Moessbauer spectra include two types of components:
- components without super fine magnetic structure – quadruple doublets
- components with expressed super fine magnetic structure – sextets

In the present case the models for computer calculating of the Moessbauer spectra include only sextets (three sextet components). The values of the determined parameters of the six-component lines in the spectra (marked as Sxt 1, Sxt 2 and Sxt 3 in Table 2) can be attributed to the presence of iron-oxide phases – α–hematite (Sxt 1) and magnetite (Sxt 2 and Sxt 3). These two phases characterize with well expressed super fine magnetic structure (sextet components).

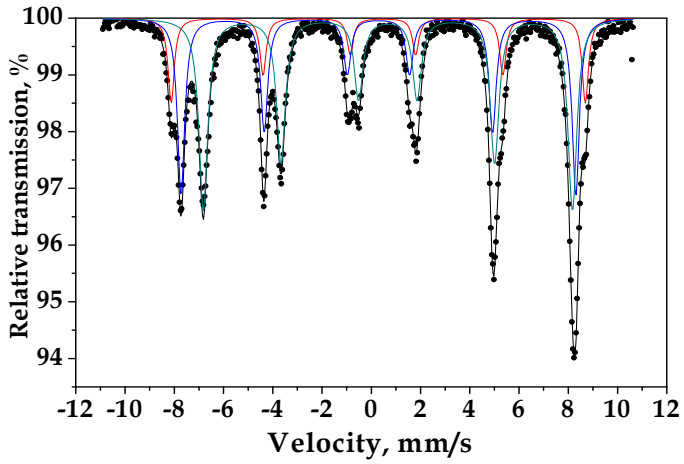


Fig. 12. Moessbauer spectra of CPs taken from SGs (average results)

The registered parameters for Sxt 1 demonstrate that all iron ions exist in the third oxidized state and octahedral surrounding. The appearance of two components of the magnetic phase, Sxt 2 and Sxt 3, is due to the existence of Fe in the third oxidized state and in tetrahedral surrounding. This is the reason for the presence of Sxt 2 in the spectrum as well as to a spectral appearance of iron ions with interstitial oxidized state ($Fe^{2.5+}$) in octahedral

position (its spectral expression is Sxt 3). As well known, magnetite is one of the widely distributed mixed valence compound with high frequency electron exchange between the ions $3d^6Fe^{2+} \leftrightarrow 3d^5Fe^{3+}$. Magnetite characterizes with cubic symmetry and the results demonstrate that it is with low degree of non-stoichiometry - Fe_3O_{4-x}, while the vacancies are localized most probably in the octahedral places.

3.4 XPS investigations

The results described below are related to powder CPs selected from different places in the volume of the SGs during the shut-down procedure. The data presented in Table 3 are summarized for all test samples.

Element	O 1s	Fe 2p	Cu 2p	Cr 2p	Ni 2p
Binding energy, eV (main peaks)	529,9 531,2	711,1	932,9 934,5	576,8	855,2
Percentage, at.%	76	8	14	1	1

Table 3. Results from XPS investigations of CPs taken from the inner volume of SGs (average results from all investigated samples)

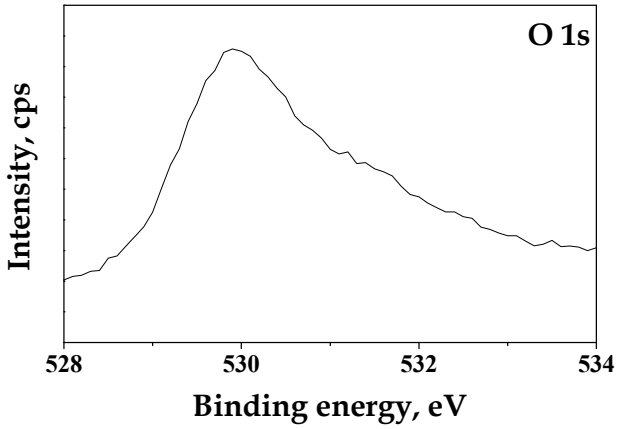


Fig. 13. XPS measurements for O1s from the CPs of the SGs (average results)

The XRD spectra obtained for different registered elements are presented with more details in the Figures 13 – 17. In general, these experimental results are in correlation with the Moessbauer investigations and also in fair agreement with the conclusions from the potentiodynamic polarization curves and SEM images. The presence of Cu is not a desired event since it could lead to local corrosion damages being more cathodic compared to iron.

As can be observed the main element with the highest content (in atomic percent) is oxygen followed by copper and iron as well as by much smaller amounts of Cr and Ni. In such a case it could be expected that the compounds on the surface will be mainly iron, copper or iron-copper oxides and hydroxides. The very low amounts of chromium and nickel compounds clearly demonstrate that their participation in the occurred corrosion processes is weaker expressed.

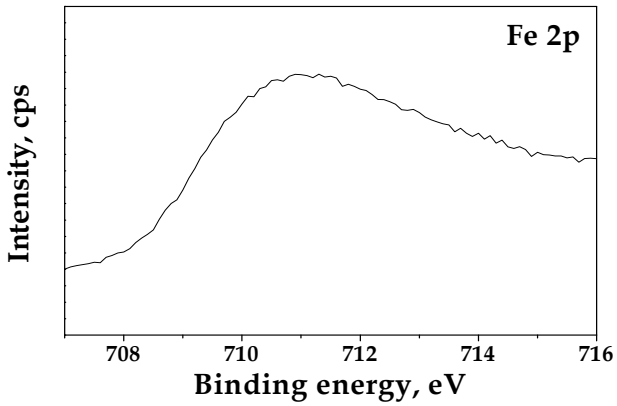


Fig. 14. XPS measurements for Fe2p from the CPs of the SGs (average results)

Element	Binding energy E, eV/ (at.%)	Possible compounds
O 1s	530,1 / (56,3) 531,6 / (37,0) 533,2 / (6,7)	Fe_2O_3; FeO; CuO FeOOH; $Cu(OH)_2$ H_2O
Cu 2p	932,8 / (57,0) 934,4 / (43,0)	Cu_2O; Cu. CuO; $Cu(OH)_2$; $CuCr_2O_4$

Table 4. Results after mathematical deconvolution of XPS investigations of CPs

However, the iron content is not very high which means that the used construction is to a definite degree protected at these conditions. The alloying elements nickel and chromium which are also registered during the investigations show practically their positive influence for the increased corrosion resistance of the used alloy.

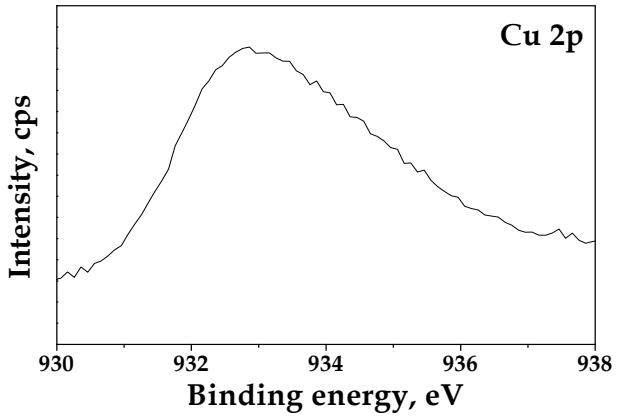


Fig. 15. XPS measurements for Cu2p from the CPs of the SGs (average results)

Aimed to receive clear results the XPS spectra obtained are additionally mathematically calculated using a special software aimed to determine also another peaks of possible

compounds that present in the sample but are not enough visible. This additional mathematical treatment known as "deconvolution process" gives more adequate information about the real situation in the samples. The results obtained are presented in Table 4 for two of the main elements – O1s and Cu2p.

The experimental data for the binding energy values of the main component presenting in the samples - oxygen – support the expectation that predominantly iron and copper oxides and hydroxides as well as water amounts appear - see Tables 3 and 4.

It follows from these data that the most possible iron compounds seems to be the oxides of Fe^{2+} and Fe^{3+} ions both latter presenting also in the composition of magnetite. This is a very important result as far as being in good correlation with the Moessbauer spectroscopy. Additionally, FeOOH has been detected but its amount seems to be lower compared to the quantities of bi- and trivalent iron oxides. The Cu spectrum demonstrates that this element has been detected mainly as pure metal and in its first oxidized state – in the form of Cu_2O and Cu. Their concentrations are comparable with the amounts of bivalent copper compounds – CuO, $Cu(OH)_2$ and most probably $CuCr_2O_4$.

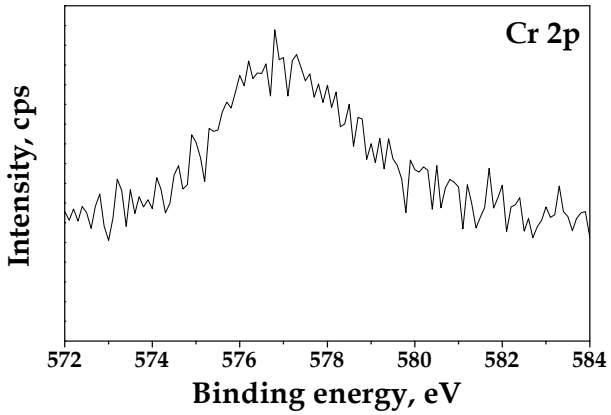


Fig. 16. XPS measurements for Cr2p from the CPs of the SGs (average results)

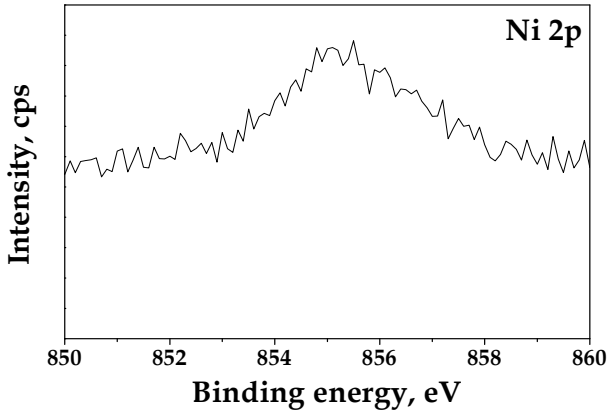


Fig. 17. XPS measurements for Ni2p from the CPs of the SGs (average results)

As already discussed above the other two elements (Ni and Cr) detected by XPS method are in much smaller amounts – about 1 at%. Nickel is in the form of Ni^{2+} ions, most probably as oxides and hydroxides - $NiO/Ni(OH)_2$. Chromium is registered in its third and sixth valencies (Cr_2O_3 and CrO_3) in a correlation between them of about 7:1. As can be registered these results differ to a certain degree with the results in Table 3 due to the mathematical treatment but from other side they confirm the general tendency of appearance of iron and copper hydroxides and other compounds. Main positive conclusion is that they are in agreement with the other used methods.

4. Conclusion

At model conditions and water chemistry regime applied, the corrosion current for austenitic stainless steel 18Cr10NiTi used in the construction is much lower compared to the low-alloyed steel. The investigations carried out at two operating temperatures and also at conditions close to the exploitation ones demonstrate the influence of selected corrosion activators on the most dangerous cases in the exploitation practice.

Having in mind the obtained results it can be concluded that the application of monoethanolamine significantly decrease the corrosion current density and leads to better corrosion resistance of the used steel. This compound also leads to appearing of middling slime with higher dispersion that can be easily removed from these places.

The corrosion processes occurring in the narrower cracks are more intensive compared to the processes in the wider cracks especially at higher temperatures and lead to faster destroying of the equipment at operating conditions. The application of MEA which is a "mixed" type inhibitor decrease partially the cathodic (oxygen reduction) and also the anodic (metal dissolution) reactions ensuring lower corrosion current density values and better protection against corrosion.

The presence of the protective layer of corrosion products of iron on the surface of the equipment is confirmed using Moessbauer spectroscopy and XPS investigations which are in fair agreement with the results obtained by the electrochemical polarization technique, SEM and EDX analysis. Taking into account that at real operation conditions in NPP "Kozloduy" and prolonged exploitation of the steam generators a film composed mainly of magnetite and hematite is readily formed on the pipe surfaces, it can be expected that the actual corrosion rate should be much lower.

Very important factor which influence the corrosion damages is the iron and copper concentration in the water of condense-feeding part of the SGs. This also affects their content in the corrosion products. Better exploitation conditions will be realized by periodical removing of the deposited corrosion products on the pipe bundles and on the other places in the steam generators where they appear.

5. References

Andreeva M. et. al., (2008). Overview of plant specific severe accident management strategies for Kozloduy nuclear power plant, WWER-1000/320, *Annals of Nuclear Energy*, v. 35, pp. 555–564, ISSN: 0306-4549.

R.W. Staehle and J.A. Gorman, (2003), Quantitative Assessment of Submodes of Stress Corrosion Cracking on the Secondary Side of Steam Generator Tubing in PWRs;

Part I, II & III, *Corrosion,* v. 59, No. 11, pp. 931-994; Houston, TX: NACE, ISSN 0010-9312

R.W. Staehle and J.A. Gorman, (2004), Quantitative Assessment of Submodes of Stress Corrosion Cracking on the Secondary Side of Steam Generator Tubing in PWRs; Part I, II & III, *Corrosion,* v. 60, No. 1, pp. 5-63; Houston, TX: NACE, ISSN 0010-9312

R.W. Staehle and J.A. Gorman, (2004), Quantitative Assessment of Submodes of Stress Corrosion Cracking on the Secondary Side of Steam Generator Tubing in PWRs; Part I, II & III, *Corrosion,* v. 60, No. 2, pp. 115-180; Houston, TX: NACE, ISSN 0010-9312

J. Congleton, T. Shoji and R.N. Parkins, (1985), "The Stress Corrosion Cracking of Reactor Pressure Vessel Steel in High-Temperature Water, *Corrosion Science*, v. 25, No. 8/9, p. 633-650, ISSN: 0010-938X .

Sviridenko, I., (2008). Heat exchangers based on low temperature heat pipes for autonomous emergency WWER cooldown systems, *Applied Thermal Engineering*, v. 28, pp. 327 – 324, ISSN: 1359-4311

Hadavi, S.M.H., (2008). WWER-1000 shutdown probabilistic risk assessment: An introductory insight, *Annals of Nuclear Energy,* v. 35, pp. 196-206, ISSN: 0306-4549

Viehrig, H.-W., et. al., (2006). Application of advanced master curve approaches on WWER-440 reactor pressure vessel steels, *International Journal of Pressure Vessels and Piping*, v. 83, pp. 584–592, ISSN: 0308-0161

Slugen, V., et. al., (2005). Corrosion of steam generator pipelines analysed using Moessbayer spectroscopy, *Nuclear Engineering and Design*, v. 235, pp. 1969–1976, ISSN: 0029-5493

Raichevski G. et al., Corrosion investigations and SEM studies of austenitic stainless steel used in the steam generators of the Nuclear Power Plant "Kozloduy", First International Conference "Corrosion and Material Protection", EFC event No. 294, paper No. 72, ISBN 978-80-903933-0-1, Prague, Czech Republic, 01 – 04 October 2007.

Hojna A. et. al, Electrochemical evaluation system – modular system for long-term on-line monitoring of corrosion processes in structural crevices of steam generator secondary circuits, First International Conference "Corrosion and Material Protection", EFC event No. 294, paper No. 83, ISBN 978-80-903933-0-1, Prague, Czech Republic, 01 – 04 October 2007.

Kaczorowski D. et. al., (2006). Water chemistry effect on the wear of stainless steel in nuclear power plant, *Tribology International*, v. 39, pp. 1503 – 1508, ISSN: 0301-679X.

Zubchenko, A.S., et. al., (2004). Effect of nickel content on mechanical properties and fracture toughness of weld metal of WWER-1000 reactor vessel welded joints, *International Journal of Pressure Vessels and Piping*, v. 81, pp. 713 – 717, ISSN: 0308-0161

Slugen, V., et. al., (2002). Moessbauer spectroscopy used for testing of reactor steels, *NDT&E International* (*Independent Nondestructive Testing and Evaluation*), v. 35, pp. 511 – 518, ISSN: 0963-8695

Lunin, G.L., et. al., (1997)., Status and further development of nuclear power plants with WWER in Russia, *Nuclear Engineering and Design*, v. 173, pp. 43 – 57 , ISSN: 0029-5493

N.D. Budiansky, et al., (2005), Detection of Interactions Among Localized Pitting Sites on Stainless Steel Using Spatial Statistics, *Journal of the Electrochemical Society*, v.152, No. 4, pp. B152-B160, ISSN: 0013-4651.

14

Collapse Behavior of Moderately Thick Tubes Pressurized from Outside

Leone Corradi, Antonio Cammi and Lelio Luzzi
Politecnico di Milano - Department of Energy
Enrico Fermi Center for Nuclear Studies (CeSNEF)
Italy

1. Introduction

This study originated from a specific problem that arose in conjunction with the IRIS (International Reactor Innovative and Secure) project (Carelli et al., 2004; Carelli, 2009). IRIS adopts an integrated primary system reactor (IPSR) configuration with all the primary loop components of a classical Pressurized Water Reactor (PWR) contained inside the vessel (Fig. 1). Among the reactor core internals are the steam generator (SG) units (Cinotti et al., 2002) with the primary fluid flowing outside the tube bundles and subjecting them to significant external pressure. In this situation buckling affects the tube collapse modality and codes become extremely conservative, to the point that up to five years ago design procedures based on the *ASME Boiler & Pressure Vessel* code (Section III) required an external diameter to thickness ratio (D/t) less than 8.5, leading to an increased thermal resistance in the heat exchange process between primary and secondary fluids, with detrimental consequences on the dimensioning of the heat transfer surface. A reduction in the tube thickness would allow the reduction of the overall heat transfer surface needed to exchange the same amount of power, with consequent saving on tube lengths and/or number of tubes. On the other hand, if the design of the steam generator units is not modified, an increase in the exchanged thermal power and a consequent up rating of the reactor can be obtained.
Besides IRIS, other recent proposals for next generation power plants based on PWR technology consider an IPSR design (Ingersoll, 2009; Karahan, 2010; Ninokata, 2006). Such integrated design is particularly suitable for small sized units, i.e., reactors with a power less than 300 MWe following the IAEA's definition (IAEA, 2007). A significant number of small sized PWR IPSRs is currently under development (e.g., RITM-200, ABV, CAREM, SMART, MRX, NHR-200, Westinghouse SMR, mPower, NuScale, see http://www.world-nuclear.org/info/inf33.html). Packing all the PWR primary components into the reactor pressure vessel (RPV) (Fig. 2) offers several advantages (Ingersoll, 2009): (i) all large coolant pipes are eliminated (only small feed water and steam outlet pipes penetrate the vessel wall); (ii) the total inventory of primary coolant is much larger than for an external loop PWR (this feature increases the heat capacity and thermal inertia of the system and hence yields a much slower response to core heat-up transients); (iii) typically the heat exchangers are placed above the core creating a relatively tall system that facilitates more effective natural circulation of the primary coolant in the case of a coolant pump failure; (iv) the vessel accommodates a relatively large pressurizer volume that provides better control

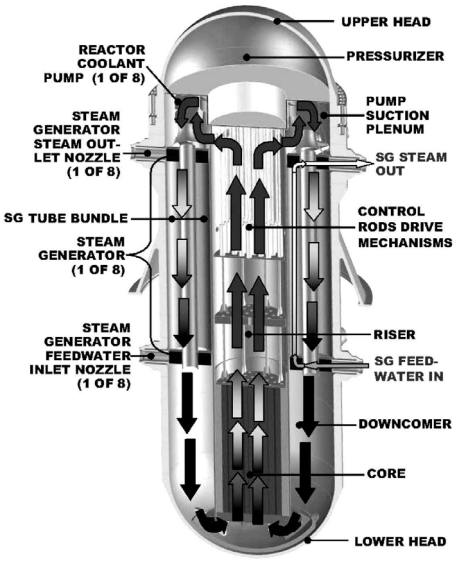
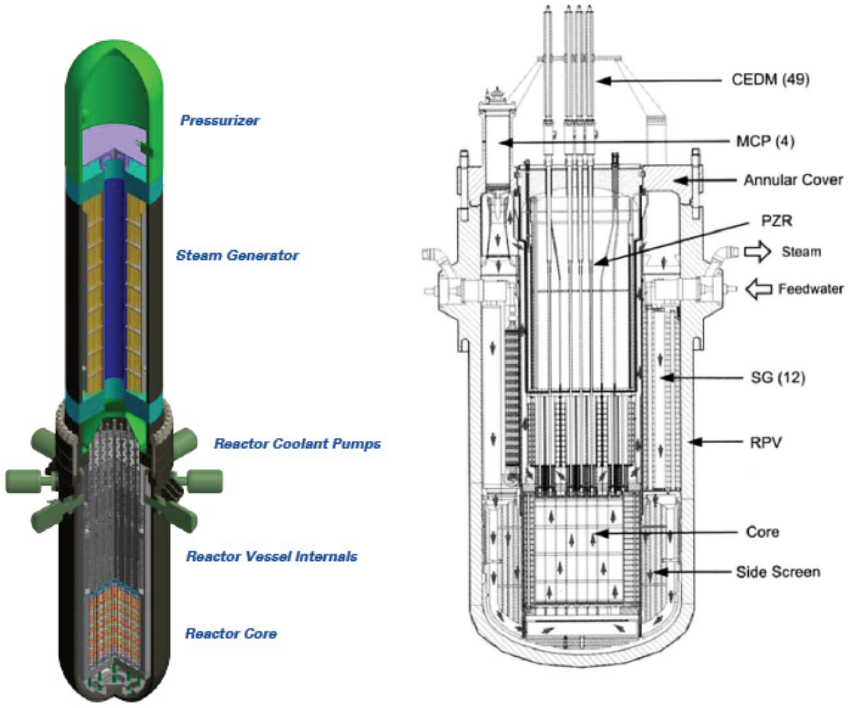


Fig. 1. IRIS reactor pressure vessel module (Reproduced from (Luzzi & Di Marcello, 2011))

of under/overpressure transients; and (v) the extended riser area provides the possibility for internal placement of the control rod drive mechanisms, thus avoiding the potentially serious accident scenario represented by the rod ejection.

All these solutions entail the presence of tubes or pipes pressurized from outside and, as for IRIS, their sizing has to face the severity of the code. It was felt that ASME code requirements were exceedingly conservative and both numerical (Corradi et al., 2009) and experimental (Lo Frano & Forasassi, 2009) investigations were performed to assess the collapse pressure of the tubes. The first results obtained within this framework contributed to the approval of Section III Code Case N-759 (ASME, 2007), which permits considerable thickness saving. From a regulation point of view the problem of excessive thickness was overcome, but the question of the collapse behavior of tubes in this thickness range remains open. These tubes are thicker not only than the very thin shells typical of aerospace applications, but also than the moderately thin tubes employed in oil industry as pipes or casings. On the other hand, they are not as thick as those that are encountered in high pressure technology and their behavior has been the object of a limited amount of study so far.

Very thin shells fail because of elastic buckling and very thick tubes because of plastic collapse, so that their ultimate pressures can be established on theoretical ground, by exploiting well established results (Budiansky, 1974; Mendelson, 1968). In the intermediate range, on the contrary, plasticity and buckling interact and, in principle, the strength of the tube can only be assessed numerically. Attempts at reproducing numerical results with empirical design formulas were also made at the end of the last century (e.g., (Haagsma & Schaap,



(a) (b)

Fig. 2. Integrated primary system reactors of small size. (a) Westinghouse SMR-200 MWe (Small Modular Reactor); (b) SMART-90 MWe (System Integrated Modular Advanced Reactor). Reproduced from http://www.westinghousenuclear.com/smr/fact_sheet.pdf and from (Ninokata, 2006), respectively

1981; Tamano et al., 1985; Yeh & Kyriakides, 1988)): such results are adequate in the range of interest for oil industry, but become questionable for the thicker tubes required by the nuclear applications mentioned above. In this range, collapse is dominated by yielding, but interaction with buckling is still significant and reduces the pressure bearing capacity by an amount that cannot be disregarded when safety is of primary concern.

The problem is similar to that of beam columns of intermediate slenderness, which also fail because of interaction between yielding and buckling and that have been studied in detail. A simple predictive formula was proposed in this context, which turns out to be reasonably accurate for any slenderness and several code recommendations are based on it (e.g., (EUROCODE 3, 1993)). An attempt at adapting such formula to the case of tubes was made in (Corradi et al., 2008), but a direct modification was successful only in the medium thin tube range, where the formula appears as a feasible alternative to other proposals. With increasing thickness the formula becomes conservative and only provides a, often coarse, lower bound to the collapse pressure. A correction was proposed which, however, is to a large extent empirical and based on fitting of numerical results.

In this study a different proposal is advanced, which is felt to better embody the physical nature of the phenomenological behavior. Comparison shows that tubes behave essentially as columns for $D/t \geq 25-30$, but differences make their appearance and grow up to significant values as this ratio diminishes. One reason is of geometric nature: the curvature of the tube wall increases with diminishing D/t ratio and the analogy with a straight column no longer applies. Another source of discrepancy is the stress redistribution capability that thick tubes, in contrast to columns, possess and can exploit with significant benefit. This aspect is not purely geometric: stress redistribution *capability* is still function of thickness, but the possibility of exploiting it is influenced by material properties as well. By properly interpreting these aspects, a formula is obtained that appears reasonably simple and accurate. In addition, it is felt that it provides a deeper understanding on the collapse behavior of cylindrical shells in a thickness range so far overlooked.
A comment on terminology is in order. Labels like "thick" or "thin" when applied to tubes are to some extent ambiguous, since they are used in a different sense in different contexts. A pipeline in deep sea water would be considered as a thick tube by an aerospace engineer and as thin one by high pressure technology people. Often, the term "thin tube" is used when thin shell assumptions, which consider stresses to be constant over the thickness, apply, but this definition also becomes questionable outside the elastic range. In this study, reference is made to the failure modality. A tube is called *thin* if it fails because of elastic buckling and *thick* when only plastic collapse is relevant. In the intermediate region the two failure modalities interact, with different weight for different slenderness. In *moderately thin* (or *medium thin*) tubes, buckling is the critical phenomenon even if plasticity plays some role; similarly, in *moderately thick* tubes failure is dominated by yielding, but interaction with buckling has non negligible effects. The separation line is not very sharp (in oil industry applications, for instance, the two phenomena have comparable weight), but the tubes of prominent interest in this study definitively belong to the *moderately thick* range.

2. Collapse of cylindrical shells pressurized from outside

2.1 Basic theoretical results

Consider a cylindrical shell of nominal circular shape, with outer diameter D and wall thickness t, subjected to an external pressure q. The shell is long enough for end effects to be disregarded. The material is isotropic, elastic-perfectly plastic and governed by von Mises' criterion. E and ν are its elastic constants (Young modulus and Poisson ratio, respectively) and σ_0 denotes the tensile yield strength.
In the theoretical situation of a perfect tube, the limit pressure is given by the smallest among the following values

$$\textit{Elastic buckling pressure} \qquad q_E = 2\frac{E}{1-\nu^2}\frac{1}{\frac{D}{t}\left(\frac{D}{t}-1\right)^2} \tag{1a}$$

$$\textit{Plastic limit pressure} \qquad q_0 = 2\sigma_0\frac{t}{D}\left(1+\frac{1}{2}\frac{t}{D}\right) \tag{1b}$$

The first expression is well known (Timoshenko & Gere, 1961), while equation (1b) was established in (Corradi et al., 2005) and is a very good approximation to the exact value for $D/t > 6$.

Equations (1) apply to possibly thick tubes, which demand that stress variation over the thickness be considered. Nevertheless, the average value S of the hoop stress σ_ϑ is a meaningful piece of information. Its value is dictated by equilibrium only and reads

$$S = \frac{1}{2} q \frac{D}{t} \tag{2}$$

For thick cylinders, peak stresses may exceed significantly the value (2), which is simply an alternative, often convenient, way to refer to pressure. In particular, the theoretical limits (1) may be replaced by the expressions

$$S_E = \frac{E}{1-\nu^2} \frac{1}{\left(\frac{D}{t}-1\right)^2} \qquad S_0 = \sigma_0 \left(1 + \frac{1}{2}\frac{t}{D}\right) \tag{3}$$

which are obtained by substituting in equation (2) either of the values (1) for q.
As the thickness decreases, local values approach their average and equation (2) becomes meaningful as a stress intensity measure for sufficiently thin tubes, which are usually studied by assuming $\sigma_\vartheta \approx S$. Also, the difference between the outer face of the tube, where the pressure acts, and the middle surface, where the resultant of hoop stresses is applied, is ignored. Within this framework, equations (1) become

$$\textit{Elastic buckling pressure} \qquad p_E = 2\frac{E}{1-\nu^2}\left(\frac{t}{D}\right)^3 \tag{4a}$$

$$\textit{Plastic limit pressure} \qquad p_0 = 2\sigma_0 \frac{t}{D} \tag{4b}$$

or, in terms of average hoop stress

$$F_E = \frac{E}{1-\nu^2}\left(\frac{t}{D}\right)^2 \qquad F_0 = \sigma_0 \tag{5}$$

Here (and in the sequel) p is used instead of q and F instead of S when computations are based on thin shell assumptions.
In the theoretical situation, the two critical phenomena of elastic buckling and plastic collapse are independent from each other. The quantity

$$\Lambda = \sqrt{\frac{q_0}{q_E}} = \sqrt{\frac{1}{\kappa}\left(\frac{D^2}{t^2} - \frac{3}{2}\frac{D}{t} + \frac{1}{2}\frac{t}{D}\right)} \tag{6}$$

or, if thin shell approximation is adopted

$$\Lambda = \sqrt{\frac{p_0}{p_E}} = \frac{1}{\sqrt{\kappa}}\frac{D}{t} \tag{7}$$

is known as *slenderness ratio*. The parameter

$$\kappa = \frac{1}{1-\nu^2}\frac{E}{\sigma_0} \tag{8}$$

is a dimensionless material property. $\Lambda = 1$ is the *transition value*, separating the range of comparatively thin tubes ($\Lambda > 1$, $q_E < q_0$), theoretically failing because of elastic buckling, from that of comparatively thick ones ($\Lambda < 1$, $q_0 < q_E$), when the critical situation is plastic collapse. Fig. 3 depicts schematically the two failure modalities.

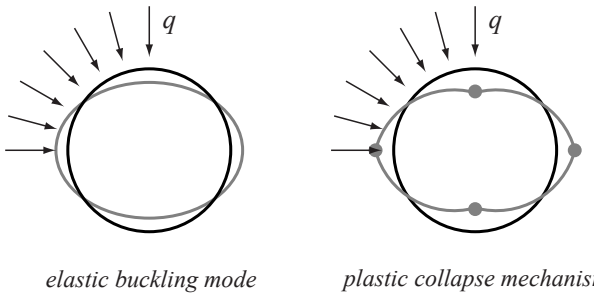


Fig. 3. Failure modalities for a long tube

2.2 Effects of imperfections

The situation above is "theoretical" in that it refers to the ideal case of a *perfect* tube. A real tube is unavoidably affected by imperfections, which introduce an interaction between plasticity and instability. As a consequence, the ultimate pressure is lower than the theoretical value. Fig. 4 depicts some aspects of the solution of a tube with an initial out of roundness (*ovality*): the pressure-displacement curve grows up to a maximum value, corresponding to failure, and then decreases; at the maximum, the tube is only partially yielded, i.e., plastic zones (in color) nowhere spread across the entire tube thickness (Fig. 4a). The "four hinge" mechanism is is attained in the post-collapse portion of the curve only (Fig. 4b). Failure occurs because of *buckling of the partially yielded tube*: even if not forming a mechanism, plastic zones reduce the tube stiffness and make the buckling load diminish. Failure corresponds to the elastic buckling of a tube of variable thickness, consisting of the current elastic portion.

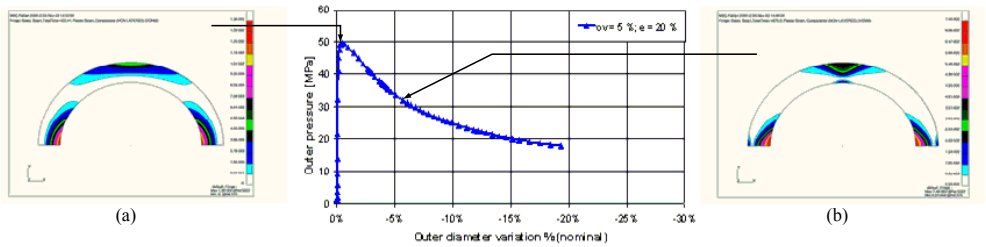


Fig. 4. Response of an initially oval tube. (a) plastic zones at failure; (b) four hinge mechanism in the post-collapse phase

To compute the failure pressure, complete elastic-plastic, large displacement analyses up to collapse are required, explicitly accounting for different kinds of possible imperfections. A systematic study was undertaken at the Politecnico di Milano and results are summarized in (Corradi et al., 2009; Luzzi & Di Marcello, 2011). Imperfections of both geometrical (initial out of roundness, non uniform thickness) and mechanical (initial stresses) nature were considered. As in a sense expected, it was found that all of them have similar consequences, causing a significant decay of the failure pressure with respect to the theoretical one for slenderness ratios close the transition value, with interaction effects diminishing as Λ departs from one in either direction. In any case, some decay was experienced in the entire range $0.2 \leq \Lambda \leq 5$, covering all situations of practical interest, except possibly high pressure technology or

aerospace engineering. The slenderness ratio of tubes for the aforesaid nuclear applications is low, but not enough to disregard the effects of interaction with instability: the IRIS steam generator tubes bundles, if sized according to Code Case N-759, correspond to $\Lambda \approx 0.4$.
When the study was started, Code Case N-759 was not available and ASME Section III rules required an external diameter to thickness ratio $D/t = 8.27$ ($\Lambda \approx 0.25$). Such a design is surprisingly severe and it was felt that the code assumed an *a-priori* conservative attitude for tubes belonging to a range scarcely studied both from the numerical and the experimental points of view, reflecting a substantial lack of knowledge on the phenomena involved. The numerical campaign was intended as a first step toward the definition of a suitable failure pressure, a reliable reference value permitting the derivation of an allowable working pressure through the use of a proper safety factor (Corradi et al., 2008). Computations had to include imperfections (one drawback of ASME III rules was that imperfections were not explicitly considered) but, since the effects of all of them were found to be similar, only the most significant was considered. This was identified with an initial out of roundness, or *ovality*, defined by the dimensionless parameter

$$W = \frac{D_{\max} - D_{\min}}{D} \tag{9}$$

where $D_{\max}$ and $D_{\min}$ are the maximum and minimum diameters of the ellipsis portraying the external surface of the tube (see Fig. 8a in the subsequent section) and D is their average value (nominal external diameter). To the failure pressure q_C computed in this way (a reasonable choice for the reference value) a safety factor is applied so as to reproduce ASME Section III sizing for medium thin tubes, a well known and well explored range, in which the code can be assumed to consider the proper safety margin (see (Corradi et al., 2008) for details). If the same factor is applied to thicker tubes as well, significant thickness saving is achieved without jeopardizing safety.
The requirement that the reference pressure be computed numerically makes the procedure cumbersome and an attempt at reproducing numerical results with an empirical formula was made (Corradi et al., 2008). The formula is adequate for practical purposes, but the approach is not completely satisfactory for a number of reasons: (i) the formula is involved and a simpler expression is desirable; (ii) its empirical nature does not help the understanding of the mechanical aspects of the tube behavior, and (iii) the formula is not equally accurate for all materials. Its coefficients were determined by considering the material envisaged for the IRIS SG tube bundles, i.e. Nickel-Chromium-Iron alloy N06690 (INCONEL 690) and the formula is fairly precise for $700 \leq \kappa \leq 1100$, where κ is defined by equation (8). Some materials, however, either because of high tensile yield strength σ_0 or low Young modulus E, have values of κ significantly below the lower limit; in these instances, the formula entails errors up to 10%, even if always on the safe side. Table 1 lists some of the materials investigated, with the properties employed in (Corradi et al., 2008) for computations and that will be used in this study as well. Trouble was experienced with aluminum and titanium alloys.

3. Interaction domains

3.1 Preliminary: load bearing capacity of struts

The proposal advanced in this study originates from the approach used to evaluate the collapse load of compressed columns, which is briefly outlined to introduce the procedure.
Consider the strut in compression illustrated in figure 5a. Its center line has an initially sinusoidal shape of amplitude U. The critical section, obviously, is the central one, where

material	E (GPa)	σ_0 (MPa)	ν	κ	$\left(\frac{D}{t}\right)_{\Lambda=0.2}$
stainless steel UNS S31600	200	200	0.31	1106	7.44
nickel-chromium-iron alloy N06690	183	240	0.29	832	6.56
aluminum alloy UNS A96061	70	240	0.35	332	4.46
titanium alloy UNS R56400	110	830	0.34	150	3.28

Table 1. Material properties for the considered materials

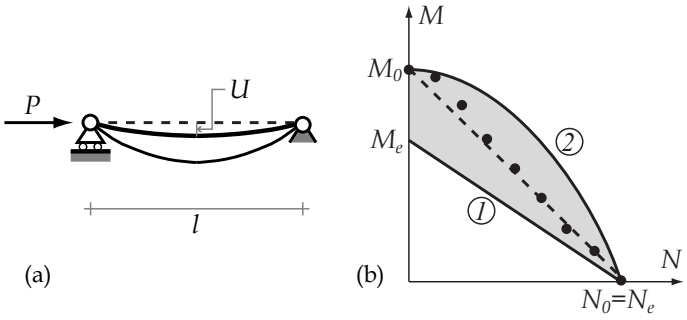


Fig. 5. Compressed column with initial imperfection

the axial force is $N = P$ (compression positive) and the bending moment M is expressed as

$$M = PU\frac{1}{1-\frac{P}{P_E}} \tag{10}$$

where P_E is the Euler buckling load ($P_E = \pi^2 EI/l^2$) and $1/\left(1 - P/P_E\right)$ the *magnification factor*. Equation (10) is exact in the elastic range since the initial imperfection has the same shape as the buckling mode (Timoshenko & Gere, 1961).

The behavior of the cross section is subsumed by the *interaction diagrams* in figure 5b. Line 1 is the *elastic limit* and for $N - M$ values on it one fiber is about to yield; line 2 is the *limit curve*, bounding the domain of $N - M$ combinations that can be borne. The gray zone is the elastic plastic region, corresponding to partially yielded sections. N_e and M_e are the values that individually bring the section at the onset of yielding, N_0 and M_0 the corresponding values exhausting the sectional bearing capacity. Obviously, it is $N_e = N_0$, since in pure compression stresses are uniform. The elastic limit is given by

$$\frac{N}{N_e} + \frac{M}{M_e} = 1 \tag{11a}$$

while the equation of the limit curve depends on the sectional shape. For rectangular cross sections one has

$$\left(\frac{N}{N_0}\right)^2 + \frac{M}{M_0} = 1 \tag{11b}$$

with $M_0 = \frac{3}{2}M_e$.

By substituting equation (10) for M into (11a), a quadratic equation for P is obtained, which is easily solved to give the load P_{el} exhausting the elastic resources of the strut and bounding from below its load bearing capacity P_C. The same procedure applied with (11b) replacing (11a) provides an *upper bound* to P_C. In fact, at collapse some fibers of the central section are still elastic (Fig. 6) and the corresponding $N - M$ point is *inside* the limit domain. Some collapse situations are indicated by dots in figure 5b (the representation is qualitative, the location of the points being influenced to some extent by the strut slenderness). Observe also that the expression (10) for the maximum bending moment looses its validity outside the elastic range.

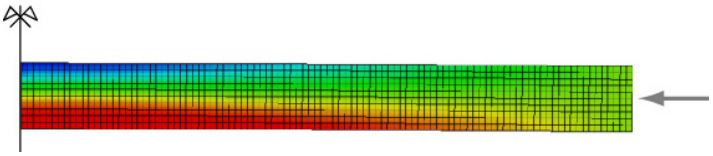


Fig. 6. Typical column at collapse: plastic strains develop in the red zone

A reasonable approximation to the collapse load is obtained by assuming that the $N - M$ points at collapse are located on the straight line

$$\frac{N}{N_0} + \frac{M}{M_0} = 1 \tag{12}$$

(dashed in figure 5b) and that the elastic relation (10) holds up to this point. One obtains

$$P_C = \frac{1}{2}\left(N_0 + P_E\left(1 + U\frac{N_0}{M_0}\right) - \sqrt{\left(N_0 + P_E\left(1 + U\frac{N_0}{M_0}\right)\right)^2 - 4N_0P_E}\right) \tag{13}$$

For rectangular $(b \times h)$ cross sections it is $N_0 = \sigma_0 bh$, $M_0 = \frac{1}{4}\sigma_0 bh^2$ and the column slenderness can be written as $\lambda = 2\sqrt{3}\frac{l}{h}$. By considering the slenderness ratio

$$\Lambda = \sqrt{\frac{N_0}{P_E}} = \frac{\lambda}{\lambda_0} \tag{14}$$

where

$$\lambda_0 = \pi\sqrt{\frac{E}{\sigma_0}} \tag{15}$$

is the *transition slenderness* (a material property) and by introducing the dimensionless imperfection measure

$$W = \frac{U}{l} \tag{16}$$

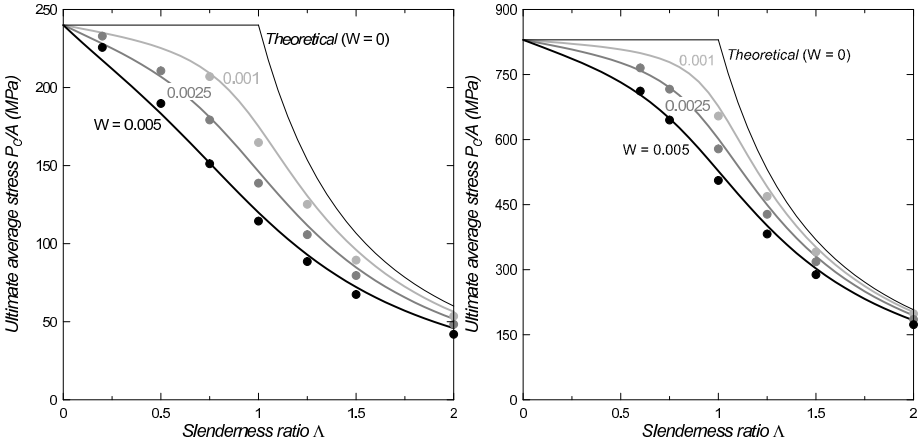
one can write equation (13) in the form

$$P_C = \frac{1}{2}\left(N_0 + P_E\left(1 + Z\right) - \sqrt{\left(N_0 + P_E\left(1 + Z\right)\right)^2 - 4N_0P_E}\right) \tag{17a}$$

with

$$Z = U\frac{N_0}{M_0} = \frac{2}{\sqrt{3}}\lambda_0 \Lambda W \tag{17b}$$

Equation (17) is a good approximation to the numerically computed failure load of compressed columns, as depicted in Fig. 7 where results for two materials with strongly different properties and a few initial imperfection magnitudes are compared. Several codes, including EUROCODE 3, base their recommendations on it (Dowling, 1990).



(a) INCONEL 690 ($\lambda_0 = 86.8$) (b) Titanium alloy ($\lambda_0 = 36.2$)

Fig. 7. Formula (17) vs computed results (dots). Λ defined by equation (14)

3.2 Interaction domain for an oval tube

Consider now a cylindrical shell with an initial imperfection controlled by W, equation (9), as illustrated in figure 8. Because of W, the external pressure q will cause, besides compressive hoop stresses, a bending moment with peak values given by the relation

$$M = M_I \frac{1}{1 - \frac{q}{q_E}} \tag{18a}$$

where

$$M_I = \frac{1}{8} q D^2 W \tag{18b}$$

is the value predicted within the small displacements framework (geometric linearity) and q_E is the Euler buckling pressure (1a). The expression (18) for M is exact in the elastic range if the initial imperfection has the same shape as the buckling mode (Timoshenko & Gere, 1961).
As for beam columns, the behavior of the tube wall can be interpreted on the basis of suitable interaction domains, with the external pressure q playing the role of the compression force and equation (18) replacing (10) to express the peak value of the bending moment. The domains are sketched in figure 8b: as in the equivalent picture for the strut, line 1 bounds the elastic

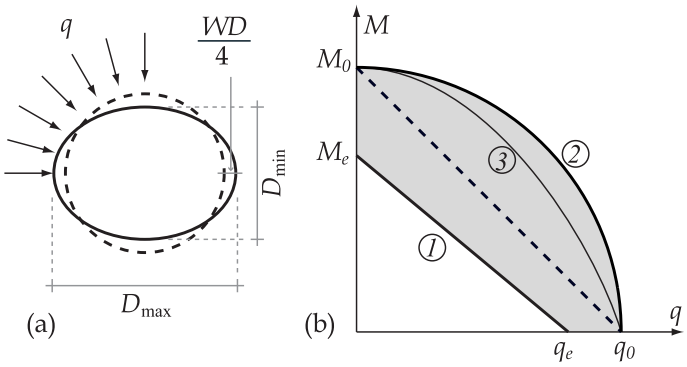


Fig. 8. Interaction domain for the tube cross section

region and line 2 is the limit curve. Reference values are assumed as follows (Corradi et al., 2008)

$$q_e = 2\sigma_0 \frac{t}{D}\left(1 - \frac{t}{D}\right) \qquad q_0 = 2\sigma_0 \frac{t}{D}\left(1 + \frac{1}{2}\frac{t}{D}\right) \tag{19a}$$

$$M_e = \frac{\sigma_0}{\sqrt{1-\nu+\nu^2}} \frac{\frac{1}{4}\left(b^2-a^2\right)^2 - a^2b^2\left(\ln\frac{b}{a}\right)^2}{2b^2\ln\frac{b}{a} - \left(b^2-a^2\right)} \qquad M_0 \approx \frac{\sigma_0 t^2}{2\sqrt{3}} \tag{19b}$$

where q_e, M_e and q_0, M_0 are the pressure and moment values that individually bring the tube at the onset of yielding and exhaust its load bearing capacity. $b = D/2$ and $a = b - t$ are the external and internal nominal radii, respectively.

The values above refer to materials governed by von Mises' criterion. Elastic stresses are computed from the well known plane solutions for a round cylinder under external pressure and for a curved beam subject to constant bending moments (Timoshenko & Goodier, 1951) and the values of q_e, M_e are obtained on this basis. q_0 is given by equation (1b), rewritten for completeness; the value $(19b)_2$ of M_0 actually refers to a straight beam and, for tubes thick enough to demand that curvature be considered, entails an error not completely negligible but acceptable: bending moments being caused by imperfections, only the portion of the domains close to the q axis is of interest.

The interaction domains for the tube and the strut of rectangular cross section exhibit some differences that become significant with increasing tube thickness. First of all, while the ratio M_0/M_e maintains more or less the value of 1.5, q_0 exceeds q_e by an amount that must be considered for $D/t < 25$. Moreover, in thick tubes the hoop stresses due to pressure are not uniform, which provides additional stress redistribution capabilities, so that the limit curve is expected to be external to that of the equivalent strut (the situation is sketched in figure 8b, where curve 3 (thinner) portrays the parabola (11b) with q replacing N). As a consequence, the region of partially yielded tubes (in gray), which contains the collapse situations, widens considerably, augmenting the uncertainties in estimating the failure pressure.

Nevertheless, the extension to tubes of the beam-column procedure is spontaneous and an attempt in this sense is made by introducing equation (18) for M in the linear expression

$$\frac{q}{q_0} + \frac{M}{M_0} = 1 \tag{20}$$

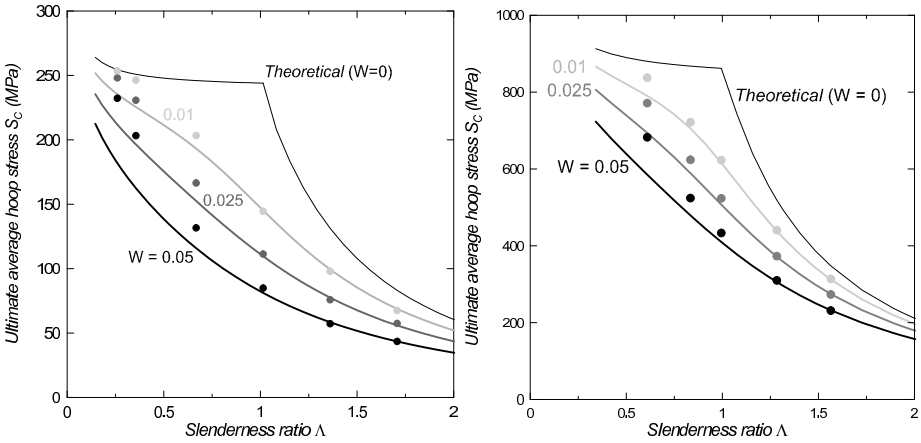
corresponding to the dashed segment in Fig. 8b. As for columns, a second order equation is obtained; its smallest root reads

$$q_C = \frac{1}{2}\left[q_0 + q_E\,(1+Z) - \sqrt{(q_0 + q_E\,(1+Z))^2 - 4q_0 q_E}\right] \tag{21a}$$

with

$$Z = \frac{\sqrt{3}}{2}\left(\frac{D}{t} + \frac{1}{2}\right) W \tag{21b}$$

The analogy with equation (17a) is immediately apparent.



(a) INCONEL 690 ($\kappa = 832$) (b) Titanium alloy ($\kappa = 150$)

Fig. 9. Formula (21) vs computed results (dots). Λ defined by equation (6)

The results provided by equation (21) are plotted in Fig. 9 (solid lines). Dots refer to the results computed in (Corradi et al., 2009), where indication on the assumptions made, the finite element model used and the solution procedure adopted can be found. For graphical purposes, the ultimate pressure is expressed in terms of the average hoop stress equation (2). Agreement is good for $\Lambda > 1$ but, as the thickness increases, lower bounds rather than good approximations are obtained. It can be concluded that thin or moderately thin tubes behave essentially as straight columns of rectangular cross section, but some fundamental aspects of the structural response change drastically as the thickness increases beyond a certain limit.
A first reason for this change is of purely geometric nature, i.e. it depends on the value of D/t only. For comparatively large values the tube wall behaves essentially as a straight column, but curvature increases with diminishing D/t and differences become more and more evident. Secondly, thick tubes exhibit stress redistribution capabilities that columns do not have and this provides additional resources in terms of overall strength. It must be observed that stress redistribution *capability* depends on D/t only, but the possibility of actually exploiting it is conditioned by tube slenderness Λ which, as equation (6) shows, depends on both D/t and the material properties subsumed by the dimensionless parameter κ, equation (8).
Fig. 9 indicates that the second effect is dominant. Because of the strong difference in κ, $\Lambda = 1$ corresponds to $D/t \approx 30$ for INCONEL 690 and to $D/t \approx 13$ for titanium alloy. The two

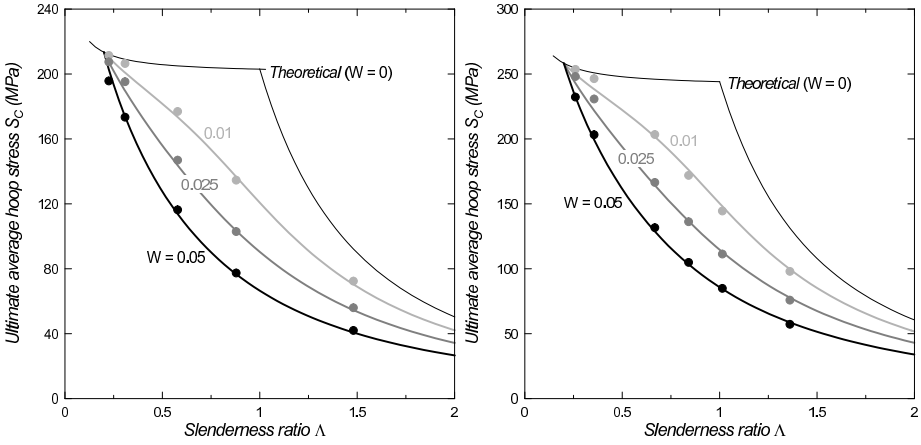
pictures do show some differences, but not as strong as the discrepancy between the two D/t values would suggest: computed results for titanium depart from formula predictions for a slightly greater Λ than for INCONEL, but the overall response seems to depend more on Λ as a whole than on D/t only.
In any case, equation (21a) provides conservative estimates for the failure pressure of thick tubes. In (Corradi et al., 2008) this result was considered effectively as a *lower bound* and a corresponding upper bound, consisting in the plastic collapse load of the ovalized tube computed by neglecting geometry changes, was associated to it. The two bounds were combined by introducing a suitable weighting factor, determined by fitting a number of computed results for tubes of different materials (including those listed in Table 1). As it was already mentioned, the procedure produced acceptable results, but it is felt that it could be both simplified and improved.

4. The proposed procedure

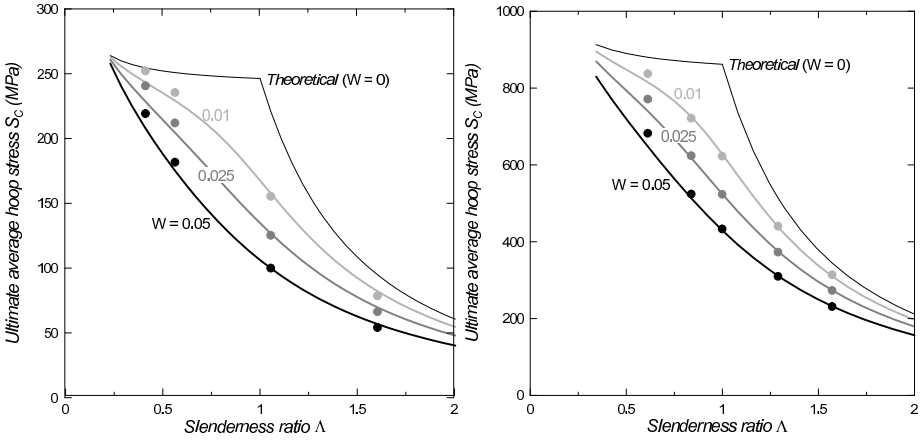
Both equations (17) for columns and (21) for cylindrical shells predict that the failure pressure coincides with the theoretical limit when the relevant parameter Z vanishes. This obviously occurs for any slenderness ratio when no imperfections are present ($W = 0$) but, independently of the presence of imperfections, both structures are expected to become stocky enough to make negligible interaction with buckling. In other words, it should be $Z = 0$ for any W when slenderness attains a sufficiently low value.
Equation (17b) in fact implies $Z \to 0$ for $\Lambda \to 0$, so that one obtains $P_C = N_0$ for infinitely stocky columns, independently of the imperfection amplitude. However, to give $Z = 0$ for $W \neq 0$, equation (21b) requires $\frac{D}{t} = -\frac{1}{2}$, a value with no physical meaning. This is another reason for the increasingly conservative nature of the approximation as Λ diminishes.
The remarks above suggest that the approximation can be improved by operating on the expression (21b) of Z so to make it vanish for sufficiently small D/t. A minimal choice is $D/t = 2$, the lowest possible value, which however turns out to be still too restrictive. Moreover, the discussion in the preceding section shows that, to obtain an approximation reasonably accurate for all materials, slenderness ratio Λ is preferable to D/t as a measure of the limit stockiness. On the basis of the numerical results in (Corradi et al., 2009), this can be reasonably identified with $\Lambda = 0.2$ and the corresponding values of D/t for different materials can be obtained by numerically solving equation (6). For the materials considered in Table 1, the resulting values, labeled as $(D/t)_{\Lambda=0.2}$, are listed in the last column.
The correction consists in replacing the term $(D/t + 1/2)$ with $(D/t - (D/t)_{\Lambda=0.2})$ in equation (21b). This improves the approximation for $\Lambda < 1$, but shifts the curves upward everywhere by an amount of some significance, even if not dramatic and diminishing with increasing Λ. For compensation, the imperfection amplitude is artificially increased by multiplying W by a factor that, empirically, was identified with 1.2. Thus, the expression for Z becomes

$$Z = \frac{\sqrt{3}}{2}\left(\frac{D}{t} - \left(\frac{D}{t}\right)_{\Lambda=0.2}\right) 1.2\, W \tag{22}$$

If equation (21b) is replaced by the expression above, equation (21a) produces the results depicted in Fig. 10 for the four materials listed in Table 1, covering a range of values of κ that can be regarded as exhaustive. Plots depart from the value of Λ corresponding to $D/t = 5$, which is different for different materials. Formula predictions show a very good agreement throughout with numerical results (dots). For materials with low values of κ (aluminum and



(a) Stainless steel ($\kappa = 1106$) (b) INCONEL 690 ($\kappa = 832$)



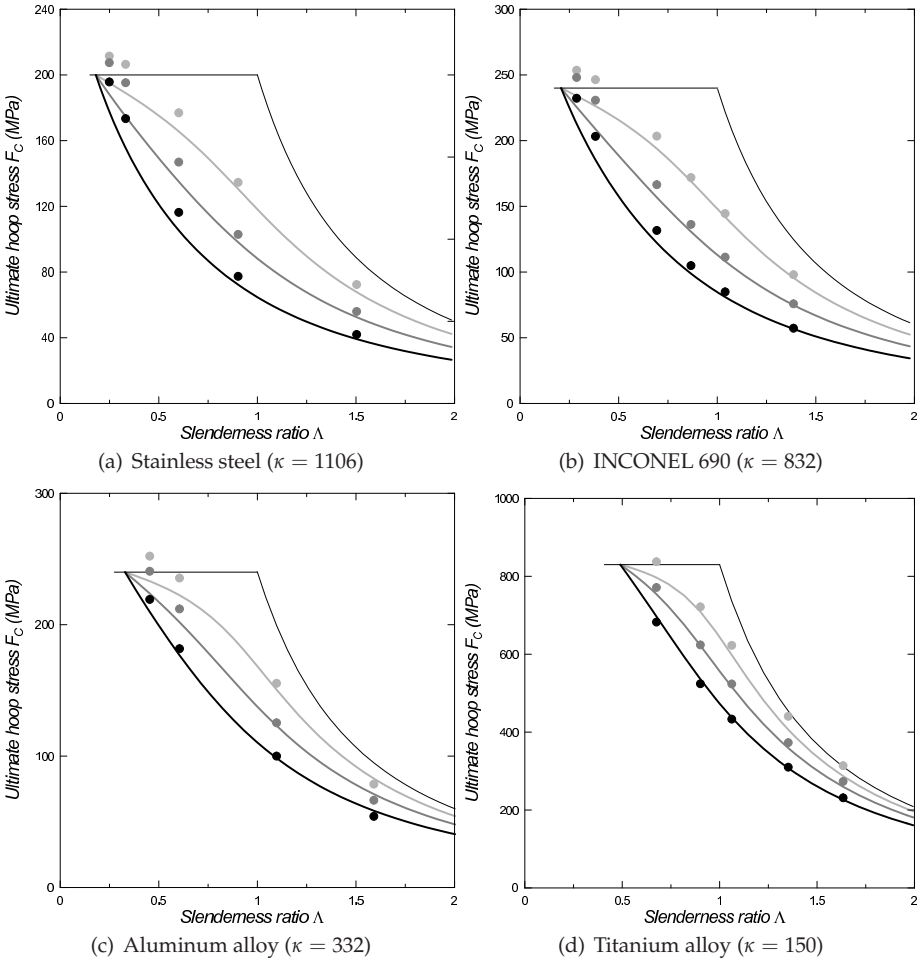
(c) Aluminum alloy ($\kappa = 332$) (d) Titanium alloy ($\kappa = 150$)

Fig. 10. Proposed formula vs computed results. Λ defined by equation (6)

titanium alloys) they remain a little conservative for stocky tubes, but improvement with respect to the unbridged formula is significant.

5. Shortcomings of thin shell approximation

In the formula above the theoretical limit values are defined by equations (1) and, as a consequence, the slenderness ratio by equation (6). Use of these expressions is mandatory in a context that includes thick and moderately thick tubes, in that they incorporate the effects of stress redistribution over the wall thickness, which were seen to be significant and which the "thin shell" equations (4), (7) ignore. Nevertheless, the latter expressions often are preferred and the implications of their use are worth exploring.



(a) Stainless steel ($\kappa = 1106$) (b) INCONEL 690 ($\kappa = 832$)

(c) Aluminum alloy ($\kappa = 332$) (d) Titanium alloy ($\kappa = 150$)

Fig. 11. Results obtained with thin shell approximation. Λ defined by equation (7)

Formally, modifications are straightforward. It suffices to replace in equation (21a) q_0 and q_E with p_0 and p_E, as defined by equations (4). One obtains

$$p_C = \frac{1}{2}\left[p_0 + p_E\,(1+Z) - \sqrt{(p_0 + p_E\,(1+Z))^2 - 4p_0 p_E}\right] \tag{23a}$$

The value of Z still could be given by equation (22), whith $(D/t)_{\Lambda=0.2}$ computed from equation (7). However, the choice for a limiting value of D/t associated to a material independent slenderness ratio is justified by the dominant effect of stress redistribution, associated with Λ. Thin shell approximation does not account for stress redistribution and the only cause of departure of the tube response from that of the straight column is the geometric curvature, so that a limiting value of D/t seems the most appropriate choice. An acceptable

compromise, valid for all materials, turns out to be $(D/t)_{\lim} = 6$ and one can write

$$Z = \frac{\sqrt{3}}{2}\left(\frac{D}{t} - 6\right) 1.2\, W \tag{23b}$$

The results provided by equations (23a) are depicted in Fig. 11. Results are not as accurate as those in Fig. 10, but still acceptable in the moderately thin and thin tube range. As well expected, predictions become grossly conservative with increasing thickness, which underlines the importance of properly accounting for stress redistribution. To assess the pressure bearing capacity of the tube, thin shell theory is adequate only to more than moderately thin tubes (i.e., thinner than those for which an elastic solution still is acceptable) and a formulation aiming at covering the entire slenderness range must consider more precise expressions. One cannot even claim that these shortcomings are compensated by greater simplicity: equations (23a) are not simpler; only they are based on more usual definitions.

6. Conclusions

Long cylindrical shells subjected to external pressure have been considered. The study was motivated by the necessity of assessing the collapse behavior of the moderately thick tubes involved by some recent nuclear power plant proposals, but tubes of any slenderness were considered, even if little attention was devoted to very thin tubes, which buckle when still elastic according to well known modalities and that do not need additional investigation.

In previous papers it was demonstrated that a reliable reference value for the pressure causing tube failure can be obtained by performing complete non linear finite element computations under suitable assumptions. Purpose of this study was the derivation of an accurate and simple formula permitting the definition of this value without performing numerical analyses. It does not seem too daring to state that this goal has been attained with equations (21a), (22): the formula is fairly simple and the results it provides are in good agreement with numerical outputs for different materials, imperfection amplitudes and slenderness ratios. Obviously, only a few materials, imperfections and slendernesses have been checked, but the range of parameters used is wide enough for this statement to be considered of general validity.

The formula can be used both for preliminary design purposes and as a reliable reference value for the definition of allowable working pressure. This second aspect, however, no longer is a must: since Code Case N-759 was approved, tubes can be sized adequately by using existing regulations and alternatives are not required. In the authors' opinion, however, this fact does not diminish the interest of the result achieved. The ingredients used to build the formula enlighten some aspects of the collapse behavior of moderately thick tubes, a range so far little explored. Tubes of intermediate slenderness fail because of interaction between buckling and plasticity, but differences show up at slendernesses about the transition value. Medium thin tubes behave essentially as straight columns and column formulas can be employed with straightforward modifications. As thickness increases, however, the geometry dependent effect of curvature and the slenderness dependent effect of stress redistribution enter the picture and the tube wall no longer behaves as a straight beam. To account for these aspects, a correction was introduced to the original formula. The accuracy of the consequent results can be taken as the indication that the fundamental aspects of the mechanical behavior are correctly represented.

7. Acknowledgment

Authors express their gratitude to Mr Giovanni Costantino and Mr Manuele Aufiero for performing some of the computations used in this study.

8. References

ASME (2007). Code Case N-759: Alternative rules for determining allowable external pressure and compressive stresses for cylinders, cones, spheres and formed heads, In: *BPVC-CC-N-2007 Nuclear Code Cases; Nuclear components*, ASME International, 2007. ISBN 0791830764.

Budiansky B. (1974). Theory of buckling and post-buckling behaviour of elastic structures. *Advances in Applied Mechanics*, Vol. 14, Issue C, 1974, 1-65, ISSN 0065-2156.

Carelli M.D.; Conway L.E.; Oriani L.; Petrovic B.; Lombardi C.V.; Ricotti M.E.; Baroso A.C.O.; Collado J.M.; Cinotti L.; Todreas N.E.; Grgic D.; Moraes M.M.; Boroughs R.D.; Ninokata H.; Ingersoll D.T. & Oriolo F. (2004). The design and safety features of the IRIS reactor. *Nuclear Engineering and Design*, Vol. 230, No. 1-3, May 2004, 151-167, ISSN 0029-5493.

Carelli M.D. (2009). The exciting journey of designing an advanced reactor. *Nuclear Engineering and Design*, Vol. 239, No. 5, May 2009, 880-887, ISSN 0029-5493.

Cinotti L.; Bruzzone M.; Meda N.; Corsini G.; Lombardi C.V.; Ricotti M. & Conway L.E. (2002). Steam generator of the International Reactor Innovative and Secure, *Proceedings of the Tenth International Conference on Nuclear Engineering (ICONE)*, Arlington, VA, Paper No. ICONE10-22570.

Corradi L.; Luzzi L. & Trudi F. (2005). Collapse of thick cylinders under radial pressure and axial load. *ASME Journal of Applied Mechanics*, Vol. 72, No. 4, July 2005, 564-569, ISSN 0021-8936 .

Corradi L.; Ghielmetti C. & Luzzi L. (2008). Collapse of thick tubes pressurized from outside: an accurate predictive formula. *ASME Journal of Pressure Vessel Technology*, Vol. 130, No. 2, May 2008, Paper 021204_1-9, ISSN 0094-9930.

Corradi L.; Di Marcello V.; Luzzi L. & Trudi F. (2009). A numerical assessment of the load bearing capacity of externally pressurized moderately thick tubes. *International Journal of Pressure Vessels and Piping*, Vol. 86, No. 8, August 2009, 525-532, ISSN 0308-0161.

Dowling P.J. (1990). New directions in European structural steel design, *Journal of Constructional Steel Research*, Vol. 17, No. 1-2, July 1990, 113-140, ISSN 0143-974X.

EUROCODE 3 (1993). Design of Steel Structures. Part 1.1: General Rules and Rules for Building, ENV 1-1, Brussels, 1993.

Haagsma S.C. & Schaap D. (1981). Collapse resistance of submarine pipelines studied. *Oil & Gas Journal*, No. 2, February 1981, 86-91, ISSN 0030 1388.

IAEA (2007). *Status of Small Reactor Designs without Onsite Refueling*, International Atomic Energy Agency, IAEA-TECDOC-1536, ISSN 1011-4289.

Ingersoll D.T. (2009). Deliberately small reactors and the second nuclear era, *Progress in Nuclear Engineering*, Vol. 51, No. 4-5, May-July 2009, 589-603, ISSN 0149-1970.

Karahan A. (2010). Possible design improvements and a high power density fuel design for integral type small modular pressurized water reactors, *Nuclear Engineering and Design*, Vol. 240, No. 10, October 2010, 2812-2819, ISSN 0029-5493.

Lo Frano R. & Forasassi G. (2009). Experimental evidence of imperfection influence on the buckling of thin cylindrical shells under uniform external pressure. *Nuclear Engineering and Design*, Vol. 239, No. 2, February 2009, 193-200, ISSN 0029-5493.

Luzzi L. & Di Marcello V. (2011). Collapse of nuclear reactor SG tubes pressurized from outside: the influence of imperfections. *ASME Journal of Pressure Vessel Technology*, Vol. 133, No. 1, February 2011, Paper 011206_1-6, February 2011, ISSN 0094-9930.

Mendelson A. (1968). *Plasticity: Theory and Applications*, McMillan, New York.

Ninokata H. (2006). A comparative overwiev of the thermal hydraulic characteristics of integrated primary systems nuclear reactors, *Nuclear Engineering and Technology*, Vol. 38, No. 1, February 2006, 33-44, ISSN 1738-5733.

Tamano T.; Mimaki T. & Yanagimoto S. (1985). A new empirical formula for collapse resistance of commercial casing, *Nippon Steel Technical Report*, No. 26, 1985, 19-26, ISSN 0300 306X.

Timoshenko S. & Gere J.M. (1961). *Theory of Elastic Stability*, McGraw-Hill, New York.

Timoshenko S. & Goodier J.N. (1951). *Theory of Elasticity*, McGraw-Hill, New York.

Yeh M.H. & Kyriakides S. (1988). Collapse of deepwater pipelines. *ASME Journal of Energy Resources Technology*, Vol. 110, No. 1, March 1988, 1-11, ISSN 0195-8954.

15

Resistance of 10GN2MFA-A Low Alloy Steel to Stress Corrosion Cracking in High Temperature Water

Karel Matocha[1], Petr Čížek[1], Ladislav Kander[1] and Petr Pustějovský[2]
[1]*Material & Metallurgical Research, Ltd., Ostrava,*
[2]*VÍTKOVICE Heavy Machinery, J.S.C., Ostrava,*
Czech Republic

1. Introduction

In the second half of the last century the pressure vessel and the cold and hot collector bodies, ranking among the most stress and corrosion exposed components of the WWER 1000 (Water-Water Energy Reactor) horizontal steam generator (SG), were manufactured of the low alloy steel of type 10GN2MFA, the chemical composition of which is shown in tab. 1 (IAEA-EBP-WWER-07, 1997).

C	Mn	Si	S	P	Cr	Ni	Mo	V	Cu
0,08-0,12	0,80-1,10	0,17- 0,37	≤ 0,020	≤0,020	≤0,030	1,8-2,3	0,40-0,70	0,03-0,07	≤0,30

Table 1. Chemical composition of 10GN2MFA steel in line with standard element mass content, [%]

In the period from 1986 to 1995 cracks due to stress corrosion cracking (SCC) have been revealed after 7000 to 60 000 hours of operation in the ligaments between tube holes on the cold collectors in 25 steam generators at 9 units operated in the former Soviet Union (WWER SC 076,1993; Matocha, Wozniak, 1996, IAEA-EBP-WWER-07, 1997).
The cause analysis of collector cracking, carried out in the former Soviet Union, showed that the steel used, in all cases where cracking was experienced, has been produced by the open hearth furnace process (IAEA-EBP-WWER-07, 1997). Open heart furnace melted collector metal had local impurity concentrations, including manganese sulphide, which have deleterious effect on SCC resistance of low alloy steels in high temperature water environment.
Since the beginning of 1989 a number of modifications have been introduced in the former Soviet Union in the steam generators or in their operational conditions process (IAEA-EBP-WWER-07, 1997). One of them was the modification of the steelmaking process. Only steel produced by electroslag remelting has been used for the manufacturing of collector bodies. The electroslag remelting process (ESR) allows to enhance considerably micro-cleanliness of the steel and consequently the resistance of the 10GN2MFA steel to SCC in high temperature water environment.
In the period from 1991 to 1994, the eight steam generators were manufactured in VÍTKOVICE, J.S.C. for WWER 1000 Temelín NPP. The collector bodies of Temelín NPP

steam generators were made of doubly vacuum treated 10GN2MFA steel (first on DH equipment and then by pouring in vacuum) so as to minimize the gas concentration and to secure a homogeneous chemical composition. Steps were taken in all heats to keep down the content of impurity elements. The typical chemical composition of the steel is shown in tab. 2 (WWER SC 076, 1993).

C	Mn	Si	S	P	Cr	Ni	Mo	V	Cu
0,10	0,86	0,27	0,008	0,009	0,19	2,16	0,43	0,05	0,08

Table 2. The typical chemical composition of 10GN2MFA steel of Temelín NPP collector bodies [mass %]

No occurrence of environmentally assisted cracking has been noticed till now after approximately 79 000 hours of operation of the steam generators at Temelín NPP.
According to contemporary Russian technical specifications TU 0893-014-00212179-2004, the WWER 1000 SG collector body must be manufactured of the electroslag remelted 10GN2MFA steel (10GN2MFA-S). Sulphur content must be kept lower than 0,005% and the concentration of phosphorus must be lower than 0,008%. The requirement for the chemical composition and tensile properties of the 10GN2MFA-S steel are summarized in tab.3 and tab.4.

	C	**Mn**	**Si**	**P**	**S**	**Cu**	**Ni**	**Cr**	**Mo**	**V**	**Ti**	**Al**
min.	0,08	0,80	0,17	-	-	-	1,80	-	0,40	0,03	-	0,005
max.	0,12	1,10	0,37	0,008	0,005	0,30	2,30	0,30	0,70	0,07	0,015	0,035

Table 3. The requirements for the composition of the steel used for the manufacture of the WWER 1000 collector body.

Test temperature	Yield point	Yield strength	Tensile strength	Elongation	Reduction of Area
+20°C	345-590 MPa		540-700 MPa	18%	60%
350°C		min. 295 MPa	min. 490 MPa	15%	55%

Table 4. Requirements for the tensile properties of the 10GN2MFA steel.

In the course of the nineties of the last century, the significantly modernised technology of the 10GN2MFA steel production (electric arc furnace, refining in ladle furnace, vacuum degassing, bottom pouring under argon protection into mould) was established in VÍTKOVICE Heavy Machinery, J.S.C. It enables to guarantee the content of sulphur lower than 0,005% and the content of phosphorus lower then 0,008%.
However Vítkovice Heavy Machinery J.S.C. does not own the electroslag remelting plant. To be able to manufacture the WWER 1000 collector bodies for AES 2006 reactor plant, it was necessary to prove that there is no difference between electroslag remelted 10GN2MFA steel (10GN2MFA-S) and the steel manufactured by modernised VÍTKOVICE Heavy Machinery technology (10GN2MFA-A) from the point of view of fracture behaviour and resistance to stress corrosion cracking in high temperature water environment. For this purpose three forgings of WWER 1000 SG collector body were manufactured in VÍTKOVICE Heavy Machinery, J.S.C. (see Fig.1).

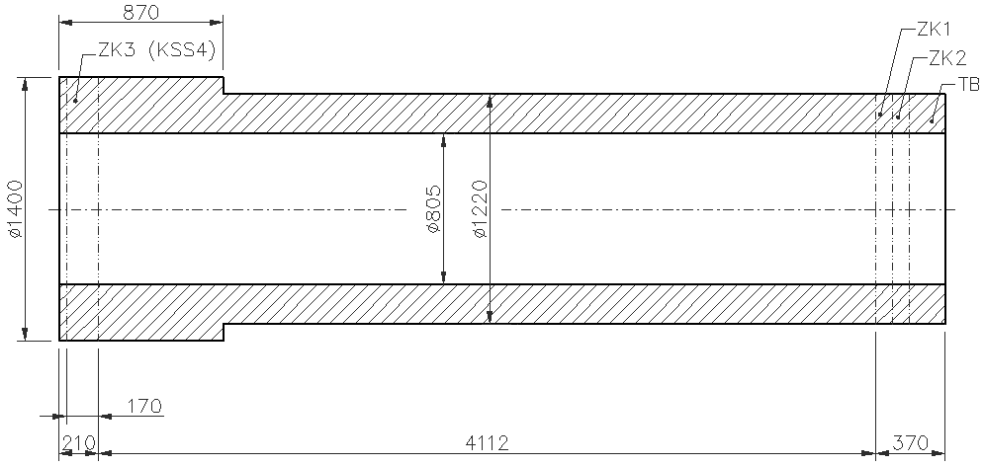


Fig. 1. Dimensions of the collector body after rough machining

This work summarizes the results of the evaluation of the tensile properties, fracture behaviour and resistance to stress corrosion cracking in high temperature water environment of the collector body forging No 1 material. The resistance to SCC in high temperature water environment was evaluated both in laboratories of MATERIAL & METALLURGICAL RESEARCH, Ltd. (M&MR, Ltd.) and also in laboratories of CNIITMASH, a joint-stock company, Moscow (Russia). The testing procedure used was proposed by CNIITMASH. The obtained results were compared with the results published by CNIITMASH for the 10GN2MFA-S steel (Yukhanov,V.A. & co-workers, 2009, Yukhanov,V.A. & co-workers, 2010).

2. Testing material and experimental techniques

The testing ring ZK1 (see Fig.1) after simulated post weld heat treatment was used as a testing material. Its chemical composition is shown in tab.5.

C	Mn	Si	P	S	Cu	Ni	Cr	Mo	V	Ti	Al
0,12	0,90	0,22	0,006	0,002	0,05	1,88	0,048	0,55	0,042	0,004	0,021

Table 5. The chemical composition of the testing material.

Metallographic analysis proved very low content of non-metallic inclusions of globular shape. The fine grained microstructure of the steel is bainitic with the average initial austenitic grain size G= 6-7.
Fig.2 illustrates schematically the location and orientation of the test specimens used for the determination of tensile and fracture characteristics (impact tests, fracture mechanics tests) and the round bar tests used for the evaluation of resistance to SCC by slow strain rate test.

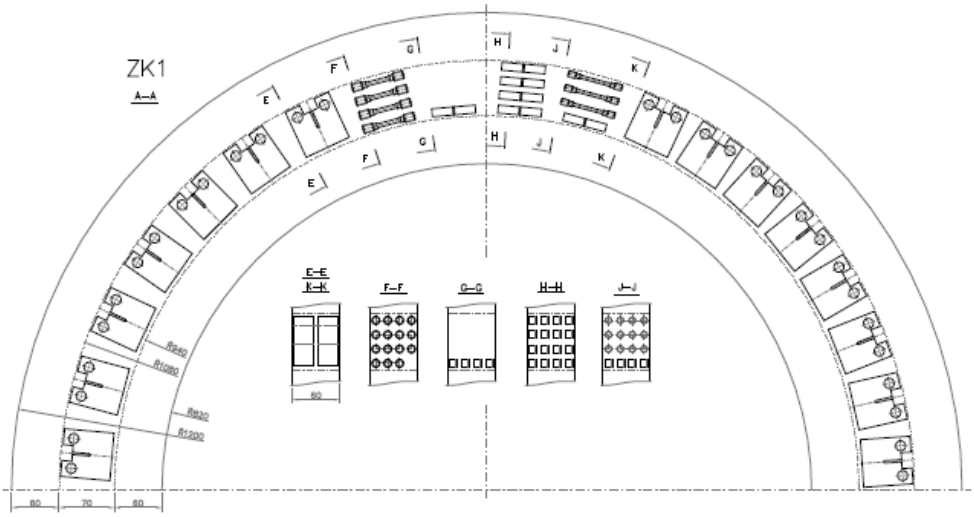


Fig. 2. Orientation of the test specimens in the testing block

As the characteristic feature of the fracture behaviour of 10GN2MFA steel at temperatures from 20°C to 320°C was found to be a ductile stable crack growth (Matocha,K., Wozniak,J., 1995, Matocha,K., Rožnovská,G. & Hanus, V.,2007), the variation in J with crack growth (J-R curve) was investigated at 290°C using multiple specimen method . 1CT (compact tension, 25 mm in thickness) specimens were used for the determination of J_{IC}. Tests were carried out in a stroke control at the speed of 0,5 mm/min. in accordance with GOST 25.506-85 on MTS 500 kN servo-hydraulic testing machine. Loading of the test specimens was realized in the three zone furnace specially developed for the determination of the fracture toughness at elevated temperatures. The design of the furnace enables to measure the crack mouth opening during the test (see Fig.3).

The slow strain rate tests of both round bar test specimens 5 mm in diameter and fatigue pre-cracked 1CT test specimens (ISO 7539-7:2005, ISO 7539-9:2003) were used for the evaluation of the resistance of the 10GN2MFA-A steel to SCC in high temperature water environment (260°C, 290°C). The testing parameters for the evaluation of resistance of the steel against stress corrosion cracking (strain rates, initial dissolved oxygen content) were proposed by CNIITMASH.

SCC tests were performed in static autoclave 11 l in volume fitted with INOVA servo-hydraulic testing machine. Slow strain rate tests of round bar test specimens were carried out under stroke control in demineralised water at temperatures 260°C and 290°C at strain rate $1{,}4.10^{-7}$ s^{-1}. Initial concentrations of dissolved oxygen 0,5 ppm, 1,5 ppm and 4,5 ppm were obtained by adding 50% H_2O_2 into demineralised water which was got rid of carbon dioxide. The susceptibility to SCC is evaluated on the basis of a comparison of reduction of area determined in air and in water environment at the same testing temperature. Slow strain rate tests of the fatigue pre-cracked 1CT specimens were carried out also under stroke control only at 260°C at initial concentration of dissolved oxygen 4,5 ppm and stroke rate $1{,}8.10^{-6}$ mm/s.

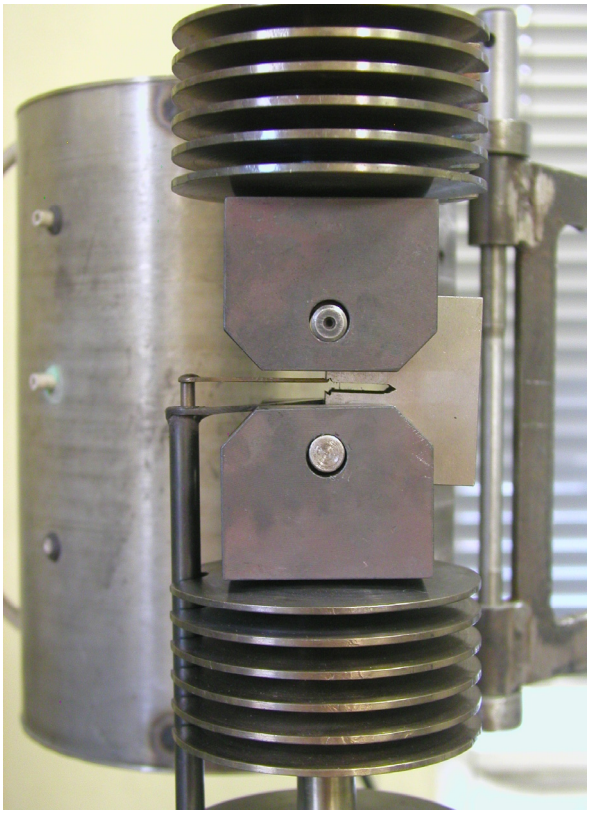


Fig. 3. The testing facility used for the determination of J-R curve in air environment at 290°C

3. Results and discussion

Summary of tensile characteristics is shown in table 6. Tensile tests at +20°C, 260°C, 290°C and 350°C were carried out on MTS 100 kN servo-hydraulic testing machine using round bar test specimens 5 mm in diameter like the test specimens used for the slow strain rate tests in high temperature water environment.

Test temperature [°C]	Yield point [MPa]	Yield strength [MPa]	Tensile strength [MPa]	Elongation [%]	R.A. [%]
+20	512		615	25,9	78
260		458	609	20,5	74
290		460	606	24,0	75
350		442	567	23,4	76

Table 6. Results of tensile tests of the studied steel at ambient and elevated temperatures.

Fig.4 and Fig.5 shows the temperature dependences of impact fracture energy and shear fracture area of the steel investigated the steel investigated in temperature range from -80°C to 320°C.

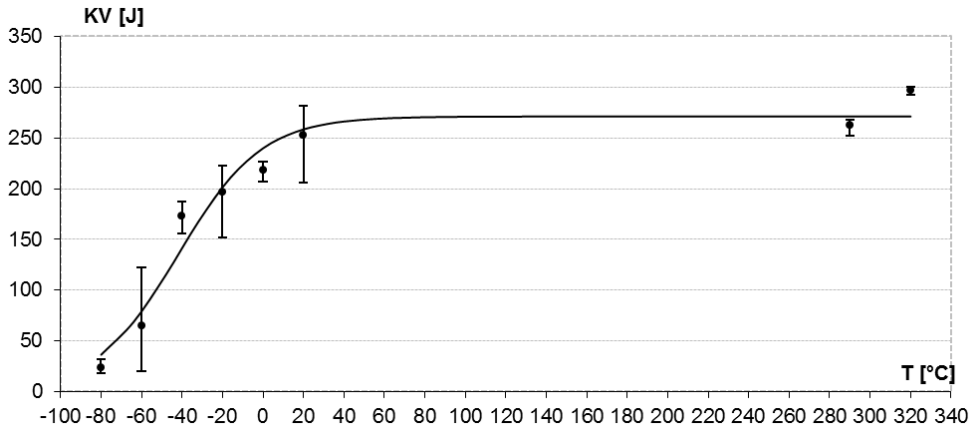


Fig. 4. Temperature dependence of impact fracture energy in the temperature range from -80°C to 320°

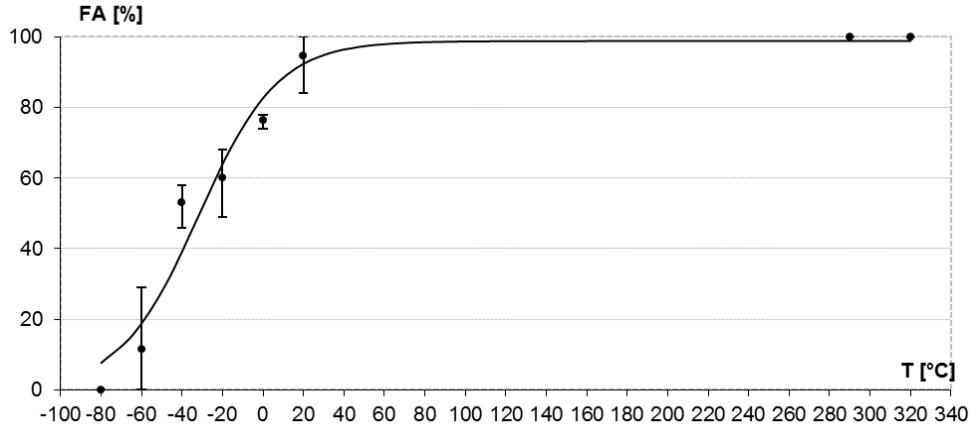


Fig. 5. Temperature dependence of shear fracture area in the temperature range from -80°C to 320°

Critical temperature of brittleness and FATT determined from the results of impact tests were found to be T_{k0} = -50°C and FATT = -31°C. Results of tensile and impact tests satisfy the requirements of the technical specifications TU 0893-014-00212179-2004.
Table 7 summarizes the results of fracture mechanics tests of the studied steel in air at 290°C performed in accordance with GOST 25.506-85. Fig.6 shows the J-R curve obtained. Fracture toughness J_{IC} determined from the J-R curve equals to J_{IC} = 237 N/mm.

No of test specimen	a_o [mm]	P_i [N]	A_{pi} [N.mm]	Δa [mm]	J_i [N/mm]
ZK1-40	26,0	74 671	211 395,1	1,68	629,8
ZK1-41	25,7	73 103	102 362,1	0,53	336,4
ZK1-42	25,6	75 662	144 246,0	0,78	447,2
ZK1-43	25,8	71 759	78 031,2	0,39	273,2
ZK1-44	25,8	72 069	172 442,1	1,28	517,2
ZK1-45	25,7	69 036	52 448,2	0,19	200,6
ZK1-47	25,9	71 008	242 460,2	1,80	700,7
ZK1-48	25,6	75 152	312 651,9	2,09	874,3

Table 7. Results of fracture mechanics tests in air at 290°C

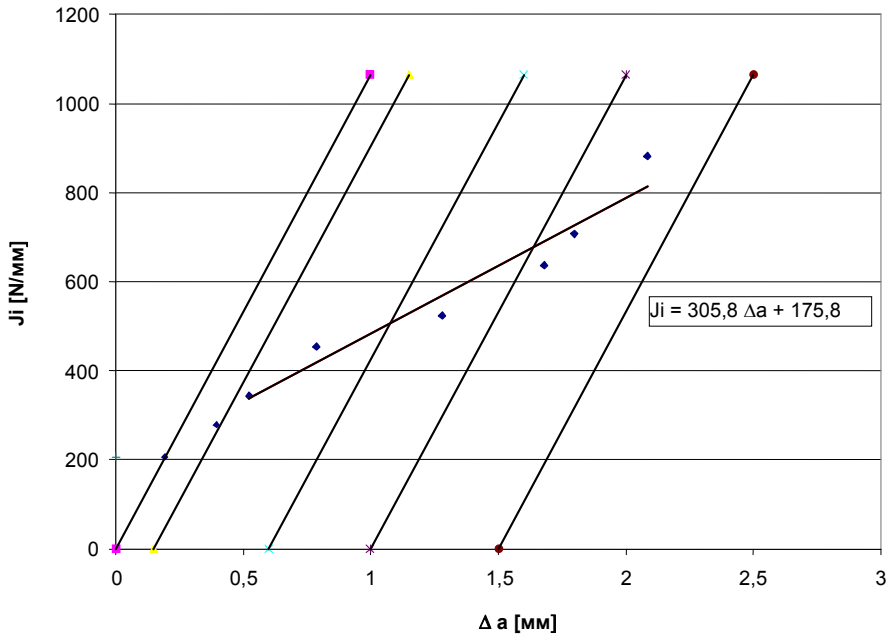


Fig. 6. J-R curve of the studied steel at 290°C determined by multiple test specimen method

This value is in a very good agreement with J_{IC} determined in air at 290°C for 10GN2MFA steel used for the manufacture of the collector bodies of Temelín NPP steam generators (Matocha,K., Wozniak,J., 1995).

Results of slow strain rate tests of round bar test specimens, made of 10GN2MFA-A steel, in high temperature water environment performed in both laboratories are summarized in table 8.

Test temperature	Initial concentration of dissolved oxygen	CNIITMASH	M&MR
		Reduction of area [%]	Reduction of area [%]
260°C	4,5 ppm	9,3	8,6
		9,6	12,3
		8,7	12,4
	1,5 ppm	69,9	> 54,1
		67,5	68,9
		66,5	
	0,5 ppm	76,3	73,2
		77,1	71,1
290°C (M&MR)	1,5 ppm		75,2
300°C (CNIITMASH)	4,5 ppm	74,8	77,0

Table 8. Results of slow strain rate tests of 10GN2MFA-A steel carried out in CNIITMAS and in M&MR, Ltd.

Table 8 shows a very good agreement between the results obtained in both laboratories for all initial concentrations of dissolved oxygen investigated. Fracture surfaces of the failured test specimens were examined by scanning electron microscope JEOL JM-5510. Fracture surfaces of the test specimens having low level of reduction of area showed the same fractographic features formerly observed on fracture surfaces of 1CT tests specimens tested in aerated distilled water (see Fig.7), (Matocha,K., Wozniak,J., 1995).

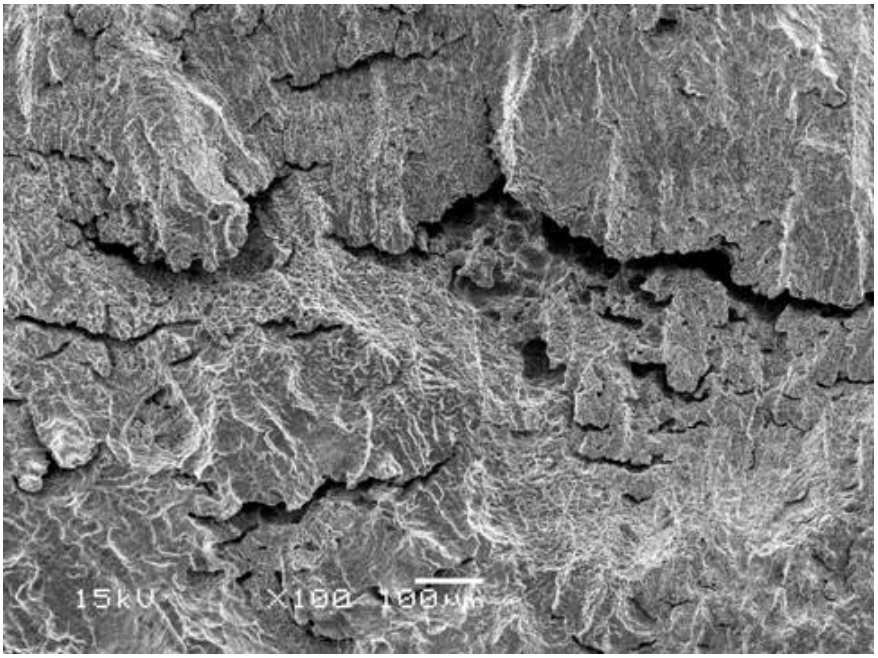


Fig. 7. Morphology of the fracture surface of the test specimen tested in water environment at 260°C (initial O_2 = 4, 5 ppm, R.A. = 12, 4%)

Metallographic evaluation of test specimen cross sections showed that the transverse microcracks observed on the fracture surface (see Fig.7) were initiated in the process zone ahead of the growing crack and probably contributed to the increase of crack growth rate (see Fig. 8), (Matocha,K., Rožnovská,G. & Hanus,V, 2007). In all cases the stress corrosion cracks were initiated from the pits (see Fig.9).

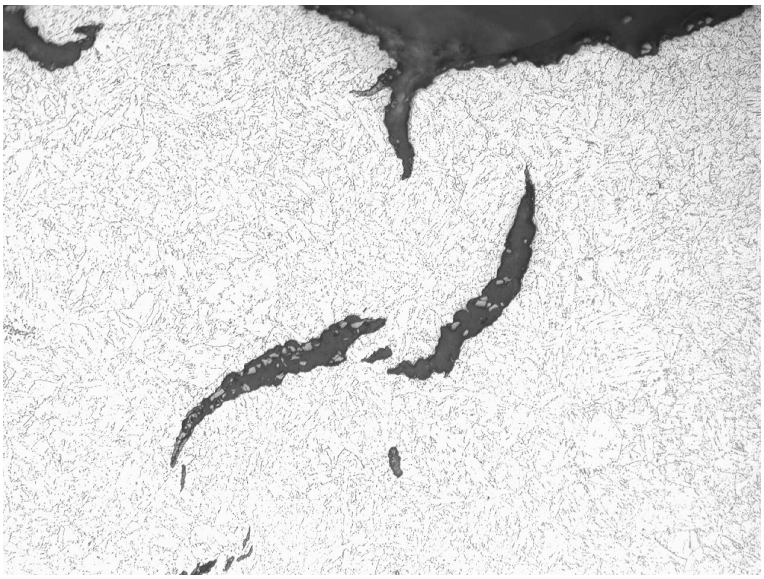


Fig. 8. The occurrence of transverse microcracks on fracture surface and in process zone of the growing crack (Matocha,K., Rožnovská,G. & Hanus,V, 2007).

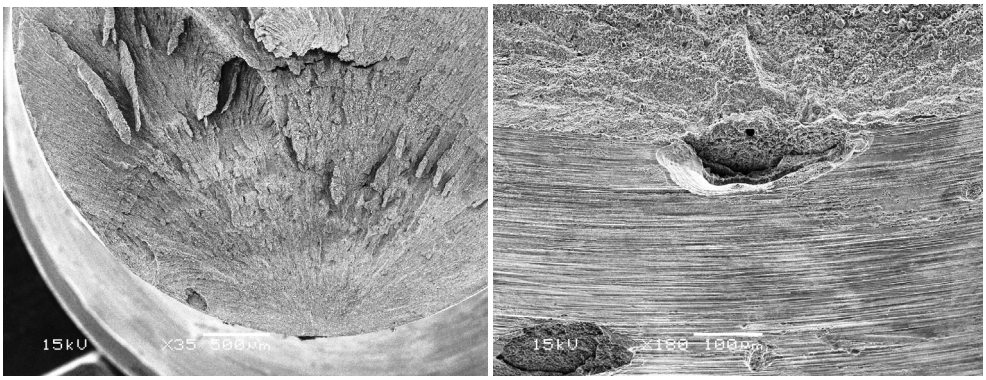


Fig. 9. Fracture surface of the test specimen tested in water environment at 260°C (initial O_2 = 4,5 ppm, R.A. = 12,4%)

Table 9 shows the comparison of the slow strain rate tests results obtained in CNIITMASH using round bar specimens manufactured of 10GN2MFA-A and 10GN2MFA-S steels at a strain rate $1{,}4.10^{-7}$ s^{-1}.

Test temperature	Initial concentration of dissolved oxygen	10GN2MFA-A	10GN2MFA-S.
		Reduction of area [%]	Reduction of area [%]
260°C	4,5 ppm	9,3	7,8
		9,6	8,7
		8,7	
	1,5 ppm	69,9	54,9
		67,5	
		66,5	
	0,5 ppm	76,3	72,3
		77,1	
300°C	4,5 ppm	74,8	75,4

Table 9. Results of slow strain rate tests of the 1OGN2MFA-A and 10GN2MFA-S performed in CNIITMASH.

The results obtained proved that the resistance of both steels against the stress corrosion cracking in high temperature water environment, evaluated by slow strain rate tests of round bars test specimens, is almost identical.
Results of slow strain rate tests of fatigue precracked 1C(T) specimens carried out in water environment at 260°C, initial dissolved oxygen content 4,5 ppm ppm and stroke rate $1{,}8.10^{-6}$ mm/s are summarized in table 10. The variation in δ (crack tip opening displacement) with crack advance Δa was investigated using a multiple specimen method. Fracture surfaces created by stable crack growth during autoclave tests were examined by scanning electron microscope JEOL JM-5510.

Load line displ. [mm]	v_{pl} [mm]	P_{max} [N]	δ [mm]	K_δ [MPa.m$^{1/2}$]	Δa [mm]
1,66	0,1	36595	0,053	100,7	3,05
1,16	0	28970	0,019	60,3	0
1,13	0	30160	0,019	59,7	0,51
1,09	0,06	27213	0,031	79,6	0
1,18	0,04	27810	0,027	74,5	1,74
1,43	0,09	31783	0,048	98,7	2,94
1,05	0,04	25428	0,024	69,6	0,35
1,86	0,33	23865	0,134	165,4	8,66
0,90	0,02	21462	0,015	54,7	0,68
1,19	0,05	26857	0,029	77,3	1,87
1,42	0,05	30003	0,033	82,8	3,27
0,78	0	16370	0,005	32,7	0
1,02	0,02	23913	0,017	59,0	0,93

Table 10. Results of slow strain rate tests of 1CT specimen in demineralised water at 260°C and initial dissolved oxygen content 4,5 ppm.

Fig. 10 shows the variation of K_δ with crack advance Δa. From the variance analysis of the results obtained follows that for the probability of 95% $K_{\delta in} = 54 \pm 16$ MPa.m$^{1/2}$. The average calculated environmentally assisted crack growth rate was found to be $v_{cor} = 1{,}8.10^{-5}$ mm/s. Fractographic analysis of the fracture surfaces created by SCC in high temperature water revealed the same fractographic features (see Fig.11) observed on fracture surfaces of round bar test specimens tested at the same initial dissolved oxygen concentration.

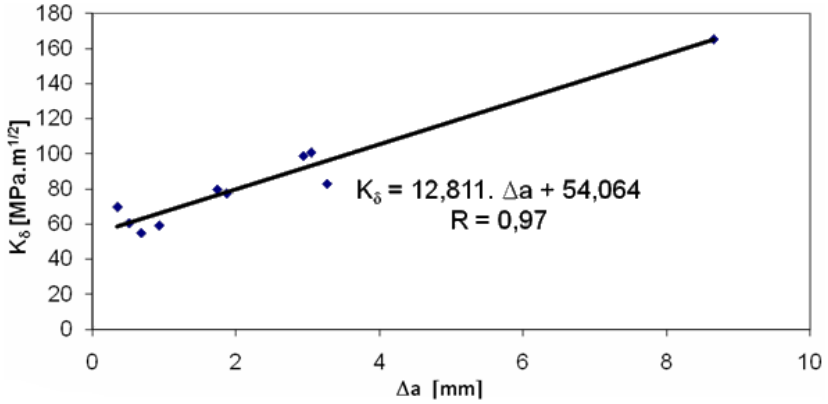


Fig. 10. The variation of K_δ with crack advance Δa (initial O_2 = 4,5 ppm, stroke rate $1{,}8.10^{-6}$ mm/s).

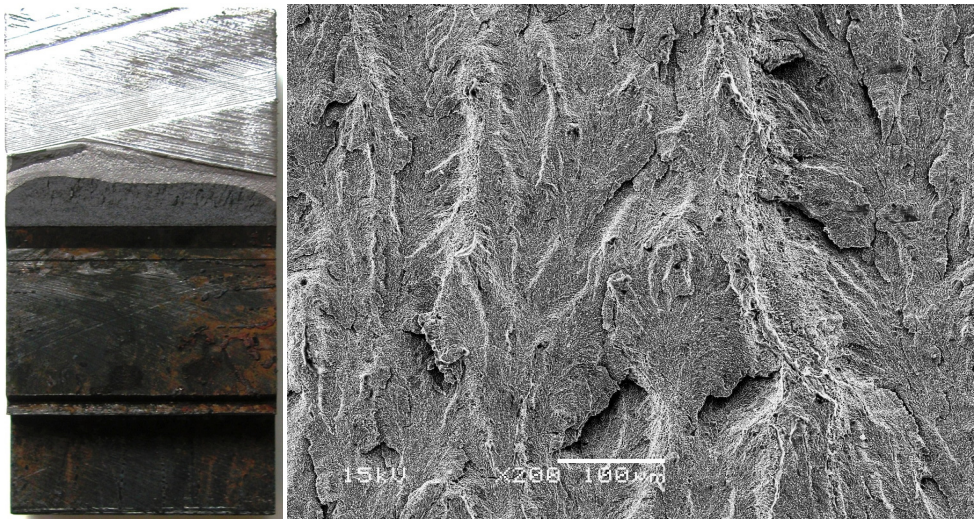


Fig. 11. Morphology of the fracture surface created by SCC, t = 260°C, O_2 = 4,5 ppm, $v_{cor} = 3{,}7.10^{-5}$ mm/s

4. Conclusion

From the results obtained in this study it follows that:

1. Results of tensile and impact tests of the 10GN2MFA-A steel manufactured by VÍTKOVICE Heavy Machinery, J. S. C., satisfy the requirements of the Russian technical specifications TU 0893-014-00212179-2004.
2. The resistance of 10GN2MFA-A and 10GN2MFA-S steels against the stress corrosion cracking in high temperature water environment with initial dissolved oxygen contents 0,5 ppm, 1,5 ppm and 4,5 ppm was found to be almost identical.

3. Stress corrosion cracks observed on the fracture surfaces of the round bar test specimens failured due to slow loading in high temperature water with increased initial dissolved oxygen content were in all cases initiated from the corrosion pits.
4. The crack growth in high temperature water with increased initial dissolved oxygen content was accompanied by the significant occurrence of transverse micro-cracks initiated in the process zone ahead of the growing crack.
5. Rates of corrosion crack growth determined by slow strain rate tests are typical for stress corrosion cracking of low alloy bainitic steels in high temperature water environment with increased initial dissolved oxygen content.

5. Acknowledgement

This paper was created in the project No.CZ.1.05/2.100/01.0040 "Regional Materials Science and Technology Centre" within the frame of the operation programme "Research and Development for Innovation" financed by the Structural Funds and from the state budget of the Czech Republic.

6. References

WWER-1000 Steam Generator Integrity. A Publication of the Extrabudgetary Programme on the Safety of WWER and RBMK Nuclear Power Plants. IAEA-EBP-WWER-07, July 1997.

Report of the Consultants Meeting on Steam generator Collector Integrity of WWER-1000 Reactors held in Vienna, Austria, 17 to 21 May 1993. Working Material IAEA, WWER SC 076, December 1993.

Matocha,K., Wozniak,J. Analysis of WWER 1000 Collector Cracking Mechanisms. Proceedings of the IAEA Specialists Meeting on Steam Generator Repair and Replacement: Practices and Lessons Learned. Ostrava, Czech Rep. 15-18 April, 1996, p. 284. IWG-LMNPP-96/1.

Yukhanov,V.A., Kazantzev,A.G., Karina,I.L., Šur,A.D. CNIITMASH report on the evaluation of material properties of collector body forgings manufactured in VÍTKOVICE, j. s.c. and the elaboration of conclusions about the possibility to use them for the manufacture of steam generators for AES 2006 reactor plant. Moscow, 2009 (in Russian).

Yukhanov,V.A., Kazantzev,A.G., Karina,I.L., Matocha,K., Čížek,P., Kavalec,M. & Korčák, A. Development and evaluation of a new modification of vessel steel 10GN2MFA-A for PGV-1000 steam generator collectors. Proceedings of the 8-th International Seminar on Horizontal Steam Generators, Podolsk, Russia, 19-21 May 2010 (CD ROM).

Matocha,K., Wozniak,J. Analysis of WWER 1000 Collector Cracking Mechanisms. Topical meeting on WWER-1000 Steam Generator Integrity,Tokyo,Japan, 25-29 September 1995.

Matocha,K., Rožnovská,G. & Hanus,V. The effect of lead on resistance of low aloy steel to SCC in high temperature water environment. *Corrosion issues in light water reactors. Stress corrosion cracking. Edited by D.Féron and J.-M.Olive.* European Federation of Corrosion Publications Number 51. Woodhouse Publishing Limited, Cambridge, England, 2007, ISSN 1354-5116.

ISO 7539-7:2005 Corrosion of metals and alloys – Stress corrosion testing – Part 7: Method for slow strain rate testing.

ISO 7539 – 9:2003 Corrosion of metals and alloys – Stress corrosion testing – Part 9: Preparation and use of pre-cracked specimens for tests under rising load or rising displacement.

GOST 25.506-85 Determination of fracture toughness (in Russia).

Part 3

Component Aging

16

Aging Evaluation for the Extension of Qualified Life of Nuclear Power Plant Equipment

Pedro Luiz da Cruz Saldanha[1,2] and Paulo Fernando F. Frutuoso e Melo[3]
[1]*Comissão Nacional de Energia Nuclear, CNEN – CGRC*
[2]*Associação Brasileira de Ensino Universitário- UNIABEU*
[3]*COPPE/UFRJ - Programa de Engenharia Nuclear*
Brazil

1. Introduction

In recent decades, the aging of nuclear power plants, the upgrading of safety systems and the concern for life extension of licensed plants close to completing 40 years of operation have been considered by regulatory agencies.

The extension of operating licenses for power reactors over 40 years has been a viable option for operators of nuclear power plants to ensure the adequacy of future capacity of power generation, in terms of economic benefits, compared with the construction of new power reactors.

Routine reviews of nuclear power plant operation (including modifications to hardware and procedures, significant events, operating experience, plant management and personnel competence) and special reviews following major events of safety significance are the primary means of safety verifications. Rereviews include an assessment of plant design and operation against current safety standards and practices, and they aim at ensuring a high level of safety throughout the plant's operating lifetime. They are complementary to the routine and special safety reviews and do not replace them (IAEA, 2009).

In 1991, the US Nuclear Regulatory Commission (NRC) issued rules and associated documentation describing how the licensee must demonstrate that the unit can continue operating for 20 years following the expiration of the 40-year license. These rules were established in 10CFR51 (NRC, 2007a), environmental protection requirements, and 10CFR54 (NRC, 2010a), technical requirements.

In 1994, the International Atomic Energy Agency (IAEA) issued recommendations, guidance and associated documentation describing how the licensee must demonstrate that the unit can continue operating through a systematic safety reassessment, named periodic safety review (PSR) (IAEA, 1994, 2003). The safety guide has the purpose of providing recommendations and guidance on the conduct of a PSR each 10 years for an existing nuclear power plant, and it is directed at plant operating organizations and regulators. In addition, some member states have initiated this reassessment to evaluate the cumulative effects of plant aging and plant modifications, operating experience, technical developments and siting aspects.

Considering this scenario, it is important to pursue knowledge and competence to assess the impact of aging and degradation mechanisms in systems and equipment for a nuclear plant, and from this knowledge, to decide on the plant acceptability considering the operational experience. The knowledge of these effects can indicate what action is applicable (renewal or repair), and then assist in corrective actions to be adopted.

The purpose of this chapter is to discuss and present an application of probabilistic models for the rate of occurrence of failures of active repairable systems using stochastic point processes, to decide for the extension of qualified life of equipment. The model application is within the context of the Plant Life Management (PLIM) and the Plant Life Extension (PLEX) or License Renewal of Nuclear Power Plants. The basic literature on PLIM and PLEX is (NRC, 2010b, 2010c, 2005; IAEA, 1994, 2002, 2003, 2004; Young 2009a, 2009b, 2009c).

The emphasis for this subject is the maintenance rule (NRC, 1997, 2000, 2007b; NEI, 1996). This rule is a prerequisite for the license renewal for US plants (Young, 2009c), and it provides an aging management tool for active equipment. Its performance criteria monitoring combined with the license renewal rule, equipment qualification, and life cycle management provides a sound basis for extending operation periods (Saldanha et al., 2001; Saldanha & Frutuoso e Melo, 2009).

The chapter is organized as follows. Section 2 discusses the aging concepts, aging management, equipment qualification, life extension and repairable systems. Section 3 presents an overview of the regulatory aspects related to plant life management and plant life extension. Section 4 discusses the aging evaluation for extension of qualified life of repairable sytems through stochastic point process (non-homogeneous Poisson model, NHPP). Section 5 presents a case study of service water pumps of a pressurized nuclear power plant. Section 6 presents conclusions and recommendations. References can be found in Section 8. It is noteworthy that NRC periodically updates paragraphs and appendices of 10 CFR Part 50 and associated Regulatory Guides. The references to these documents are related with the dates of last update described in the address http://www.nrc.gov/reading-rm/doc-collections/cfr/part050/.

2. Basic concepts

Operational experience at nuclear power plants has shown that two types of time-dependent changes occur in systems, structures and components/equipments (SSC): physical aging, or aging (that, henceforth, will be used in this text), which results in degradation (gradual deterioration) in physical characteristics; and the obsolescence that is the condition that occurs in structures and components that cease to be useful, despite being in perfect condition, owing to the emergence of other technologically more advanced ones.

Then, aging means the ongoing process by which physical characteristics of a system, structure and components or equipment (equipments and components will be used interchangeably in this text) change with time or usage. This process can proceed through a single aging mechanism or as the contribution of several mechanisms (cumulative effects).

Aging can lead to large scale degradation of physical barriers and redundant components, which can result in common-cause failures. These conditions can reduce the margins of safety equipment to values below the plant design basis or regulatory requirements, and cause damage to safety systems.

The knowledge of the aging and degradation processes and the developing of methods and guidelines for its management is important to the reliable and safe operation of nuclear power plants. The management of aging components can predict or detect the degradation of a component and take the appropriate corrective mitigation actions.

Evaluation of the cumulative effects of both physical aging and obsolescence on the safety of nuclear power plants is a continuous process and is assessed in a periodic safety review or an equivalent systematic safety reassessment program, as a renew of operating licenses, (IAEA, 2009).

The equipment qualification program of a nuclear power plant provides an effective aging management of plant components important to safety. The scope of the equipment qualification program should include equipment that performs safety functions and contributes to safety functions performance. Noteworthy is the demonstration that the equipment will continue performing its safety functions under harsh environmental conditions. These service conditions are those that exist after the postulated initiating event, and they are significantly different from normal operating conditions, which have their functionality demonstrated by performance during normal operation (pre-operational tests and periodic testing).

Qualified life (estabilished by equipment qualification program) is the period of time of normal operation for which an aging degradation would not prevent satisfactory performance of equipment if a postulated initianting event were to occur. The qualified condition of equipment established is expressed in terms of one or more measurable condition indicators for which it has been demonstrated that the equipment will meet its performance requirements, (IAEA, 2009).

Installed life is the period of time from the installation to the removal of equipment. A system, structure or equipment can have components whose qualified life can be lesser than the installed life. These components can be replaced (renewal) or undergo a repair program to maintain their qualification.

According to (IAEA, 2009), it is important to demonstrate that aging issues have been correctly taken into account for the whole planned plant lifetime, by ensuring that: (1) qualification tests take into account potential aging effects, in light of international knowledge and practice; (2) environmental conditions at the site are monitored to detect any changes from assumed values; (3) procedures for modifying qualified lifetimes are provided, especially in the case of changes from assumed values or of increasing failure frequency of some item of equipment; and (4) procedures for adapting aging tests and their duration of validity are provided.

The extension of qualified life should be approached through probabilistic methods. The evaluation based on deterministic methods defines the difference between the current state and condition of the item in the qualification phase, but does not define its probability of continuing performing its function adequately for a period longer than the one defined by its qualified life.

The combination of deterministic and probabilistic methods in the evaluation stage of aging effects is necessary because the extension of qualified life is an item of upgrading to a new life cycle since the system or equipment has operated for a number of years.

During operation, predictive, preventive and corrective maintenance work to maintain equipment in proper performance. Thus, maintenance works in both: functional equipment aspects, and how to check its availability through periodic tests laid down in the technical specifications and operating procedures.

In the operation of a nuclear power plant, the effectiveness of maintenance becomes an essential parameter in assessing system reliability (NRC, 2007a). It can be observed that the pursuit of maintenance effectiveness can not be done without: (1) adequate knowledge of aging mechanisms, (2) to assess the conditions for qualification of equipment, and (3) knowledge of equipment performance and trends in the records of operational experience.
A system may be defined as a collection of two or more components which must perform one or more functions. Nonrepairable systems are those which, when fail, must be replaced by new ones. Repairable systems are those which, after failure to perform at least one of their required functions, can be restored to perform all of their required functions by any method, other than the replacement of the entire system, and put into operation again under the same conditions before failure. By this definition, it is possible to repair without replacing any component (Asher &Feingold, 1984).
The failure process of a repairable system depends on the failure mechanism of the equipment and the maintenance policy applied. Hence, the choice of the statistical model to analyze the failure data of a repairable system should take into account the maintenance policy applied to the equipment, (Calabria et al., 2000).
The effects of aging increase the probability of equipment failures or even make it unavailable. These effects are traditionally calculated by using reliability models incorporating rates of time-dependent occurrence of failures. Thus, the behavior of components and systems in a plant are represented through changes in failure rates with time (Hassan et al., 1992).
The idea of modeling aging effects by simply considering that time-dependent failure rates can model this effect is not necessarily adequate because failure times may not be independent and identically distributed. This leads to the use of point processes and the concept of rate of occurrence of failures. Then, it is necessary to carefully evaluate the rate of occurrence of failures behavior detected in periodic testing and maintenance performance on the equipment and on the system. These actions can significantly impact aging effects studies(Saldanha et al., 2001; Saldanha & Frutuoso e Melo, 2009).

3. Regulatory aspects of plant life management and plant life extension

There are two regulation concepts for long term operation of nuclear plants. One is based on the periodic safety review, the other on the license renewal. Most European countries use the periodic safety review. The renewal of licenses is the practice adopted in the US. Both approaches have been adapted in some IAEA member states, including approaches that combine those two concepts. More detailed information about the practices and experiences of IAEA member states can be found in (IAEA, 2006, 2010).
Technical requirements and elements of the life management program remain the same for the two concepts (preventive maintenance and aging management programs, time limited aging analysis, equipment maintenance and qualification of response to obsolescence issues).
Considering regulatory aspects, it is important to control aging management practices of the operating organization and to verify the validity of forecasts for aging systems, structures and components important to safety. The environmental impact of long term operation, due to plant life management has to be assessed, although the study details depend on regulatory requirements (IAEA, 2006)

3.1 Periodic safety review

The periodic safety review is governed by the Safety Guide NS-G-2.10 (IAEA, 2003). It is a systematic review of a nuclear power plant safety analysis conducted at regular intervals

(usually 10 years) to deal with aging cumulative effects (both physical aging and obsolescence), modifications, operating experience, changes in international safety standards, to ensure a high level of safety through the operating lifetime of the plant. Thus, it has become an instrument for responding to requests by operating organizations allowed to continue operating plants beyond a deadline licensee or a specified period of safety evaluation. The process of periodic safety review is valid for nuclear power plants throughout its lifetime and guarantee that there remains a basis for valid licenses.
The regulatory system does not limit the number of periodic safety review cycles, even if the new cycle extends beyond the original design lifetime of the plant. The only condition is to demonstrate the plant safe operation for the next periodic safety review cycle, while maintaining safety margins. The periodic safety review is a tool that may be used by regulatory bodies to identify and solve safety issues (IAEA, 2006).
According to (IAEA, 2009), the review objective of aging management in a periodic safety review by operating organization is: (1) to determine whether aging in a nuclear power plant is being effectively managed so that required safety functions are maintained, and (2) whether an effective aging management program is in place for future plant operation (IAEA, 2003).
Therefore, the review of aging management within a periodic safety review therefore aims to establish whether: (1) for each SSC important to safety, all significant aging mechanisms have been identified; (2) there is a thorough understanding of relevant aging mechanisms and their effects; (3) the aging behaviour of SSCs over the period of operation is consistent with predictions; (4) there are adequate margins in respect to aging to ensure safe operation for at least the period until the next periodic safety review is due; (5) there is an effective aging management program (addressing operations, chemistry, maintenance, surveillance and inspection) in place for future plant operation, (IAEA, 2009).
Aging management for nuclear power plants is governed by safety guide NS-G-2.12 (IAEA, 2009). The objective of this safety guide is to provide recommendations for aging managing of SSC important to safety in nuclear plants. It focus on: (1) for the operating organization, providing technical support to establish, implement and make improvements of aging management program; and (2) regulatory organizations, providing technical support to prepare standards and regulatory guides to verify whether aging is being properly and effectively managed. It highlights the basic concepts of aging management and makes recommendations to make the proactive management of physical aging during the life of a nuclear power plant, presents a systematic approach to aging management of nuclear plants in operation, management of obsolescence, and aging management review in support of long-term operation.
IAEA has sponsored various programs and projects related to the aging of nuclear power plants, focusing on the management of aging, long-term operation reliability and economical aspects of life extension of licensed plants. The results of these programs are expressed in a series of technical documents related to methodologies for aging management, guide to operational data collection and maintenance for the management and evaluation mechanisms of aging (IAEA, 1992, 1992b, 1998, 1999, 2002, 2004, 2006, 2008). These documents can support and be references in the implementation of safety guide NS-G-2.12 (IAEA, 2009).

3.2 License renewal

The operating licenses of US plants have been issued with its lifetime of 40 years of operation. This limit was established by the 1954 Atomic Energy Act, which considered the perspectives of energy consumption.

The emphasis of this discussion focuses on the requirements of 10CFR54 (NRC, 2010a). The process of renewing the operating license is based on two principles: (1) the regulatory process, continued during the extended period of operation, is adequate to ensure that the licensing bases of all plants proceed at acceptable safety levels, with the possible exception of adverse effects of aging on certain systems, structures and components, and possibility of a few other issues related to safety during the extension period of operation; and (2) the licensing basis of each plant is maintained during the license renewal.
10CFR54 (NRC, 2010a) requires operators to identify all systems, structures and components: (1) that are related to safety; (2) whose failure may affect functions related to safety; and (3) that are able to demonstrate compliance with NRC fire protection requirements, 10CFR50.48 (NRC, 2007c), environmental qualification 10CFR50.49 (NRC, 2007d), heat shock 10CFR50.61 (NRC, 2007e), reactor transient shutdown reactor provided without scram 10CFR50. 62 (NRC, 2007f), and total loss of electrical power into alternating current 10CFR50.63 (NRC, 2007g).
Requirements should be included in the license renewal application in conformance to regulatory guide RG 1.188 (NRC, 2005), which will be assessed according to the standardized review plan, NUREG-1800 (NRC, 2010b). Structures and passive components are included in this item, they perform their functions without moving parts or without changes in the configuration, as follows: reactor, steam generator, piping, supports, seismic structures.
Active equipments are considered adequately monitored by the existing regulatory process where aging effects that may occur are more easily detectable, and therefore, more easily corrected, either by the testing program, or by the maintenance program, or even through the indicators highlighted in this performance (NRC, 2007b; NEI, 1996).
The important activity is the evaluation of the Analysis of Time-Limited Aging (TLAA), which is a set of analyses and calculations that involve systems, architectures and components that fall within the scope of the rule. TLAA should consider aging effects with approaches based on the original 40 years period, and: check the limits of renewal period, review or recalculate these limits by determining whether the sentence is appropriate, and demonstrate that aging effects are covered by the calculations.
License renewals are based on the determination that each plant will continue to maintain an adequate level of safety and, over plant life, this level should be increased by maintaining the licensing basis, with appropriate adjustments to consider new information from operational experience. The licensee may submit the license renewal to NRC 20 years before the end of the term of the current operating license.

4. Aging evaluation for extension of qualified life

The study of time-dependent failure rate models conducts to (Ascher & Feingold, 1969, 1978, 1984; Ascher,1992; Ascher & Hansen, 1998). They discuss and establish the terminology and notation used for performing the statistical analysis of repairable systems. They emphasize the distinction between non-repairable and repairable systems related to times to failure. They also criticize the reliability literature which considers repair as a renewal of the system to its original condition (the so-called 'as good as new' concept). They consider that this assumption is unrealistic for probabilistic modeling and leads to major distortions in the statistical analysis.
The non-homogeneous Poisson process (NHPP) and the renewal process (RP), generally homogeneous Poisson process (HPP), are commonly used models for repairable systems.

NHPP assumes that the unit is exactly at the same condition immediately after repair as it was immediately before failure ("as bad as old" concept), assuming that the repair time is negligible, while for the renewal process the repaired unit is always brought to a like new condition (as good as new concept). These two models and their implications are discussed in detail in (Ascher & Feingold, 1984).
NHPP application to the evaluation of the extension of qualified life can be useful due to the possibility, by controlling and monitoring the NHPP parameters, to identify whether the failure process is homogeneous in time. Then, by applying adequate corrective maintenance actions to mitigate aging effects, the repairable system can return to or be maintained under a homogeneous process.
Non-homogeneous Poisson point processes have wide application, like the analysis of failure occurrence trends, (Asher & Feingold, 1984), optimal replacement problems, (Bagai & Jain, 1994), warrant data analysis, (Majeske, 2003) and accident sequence precursor analysis, (Modarres et al., 1996). NHPP applications have been extensively made to the reliability growth model, (Crow, 1974). This model has been applied to the defense and aerospace industries in the US. Krivtsov (2007a) considers some practical extensions in the reliability application of this model.
It is possible to postulate a variety of point process models for the analysis of repairable systems. If one focus attention on the non-homogeneous Poisson process (NHPP), then it will be clearly seen that it is able to model time-dependent rates of failure occurrence. This model is conceptually simple and the relevant statistical methodology (maximum likelihood estimation and linear regression modeling) is well developed and easy to apply.
The importance of NHPP resides in the fact that it does not require the conditions of stationary increments. Thus, there is the possibility that events may be more likely to occur during specific time intervals. The NHPP has memory. Then, it is an adequate tool to analyze events where there may be, for example, aging.
By observing the successive failures of a repairable system, it is generally advisable to consider first the NHPP and, if there are no trends in the occurrence of failures, test the homogeneous model. The adoption a priori of a homogeneous model (iid) can lead to an inadequate reliability assessment.
The NHPP model has been tested and validated in several publications. The basis of this text, including the case study, is discussed in (Saldanha et al., 2001).
In 2007, the Reliability Enginering & System Safety journal issued a special number concerning stochastic processes (Krivtsov, 2007b). It contains papers on applications of NHPP to repairable systems (Asher, 2007; Krivtsov, 2007a; Finkelstein, 2007). In a workshop in the 17th International Conference on Nuclear Engineering, ICONE17, in 2009 in Brussels, the model called Crow-AMSAA (Pandey & Jyrkana, 2009) that uses the power law model for assessing the reliability of repairable systems was presented and discussed.

4.1 Basis for defining the model to extend qualified life of equipment

Safety analysis is the study, examination and description of a nuclear power installation behavior through its life, under normal and transient conditions, and also postulated events in order to determine safety margins provided in normal operation and transient regimes, and the adequacy of items to prevent consequences of accidents that might occur. It is essential for the safety assessment in the licensing process. The safety of the plant must be continuously monitored during operation and have constant review to maintain the level of safety. The safety analysis can be performed in two different ways that are complementary: the deterministic and probabilistic approaches.

In the deterministic safety analysis, the plant behavior, after an initiating event or malfunction, is studied with model calculations that describe the physical processes in reactor systems. The objective of this type of analysis is to verify whether allowed values of key variables of the plant are exceeded.
Probabilistic safety analysis (PSA) focuses on identifying the sequence of events that can lead to reactor meltdown, and on reliability studies of safety systems. The objective of this type of analysis is to indicate potential weaknesses in the design of systems and provide a basis for improving safety. It assumes that component failure rates are constant. When aging is explicitly considered, then changes in component failure rates as a function of aging should be considered.
In the beginning of the installation life, the repairable system has a constant rate of occurrence of failures and it is governed by a homogeneous process. The study of aging impact becomes more coherent by observing its effects. It is more appropriate to study the behavior of equipments and systems failures under the action of time, defining its probability density function and probability distribution function, and comparing with the probability density function and probability distribution function when they were not under the action of time.
Differences between probability distribution functions will be related to the increases in the probability of failure due to aging and can be used to aging control and management to reduce the impact of aging over system failures. They will be incorporated into studies of qualified life extension, and their effects on core damage frequency will be assessed.
The following is a sequence of actions that can be performed to define the model : (1) identify the nature of the system, whether repairable or non-repairable; (2) identify the stochastic point process associated to system failures: (2a) for the repairable system, the impact of aging through occurrences of failures can be assessed and it defines the model of failure intensity (ROCOF), these failures can be critical or degradation; and (2b) for the non-repairable systems, one can evaluate the impact of aging through replacements or renewals carried out.
The equipment qualification requirements are defined to repairable systems for a satisfactory performance in their qualified life, according to the requirements of the licensing basis established in the safety analysis report.
In the reliability study of repairable systems, each time between failures and the time of the last failure shall be considered. It is possible to obtain the expected number of failures for the next cycle of operation.
By monitoring and controlling the parameters of the probability model, during equipment operation, it is possible to check whether the equipment is leaving the basis of its qualification process, and so to check how the effects of time, degradation and operation modes can influence the equipment performance. The evaluation of the probability model parameters can be verified through the improvement (or degradation) of equipment. The knowledge of aging mechanisms allows reducing the values of these parameters and cast the model into a homogeneous process.
Here, the extension of qualified life must be defined by considering the operational experience of the equipment and then to perform equipment maintenance in compliance with the qualification basis. In this case, it means the return of the equipment to the "as good as new" condition. For this purpose, one should use as operation criteria goals values

the parameters of the probabilistic model. If the equipment operates within these limits, the process of occurrence of failures will be very close to some of the homogeneous processes and the influence of aging effects will be reduced.

4.2 Stochastic point process and repairable systems concepts

A stochastic point process *{N(t)}* is a collection of usually interrelated random variables, each labeled by a point *t* on the time axis and such that $N(t_2) - N(t_1) = N(t_1,t_2]$ expresses a finite nonnegative integer for all $t_2 > t_1 \geq 0$.

A stochastic point process is a mathematical model for physical phenomena characterized by highly localized events randomly distributed in a continuum. Thus, these processes can be designed to model a probabilistic experiment that arises in points, which can be named as "arrivals" on the time axis. Accordingly, failures of a repairable system can be represented as "arrivals" of a stochastic point process (Cox & Lewis, 1996).

The distinction between repairable and non-repairable systems is crucial and it should be emphasized if terminology is not adequately explained (Asher &Feingold, 1984; Ascher, 2007).

The concept of failure and how it must be taken into account quantitatively is not the same as the one for nonrepairable systems analysis. The classical idea of failure rate, traditionally employed is not adequate, as long as only the times to first component failures are considered. In this sense, after fixing a component which fails and is repaired, its particular history is investigated in terms of operational records.

The reliability figure of interest of nonrepairable systems is the survival probability. The times between failures of a nonrepairable system are independent and identically distributed (Asher & Feingold, 1984).

In the case of repairable systems, the reliability is interpreted as the probability of not failing for a given period of time. The analysis must be performed without assuring that the times between failures are independent and identically distributed. It is important to emphasize that there may be a system with repairable components which is different from a repairable system in the sense that the first may have nonrepairable components (Asher & Feingold, 1984).

The first concept to discuss is the rate of occurrence of failures. For a repairable system, let *N(t)* be the number of failures in the interval *(0, T]* and t_1 , t_2 , ..., be the system failure times and T_i *(i= 1, 2, 3....)* the elapsed time between the *(i-1)th* and the *ith* failure.

The behaviour of T_i is of great importance in reliability analyses, for it allows the determination of trends in the times between failures, increasing (sad system), decreasing (happy system) or constant (Asher &Feingold, 1984).

The rate of occurrence of failures ν(t) is defined as (Asher &Feingold, 1984; Crowder et al., 1991)

$$\nu(t) = \frac{d}{dt} E\{N(0,t]\} \tag{1}$$

It is important to distinguish the concept of rate of occurrence of failures from that of failure rate, a concept traditionally employed in reliability engineering (Asher &Feingold, 1984).

A natural estimator of ν(t) is given by (Crowder et al., 1991) and discussed by (Lai & Xie, 2006):

$$\hat{v}(t) = \frac{\{N(t, t+\Delta t]\}}{\Delta t} \tag{2}$$

for an adequate Δt. The choice of this interval is arbitrary but like the choice of interval widths for histograms, the idea is to highlight the main features of the given data.
The expected number of failures in the interval $(t_2 - t_1)$ is given by

$$E[N(t_2) - N(t_1)] = \int_{t_1}^{t_2} v(t)\, dt, \tag{3}$$

and the reliability function is

$$R(t_2 - t_1) = \exp\{-\int_{t_1}^{t_2} v(t)\, dt\}. \tag{4}$$

4.3 Model for rate of occurrence of failures (ROCOF) by Poisson processes

A stochastic point process *{N(t), t≥0}* is said to be a homogeneous Poisson process (HPP) if it satisfies the following conditions (Asher & Feingold, 1984): (*1*) *N(0)=0*; (*2*) it has independent and stationary increments; and (*3*) the number of events in each time interval follows a Poisson distribution with mean $m(t)= \mu t$, $0< \mu < \infty$, so that:

$$P\{N(t, t+s] = k\} = (\mu t)^k (k!)^{-1} \exp(-\mu t) \tag{5}$$

for all t, t≥0 and k= 0,1,.....
One possible generalization of a HPP is a nonhomogeneous Poisson process (NHPP), which has a time-dependent rate of occurrence of failures, $M'(t)= dE[N(t)]/dt = v(t)$, *t≥0,* and the events, that are not independent and identically distributed, follow a Poisson distribution with mean $m(t)= \int_0^t v(s)ds$, so that:

$$P\{N(t, t+s] = k\} = (\int_t^{t+s} v(x)dx) \quad (k!)^{-1} .exp[-\int_t^{t+s} v(x)dx\,], \tag{6}$$

By choosing a suitable parametric form for $v(t)$, a flexible model for failures of repairable systems can be obtained. In the literature, two NHPP failure models are widely used: the log-linear and the power law (Crowder et al., 1991). It is necessary to decide which of these models is preferable in each case.
The NHPP with a log-linear rate of occurrence of failures is discussed by (Cox & Lewis, 1966) and is given by:

$$v_1(t) = e^{\beta_0 + \beta_1 t}. \tag{7}$$

The ROCOF decreases (happy system) if $\beta_1 < 0$. If $\beta_1 > 0$, the ROCOF increases (sad system). If $\beta_1 \cong 0$, the ROCOF has a linear trend over short periods of time.

The second model is based on the Weibull distribution and is referred to as the power law, (Crow, 1974; Ascher and Feingold, 1984). It is given by

$$v_2(t) = \gamma.\delta.t^{\delta-1}, \ \gamma > 0, \ \delta > 0, \text{ and } t \geq 0. \tag{8}$$

If $\delta > 1$, the ROCOF increases. The ROCOF decreases when $0 < \delta < 1$. For $\delta = 2$, one has a linearly increasing ROCOF.

If failures of repairable systems were observed for the time period $(0, t_0]$ and the logged times of occurrence were $t_1, t_2, ...t_n$., then the likelihood function could be obtained by analysing the probability of observing no failures in $(0, t_1)$, one failure in $(t_1, t_1+\Delta t_1)$, no failures in $(t_1+\Delta t_1, t_2)$ and so on, until the last interval of no failures in $(t_n+\Delta t_n, t_0)$, for small $\Delta t_1.....\Delta t_n$ and taking the limit as the Δt_n go to zero. The likelihood function obtained from Eq. (3) is given by

$$L = \prod_{i=1}^{n} v_1(t) \exp[N(0, t_0)]. \tag{9}$$

Another possibility to observe failures of repairable systems is to record them until the occurrence of the *nth* failure. In this case, the likelihood function of Eq (9) or the loglikelihood is still valid but t_0 must be replaced by t_n, the time of occurrence of the *nth* failure.

To fit a NHPP with a rate of occurrence of failures given by Eq (7) to a set of failure data of a repairable system, using statistical likelihood-based methods it is necessary to obtain the loglikelihood function of Eq (9). It is given by

$$l_1 = n\beta_0 + \beta_1 \sum_{i=1}^{n} t_i - [\frac{\exp(\beta_0)\exp\{(\beta_1 t_0) - 1\}}{\beta_1} \tag{10}$$

From maximum likelihood equations $\partial l_1 / \partial \beta_1 = 0$, the maximum likelihood estimator of β_1 can be obtained by solving the transcendental equation:

$$l_1 = \sum_{i=1}^{n} t_i + \frac{n}{\beta_1} + \frac{nt_0}{1 - \exp\{-\beta_1 t\}} = 0 \tag{11}$$

After obtaining $\hat{\beta}_1$, one has

$$\hat{\beta}_0 = \ln\left\{\frac{n\,\hat{\beta}_1}{\exp\left\{\hat{\beta}_1 t_0\right\} - 1}\right\} \tag{12}$$

A natural test of hypothesis, when considering the reliability of a repairable system, is to check whether the rate of occurrence of failures is constant. For the log-linear model, it is $\beta_1 = 0$. A commonly employed hypothesis test is the Laplace test, (Crowder et al., 1991), which is based on the statistic

$$U = \left[\frac{\sum_{i=1}^{n} t_i - \frac{1}{2} n\, t_0}{t_0 \sqrt{n/12}} \right] \tag{13}$$

which under the null hypothesis approaches a standard normal distribution. If H_1: $\beta_1 \neq 0$, one rejects H_0 if $|U|$ is large. On the other hand, if H_1: $\beta_1 > 0$, one rejects H_0 if U is large, and if H_1: $\beta_1 < 0$, one rejects H_0 if $-U$ is large.
For a NHPP with rate of occurrence of failures given by Eq (8), the loglikelihood function of Eq (3) is given by

$$l_2 = \sum_{i-1}^{n} [\ln \gamma + \ln \delta + (\delta - 1) \ln t_i] - \gamma\, t_0^{\delta} \tag{14}$$

From maximum likelihood equations $\partial l_2 / \partial \gamma = 0$ and $\partial l_2 / \partial \delta = 0$, one obtains:

$$\hat{\gamma} = \frac{n}{t_0^{\hat{\delta}}} \tag{15}$$

and

$$\hat{\delta} = \frac{n}{n \ln t_0 - \sum_{i=1}^{n} \ln t_i} \tag{16}$$

In order to test whether the rate of occurrence of failures is constant, that is $\delta=1$, the following statistic is employed (Crowder et al., 1991)

$$V = 2 \sum_{i=1}^{n} \ln \left(\frac{t_0}{t_i} \right) \tag{17}$$

which under the null hypothesis follows a $\chi^2(2n)$ distribution. Large values of V indicate reliability growth $(0 < \delta < 1)$, whereas small ones indicate deterioration $(\delta > 1)$.
A question that naturally arises in this context is which model to choose for $v(t)$. It was suggested (Crowder et al., 1991) to make this choice using the models based on loglikelihood methods and after comparing with the models obtained by linear regression (graphical methods): (1) a first step is to plot the failure number i against the t_i. The lack of linearity is an indication that the rate of occurrence of failures is not constant; (2) the second step is to obtain the expression for $v_1(t)$ and $v_2(t)$ by loglikelihood methods; and (3) the last step is to choose $v(t)$ using linear regression (graphical method).
The graphical method is based on the expected number of failures until time t, $E[N(t)]$. Using Eqs. (7) and (8) in Eq (3), one has:

$$- \; v_1(t), \qquad E[N(t)] = \frac{e^{\beta_0}}{\beta_1}\{e^{(\beta_1 t)} - 1\}; \text{ and} \tag{18}$$

$$- \; v_2(t), \qquad E[N(t)] = \gamma \; t^{\delta}. \tag{19}$$

Considering that the observation period $(0, t_0]$ is divided into k arbitrary intervals of the form $(0, a_1]$, $(a_1, a_2]$, ... $(a_{k-1}, t_0]$, an estimate of $v\left[\frac{1}{2}\left(a_{j-1} + a_j\right)\right]$, is given by (Crowder et al., 1991):

$$v\left[\frac{1}{2}\left(a_{j-1} + a_j\right)\right] = \frac{N(a_j) - N(a_{j-1})}{a_j - a_{j-1}} \tag{20}$$

for $j = 1, 2, ...k$, where $a_0 = 0$ *and* $a_k = t_0$.

Making $b_j = \frac{1}{2}\left(a_{j-1} + a_j\right)$, a plot of $v(b_j)$ x b_j furnishes an indication of the shape of the rate of occurrence of failures, $v(t)$. The choice of k and a_j is left to the user. However, it is advisable to test different subdivisions of the observation interval in order to verify that the shape of the plot does not depend on the chosen subdivision.

If $v_1(t)$ is appropriate for $v(t)$, then the plot of $\ln v(b_j)$ x b_j will show a straight line with slope β_1 and intercept β_0. On the other hand, if $v_2(t)$ is appropriate for $v(t)$, the plot of $\ln v(b_j)$ x $\ln b_j$ will also show a straight line, but with slope $(\delta - 1)$ and intercept $(\ln \gamma + \ln \delta)$. If there is a competition between the models evaluated, the choice will be the model that has the highest value for the maximum likelihood function.

For NHPP models, the extension of qualified life must be defined by the return of the equipment to the "as good as new" condition using the operation criteria goals presented in Table 1. As the equipment operates within these limits, the process will be very close to the homogeneous processes.

ROCOF model	homogeneity	linear trend
$v_1(t) = e^{\beta_0 + \beta_1 t}$ (log-linear)	$\beta_1 < 0$	$\beta_1 \cong 0$
$v_2(t) = \gamma.\delta.t^{\delta-1}$ (power law) $\gamma > 0, \; \delta > 0, \text{ and } t \geq 0$	$0 < \delta < 1$	$\delta = 2$

Table 1. Operating criteria for NHPP parameters for reducing aging effects

5. An application of the NHPP model

The following analysis is related to the service water pumps (SWP) of a typical PWR nuclear power plant. Considered was the degradation or critical failures for which corrective actions in parts, subsystems, and systems inside limits of SWP, for which maintenance corrective actions were necessary (IAEA, 1988).

Failures of SWPs were revealed in operation by daily inspection (lub-oil level of pump systems components pump, discharge pressure, leaks, vibrations, etc) or malfunction indications by sensors or alarms.
The failure times employed in the analysis were generated as follows. A time period of 1725 calendar days has been considered, which is equivalent to approximately three burnup fuel cycles and reload periods. The corresponding operational time for the SWPs was 20,300 hours, (NRC, 1988).
The first step is to check whether the rate of occurrence of failures is constant. A plot of the accumulated number of failures versus the accumulated failures times (operation times) is displayed in Figure 1, which is based in data from the first and second columns of Table 2. No linearity is seen for the plotted points so that the rate of occurrence of failures is clearly time dependent.

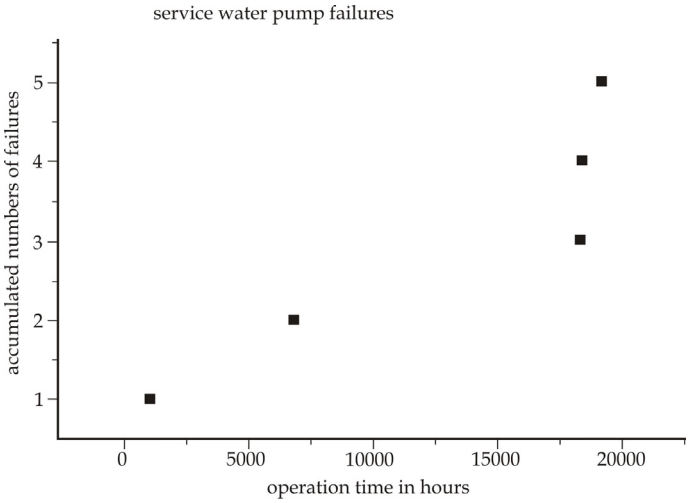


Fig. 1. Accumulated number of failures x accumulated operation time of SWP.

i	t_i	TEF	$\ln t_n/t_i$
1	1080		2,875
2	6840	5760	1,029
3	18300	11460	0,045
4	18360	60	0,042
5	19140	780	0
	Σ= 63720		Σ= 3,991

i = failure number: t_i= time to failure, in hours; and TEF= time between failures, in hours.

Table 2. Data for loglikelihood-based methods for the rate of occurrence of failure models

The second step in the analysis, it to check which model is applicable for $\nu(t)$, Eq (7) or Eq (8). The data is also presented in Table 2.
Solving Eq (10) for $v_1(t)$, one obtains $\hat{\beta}_1 = 0.000112$ and from this result, $\hat{\beta}_0 = -9.51$ from Eq (12). These values are tested through Eq (13), U= *1.98*. This value is considered large by

(Crowder et al., 1991) and the estimated parameters are considered adequate for the model (rejection of the $\beta_1=0$ hypothesis). Then, for this case, Eq (7) is given by:

$$v_1(t) = \exp\{-9.51 + 0.000112\ t\}\ \left(h^{-1}\right) \tag{21}$$

Solving Eqs. (15) and (16) for $v_2(t)$, one obtains $\hat{\delta} = 1.253$ and $\hat{\gamma} = 0.00000216$. These values are tested by Eq (17), $V = 7.982$. This value is considered small (Crowder et al., 1991) and the estimated parameters are considered adequate for the model (rejection of the $\delta = 1$ hypothesis). Then, for this case, Eq. (8) is given by:

$$v_2(t) = 2.71 \times 10^{-5}\, t^{0.253}\ (h^{-1}). \tag{22}$$

The third step in the analysis is the application of linear regression to the failure data in order to choose which model is applicable for $v(t)$, Eq (7) or Eq (8). The observation period *(0.19140]* has been divided into three distinct intervals, and the shape of $v(t)$ has been checked through the plot $v(b_j)\ x\ b_j$. Linear regression has been performed for each interval considering the plots of $\ln(v(b_j))\ x\ b_j$, for $v_1(t)$ and of $\ln(v(b_j))\ x\ \ln(b_j)$, for $v_2(t)$.

Table 3 presents the results obtained by linear regression methods for each interval and by the loglikelihood method. It may be inferred that the $v_1(t)$ model adequately fits the rate of occurrence of failures $v(t)$, considering the results of interval splitting #2 and the loglikelihood method.

Table 4 presents the performed interval splitting #2 for the data where n_f represents the number of failures in the interval. Figures 2 and 3 show the shape of $v(t)$ through the plot $v(b_j)\ x\ b_j$ and the scattering diagram for the data in Table 4, respectively.

Parameter	β_1	β_0	δ	γ
Interval splitting # 1	0.00015	-9.82	2.111	0.0005E-5
Interval splitting # 2	0.00012	-9.33	1.862	0.0007E-5
Interval splitting # 3	0.00008	-9.18	1.584	0.0760E-5
Loglikelihood method	0.000112	-9.51	1.253	2.16E-5

Table 3. Estimated parameters for choosing $v(t)$

interval (hr)	n_f	b_j	$\ln b_j$	$v(b_j)$	$\ln v(b_j)$
0-6000	1	3000	8.006	0.000017	- 8.699
6000-11000	1	8500	9.048	0.000020	- 8.517
11000-18400	2	14700	9.596	0.000270	- 8.216
18400-19140	1	18770	9.840	0.001350	- 6.607

Table 4. Division 2 for the model of Eq (7)

One can see that $v_1(t)$ is appropriate to model $v(t)$, because the fit of the line $\ln v(b_j)\ x\ \ln b_j$ values in Table 3, obtained a slope $\hat{\beta}_1 = 0.00012$, and an intercept

$\hat{\beta}_0 = -9.33$, values that are close to those obtained through the maximum likelihood method.

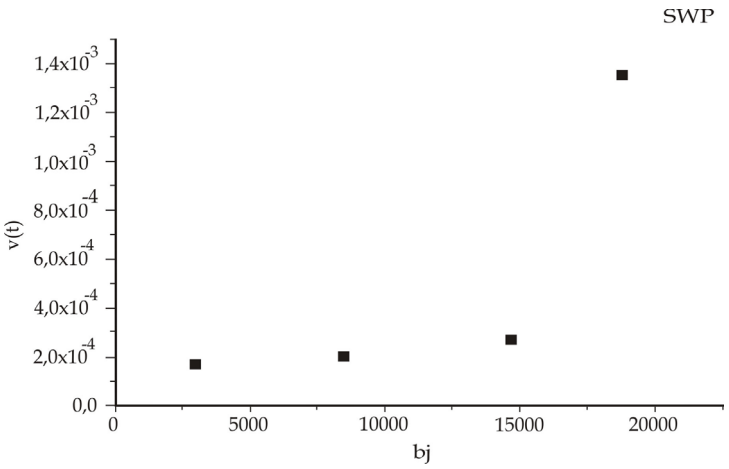


Fig. 2. Analysis of the model of Eq. (7)

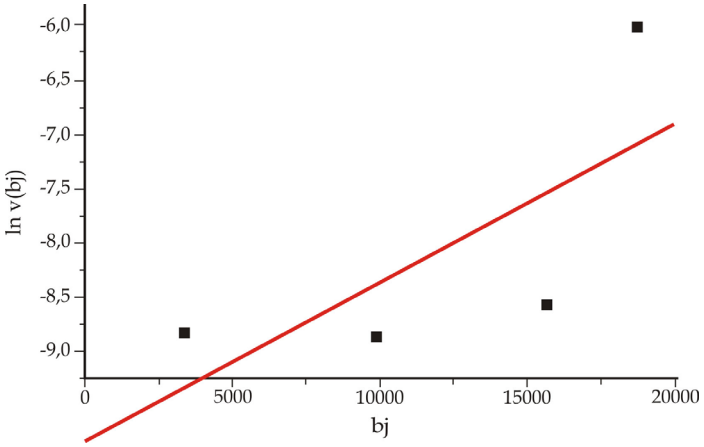


Fig. 3. Scatter diagram (data in Table 3), $\hat{\beta}_1 = 0.00012$ (slope) and $\hat{\beta}_0 = -9.329$ (intercept).

Figure 4 exhibits the shape of $v\ (t)$ for the observed period *(0, 19140]*, using Eq. (21). An increasing trend in the failure occurrence and actions on degradation factors in pump performance can be observed. Thus, it is possible to obtain the expected numbers of failures in other burnup fuel cycles from Eq. (22).

Considering that the SWPs will operate under the same model than that for the past three cycles, the evaluated time period will be *(19140, 25500]*. Using the parameters β_1 and β_2 in Eq (18), 6 failures have been obtained. Table 5 present these values. Figure 5 presents the expected cumulative number of failures in hours of operation until the C4 cycle.

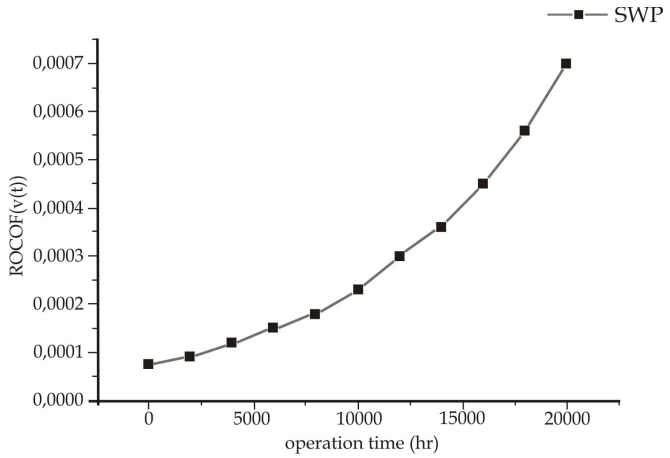


Fig. 4. Rate of occurrence of failures $v\ (t)$ in the operation period (0, 19140], in hours

interval	E[N]
19140 - 21140	1
21140 - 2314	2
23140 - 25500	3

Table 5. Expected numbers (E[N]) of failure for period (19140, 25500] until the C4 cycle

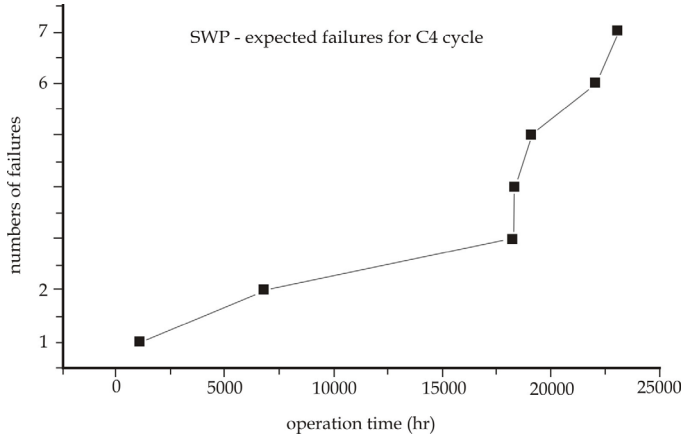


Fig. 5. Expected cumulative number of failures in hours of operation until the C4 cycle

Supposing that in order to decrease the expected number of failures for the period *(19140, 25500]* adequate aging management actions have been taken, like periodic testing and maintenance, then it would be possible to reduce the number of failures to 2 in the period, occurring, in time points, 22104 h and 23112 h, respectively.

The impact of failure reduction may be quantified by the application of the NHPP model considering the period (0, 25500]. Using the same procedure discussed before, one obtains the following expression for $v\ (t)$:

$$v_{c\,4}(t) = \exp\{-9.4 + (9.9E-5)t\}\ h^{-1} \tag{23}$$

Figure 6 shows the comparison between the rates of occurrence of failures considering the expected trend by Eq (21), Table 5, and considering management of aging and maintenance actions in the period (19140, 25500]. It can be observed that the corrective actions reduce the increasing trends in failures, ($\hat{\beta}_1 = 0.00012 > \hat{\beta}_{1C4} = 0.00009$).

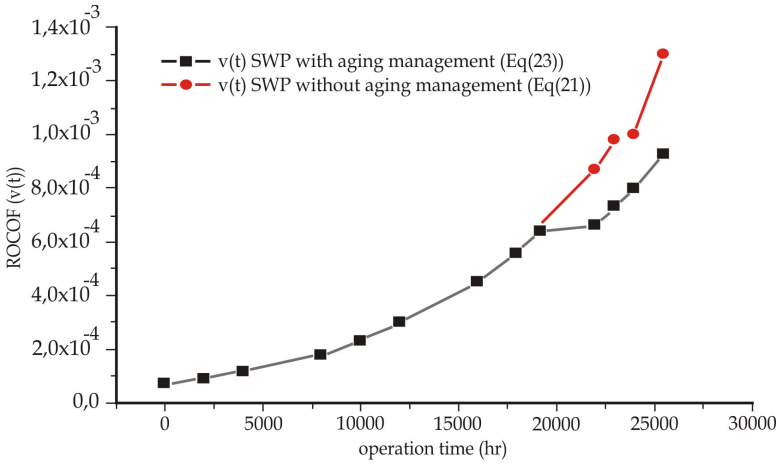


Fig. 6. Comparison of $v(t)$ by Eq (21) (trends) and Eq. (23) (released) in period *(19140, 25500]*, in operation hours

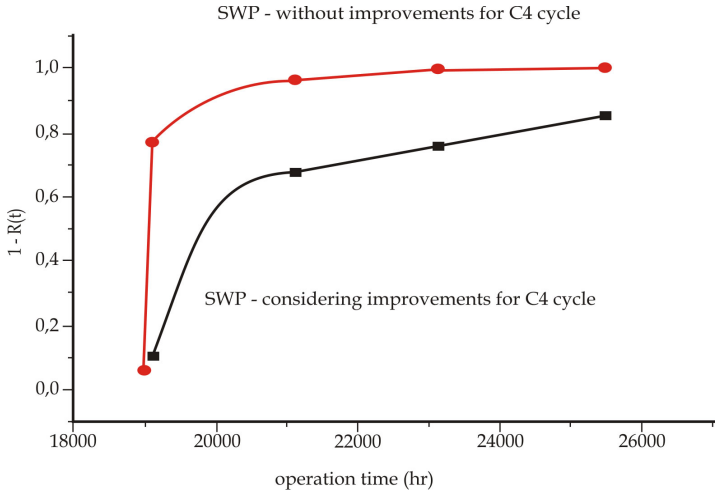


Fig. 7. Comparison of the curves of SWP failure probability with improvements and without improvements to the C4 cycle

Figure 7 shows the comparison of the SWP curves failure probability with improvements and without improvements to the C4 cycle. It can be verified that a repairable system with a

time-dependent rate of occurrence of failures can be influenced by management of aging and maintenance policies.

6. Conclusion

The equipment qualification program of a nuclear power plant provides an effective aging management of plant components important to safety. The knowledge of the aging and degradation processes and the development of methods and guidelines for its management is important to the reliable and safe operation of nuclear power plants. The management of aging components can predict or detect the degradation of a component and take the appropriate corrective mitigation actions.

The extension of qualified life is a reality. It must bear combination of deterministic and probabilistic methods in the evaluation stage of aging effects. While the evaluation based on deterministic methods defines the difference between the current state and condition of the item in the qualification phase, probability methods define whether it can continue performing its function adequately for a longer period than is defined by its qualified life.

It is possible to postulate a variety of point process models for the analysis of repairable systems for the extension of qualified life. If one focuses attention on the non-homogeneous Poisson process (NHPP), then it will be clearly seen that it is able to model time-dependent rates of occurrence of failures.

This model is conceptually simple and the relevant statistical methodology (maximum likelihood estimation and linear regression modeling) is well developed and easy to apply. By considering it, one eliminates the unrealistic assumption for the probabilistic modeling of the ROCOF that considers a renewal as a repair ('as god as new' versus 'as bad as old' hypothesis). Different models for the rate of occurrence of failures are possible and we discussed the two most common, namely, the log-linear and the power law.

The application discussed is quite simple but this is not a methodology shortcoming. Our concern was to show how to perform the analysis by taking into account a general procedure for hypothesis testing.

At first, one of those models should be adequate for the analyzed data but the hypothesis testing could reveal that none is appropriate. Sound statistical inference should always be employed for the obtained results in terms of the expected number of failures because the reliability for a given time period could be misevaluated.

The pattern of successive failures of repairable systems will define the model to be used. If failures exhibit increasing trends, one applies the NHPP (time-dependent ROCOF); otherwise, one should apply the HPP (constant ROCOF). The NHPP model adequately includes the variations in the rate of occurrence of failures behavior due to periodic testing and maintenance activities performed in repairable systems. Then, it may be used to survey aging mechanisms during the operational life of repairable systems and to the assessment of maintenance effectiveness.

The case study demonstrates that it is not difficult to perform failure data analysis, to evaluate trends and verify the impact of these trends in the reliability analysis of repairable systems. Considering the requirements of the Maintenance Rule (NRC, 2007a), in the aging management for active equipments, the study of the SWPs shows that the MPLP complies with the licensing basis.

In what concerns applications of other point processes models, (Saldanha & Frutuoso e Melo, 2009) discuss the application of the Modulated Power Law Process (MPLP) to

represent the ROCOF of repairable systems for evaluating the extention of qualified life as a generalization of the NHPP model that turns to be more realistic to represent the impact of maintenance actions.
In the area of maintenance policies, applications of Generalized Renewal Process (GRP) are very recent in the field of reliability engineering. (Kahle, 2007) discusses the optimal maintenance policies for the models of standard GRP repair. (Garcia et al, 2008) address the scheduling of preventive maintenance using genetic algorithms for systems modeled by this class of point processes. (Damaso et al, 2009) consider again the GRP for the modeling of multi-objective optimization of testing policies of aging systems.

7. Acknowledgment

One of the authors (PLCS) would like to thank the Associação Brasileira de Ensino Universitário (UNIABEU), Nova Iguaçu, Rio de Janeiro, for its financial supporting program (PROAPE).

8. References

Ascher, H. and Feingold, H., (1969), Bad-as-Old" Analysis of System Failure Data, *Annals of Assurance Sciences*, Gordon and Beach, New York, pp 49-62

Ascher, H. and Feingold, H. (1978), Is There Repair after Failure?, *Annual Reliability and Maintainability Symposium,* IEEE-77CH1429-OR, pp 153-159

Ascher, H. and Feingold, H. (1984), *Repairable System Reliability, modeling, inference, misconception and their causes*, Marcel Decker Inc, Washington

Ascher, H. (1992), Basic probabilistic and statistical concepts for maintenance of parts and systems, *IMA Journal of Analysis of Mathematics, Applied in Business and Industry 3*, 153-167

Ascher, H., Hansen, C.K., (1998), Spurious exponentiality observed when incorrectly fitting a distribution to nonstationary data, *IEEE Transactions on Reliability 47* (4),451–459

Ascher, H. (2007), Different insights for improving part and sytem reliability obtained from exactly same DEFOM failure numbers", *Reliability Engineering & System Safety*, 92, pp. 552-559

Bagai, I., Jain, K. (1994), Improvement, deterioration, and optimal replacement under age-replacement with minimal repair. *IEEE Transations in Reliability 43.* 156-162

Calalabria R., Guida M., & Pulcini G. (2000), Inference and test in modeling the failure/ repair process of repairable mechanical equipments, *Reliability Engineering & System Safety*, 67, pp. 41-53

Cox, D.R., and Lewis, P.A.W. (1966), *The Statistical Analysis of Series of Events*, Methuen, London

Crow, L.H. (1974), Reliability Analysis of Repairable Systems, *Reliability and Biometry, SIAM*, pp 379-410

Crowder, M. J., Kimber, A. C., Smith, R.L., and Sweeting, T. J. (1991), *Statistical Analysis of Reliability Data*, Chapman & Hall, London

Damaso, V. C., Garcia, P. A. A., Sant'Ana, M. C., and Frutuoso e Melo, P. F., (2009), Multiobjective opmitization of test policies for GRP-modeled aging systems. *Proccedings ESREL-09. European Safety and Reliability Association.* 2143-2149

Finkelstein, M. (2007), Shocks in homogeneous and heterogeneous populations, *Reliability Engineering & System Safety*, 92, pp. 569-574

Garcia, P. A. A., Sant'Ana, M. C., Damaso, V. C., and Frutuoso e Melo, P. F. (2008), Genetic algorithm optimization of preventive maintenance scheduling for repairable systems modeled by generalized renewal processes. *Proceedings ESREL-08. European Safety and Reliability Association*

Hassan, M., Samantha, P. & Vesely, W. (1992), Validation issues in aging risk evaluation, *NUREG/CP-0122*, Nuclear Regulatory Commission, Washington, D.C.

IAEA, (1988), *Component Reliability Data for Use in Probabilistic Safety Assessment- Technical Document 478*, International Atomic Energy, Vienna

IAEA, (1992), *Data Collection and Record Keeping for Management of Nuclear Power Plant Ageing- A Safety Practice- Safety Series 50-P-3*, International Atomic Energy Agency, Vienna

IAEA, (1992b), *Methodology for Management of Ageing of Nuclear Power Plant Components Important to Safety- Technical Report Series 338*, International Atomic Energy Agency, Vienna

IAEA (1994), *Periodic Safety Review of Operational Nuclear Power Plants, Safety Series No. 50-SG-O12*, International Atomic Energy Agency, Vienna

IAEA, (1998), *Equipment Qualification in Operational Nuclear Power Plants: upgrading, preserving and reviewing- Safety Report Series 3*, International Atomic Energy Agency, Vienna

IAEA, (1999), *AMAT Guidelines, Reference document for IAEA Aging Management Assessment Teams- Service Series nº4*, International Atomic Energy Agency, Vienna

IAEA (2002), Nuclear Power Plant Life Management, *Proceedings of International Symposium on Nuclear Power Plant Life Management, Budapest, Hungary, 4-8 november/2002*, International Atomic Energy Agency, Vienna

IAEA (2003), *Periodic Safety Review of Operational Nuclear Power Plants- Safety Guide No. NS-G-2.10* International Atomic Energy Agency, Vienna

IAEA (2004), Plant Life Management for Nuclear Power Plants, *Proceedings Regional workshop on Optimisation of Service Life of Nuclear Power Plants, Sao Paulo, Brazil, 6 to 9 July 200,*Vienna, International Atomic Energy Agency

IAEA (2006), *Plant Life Management for Long Term Operation of Light Water Reactors - Technical Report Series No 448*, International Atomic Energy Agency, Vienna

IAEA (2008), *Safe Long Term Operation of Nuclear Power Plants- Safety Report Series No 57*, International Atomic Energy Agency, Vienna

IAEA (2009), *Ageing Management for Nuclear Power Plants- Standards Series No NS-G-2.12*, International Atomic Energy Agency, Vienna

IAEA (2010), *Periodic Safety Review of Nuclear Power Plants: Experience of Member States- Technical Document, TECDOC No 1643*, International Atomic Energy Agency, Vienna

Krivtsov V. V. (2007a), Practical extensions to NHPP application in repairable system reliability analysis, *Reliability Engineering & System Safety*, 92, pp. 560-562

Krivtsov V. V. (2007b), Recent advances and applications of stochastic point process models in reliability engineering, *Reliability Engineering & System Safety*, 92, pp. 549-551

Modarres, M., Martz, H., Kaminskiy, M. (1996), The accident sequence precursor analysis. *Nuclear Science Engineering 123*, 238-258

Majeske, K. D. (2003), A non-homogeneous Poisson Process predictive model for automobile warranty claims. *Department of Operations and Management Science. University of Michigan.* Ann Arbor

NEI (1996), *Industry Guideline for Monitoring Effectiveness of Maintenance at Nuclear Power Plants- NUMARC-93-01*, Nuclear Energy Institute, Paris, USA

NRC (1988), Operating Experience Feedback Report- Service Water System Failures and Degradations, *NUREG-1275*, Vol 3, Washington, 1988

NRC (1997), *Monitoring the Effectiveness of Maintenance at Nuclear Power Plants- Regulatory Guide RG 1.160*, Nuclear Regulatory Commission, Washington, DC

NRC (2000), *Assessing and Managing Risk Before Maintenance Activities at Nuclear Power Plants- Regulatory Guide RG 1.182*, Nuclear Regulatory Commission, Washington, DC

NRC (2005), *Standard Format and Content for applications to renew Nuclear Power Plant Operating Licenses- Regulatory Guide RG 1.188*, Nuclear Regulatory Commission, Washington, DC

NRC (2007a), *Environmental Protection Regulations for Domestic Licensing and Related Regulatory Functions- 10CFR51*, Nuclear Regulatory Commission, Washington, DC

NRC (2007b), *Monitoring the Effectiveness of Maintenance at Nuclear Power Plants- 10CFR50.65*, Nuclear Regulatory Commission, Washington, DC

NRC (2007c), *Fire Protection-10CFR50.48*, Nuclear Regulatory Commission, Washington, DC

NRC (2007d), *Environmental qualification of electric equipment important to safety for Nuclear Power Plants- 10CFR50.49*, Nuclear Regulatory Commission, Washington, DC

NRC (2007e), *Fracture tougnhness requirements for protection agaist pressurized thermal shock events - 10CFR50.61*, Nuclear Regulatory Commission, Washington, DC

NRC (2007f), *Requirements for reduction on risk from anticipated transients without scram (ATWS) events for light-water-cooled Nuclear Power Plants- 10CFR50.62*, Nuclear Regulatory Commission, Washington, DC

NRC (2007g), *Loss of all alternating current power - 10CFR50.63*, Nuclear Regulatory Commission, Washington, DC

NRC (2010a), *Requirements for Renewal of Operating Licenses for Nuclear Power Plants- 10CFR54*, Nuclear Regulatory Commission, Washington, DC

NRC (2010b), *Standard Review Plan for License Renewal- NUREG-1800*, Nuclear Regulatory Commission, Washington, DC

NRC (2010c), *Generic Aging Lessons Learned*, NUREG-1801, Nuclear Regulatory Commission, Washington, DC

Pandey, M. & Jyrkama, M. (2009),Crow-AMSAA Model, *ICONE 17, Workshop on Engineering Reliability and Life Cicle Management*, July 12th, Brussels, Belgium

Saldanha, P. L. C., Simone, E. A. & Frutuoso e Melo, P. F. (2001), An application of non-homogeneous Poisson point processes to reliability analysis of service water pumps, *Nuclear Engineering and Design*, n° 210, pp. 125-133

Saldanha, P. L. C., & Frutuoso e Melo, P. F. (2009), Aging Evaluation for the extension of qualified life of nuclear Power plant equipment through Modulated Poisson Processes, *Proceedings of the 17th International Conference on Nuclear Engineering, ICONE17*, July 12-16, 2009, Brussels, Belgium

Young, G. (2009a), Status of License Renew in U. S., *Proceedings of Regional Latin America Workshop on Plant Life Management and Long Term Operation of Nuclear Power Plants, Buenos Aires, dec/09*, International Atomic Energy Agency, Vienna

Young, G. (2009b), License Renewal Rule and Aging Management of Passive SSCs, *Proceedings of Regional Latin America Workshop on Plant Life Management on Plant Life Management and Long Term Operation of Nuclear Power Plants, Buenos Aires, december/09*, International Atomic Energy Agency, Vienna

Young, G. (2009c), Maintenance Rule in the US Aging Management of Active Systems and Structures", *Proceedings of Regional Latin America Workshop Buenos Aires on Plant Life Management and Long Term Operation of Nuclear Power Plants, Buenos Aires, december/09*, International Atomic Energy Agency, Vienna

17

Non-Destructive Testing for Ageing Management of Nuclear Power Components

Gerd Dobmann
Fraunhofer-IZFP
Germany

1. Introduction

Worldwide a renaissance of the nuclear industry is obviously taken place and many countries favour nuclear power as one reliable opportunity to generate electrical energy at very low CO_2 generating rates in order to avoid the green house effect in the earth atmosphere. However, since 1986 when the Tschernobyl accident was happening, nearly nowhere new nuclear power plants were established. The People Republic of China, India and in the last decade Japan, Finland and France are the exception. In other words, existing supply chains of former manufacturers were mainly destroyed or have changed its technical application field. Furthermore, a lot of technical expertise was lost as younger generations were influenced politically to find its interest in other scientific areas other than in nuclear physics or nuclear engineering.
Even if we can observe today a change in mind in many countries concerning the acceptance of nuclear power the question seriously is to answer: Will we find enough well skilled technicians to reliably build all the planned nuclear power plants in the future?
Therefore, life extension of existing plants the more plays an important role. This is truer as we have learnt in the last decades how many potential we have for life time extension even if we take into account ageing phenomena concerning the materials as thermal ageing, fatigue and neutron embrittlement when we think at steel components in the primary circuit; as there are the reactor pressure vessel, heat exchangers, surge line, pressurizer vessel, main cooling pumps and pipe lines. However, as in different countries life extension to an over all life time of 80 years is in discussion in future we have to take into account the infrastructure, i.e. bridges nearby, important for fluid traffic, emergency current generators, the concrete components of the containment and the cooling towers but also ageing phenomena of electric cable insulation, etc.
Within these life time extension strategies the methodology of a continuously applied ageing management worldwide is seen as an important measure to guarantee nuclear safety. Besides the application of standardised non-destructive testing (NDT) technology during inservice inspection trials in order to perform a diagnosis of the material states on-line structural health monitoring of components by enhanced and intelligent NDT-sensors and sensor-networks will play a forthcoming future role.
In Germany actually code-accepted procedures to perform ageing management were finally discussed and approved by the authorities. However, research and development in the last

decade in the Nuclear Safety Research Programme of the German Ministry for Economy and Technology was continuously performed in order to develop and qualify NDT-technology for characterisation of ageing phenomena.
The here presented chapter describes the objectives of this research and the final results obtained. In any case, the methodology of the micromagnetic NDT procedures was especially developed. This methodology is suitable for materials characterisation of magnetisable steels in terms of determination of mechanical properties. There are many similarities between movements of dislocations under mechanical loads and pinning of this lattice defects at vacancies, precipitates, grain and phase boundaries, contributing to the strength of the material and the movement of magnetic domains under magnetic loads, i.e. when the material is magnetised in a hysteresis loop.
The methodology of the Micromagnetic, Multiparameter, Microstructure and Stess Analysis (3MA) is discussed which on a wide basis of different diverse as well as redundant information allows the sensitive materials characterisation.
In case of a Cu-rich steel alloy precipitation hardening is discussed in combination with thermal ageing. It is shown that superimposed fatigue loads will enhance the thermal ageing effect.
Fatiguing of austenitic stainless steel under some conditions is combined with phase transformation from the face-centred-cubic (fcc) lattice to body-centred-cubic (bcc) martensitic phase which is ferromagnetic of nature. Where the carbon content is low enough to avoid the phase transformation other NDT techniques based on electric conductivity effects or ultrasonic wave propagation phenomena have to be applied.
3MA is sensitive to characterise neutron embrittlement in pressure vessel materials. Material of western pressure vessel design as well as of Russian design were characterised which shows that a new NDT technology for inservice inspection of the pressure vessel wall from the id-surface can be developed.

2. Micromagnetic properties and micromagnetic NDT, the 3MA approach

The reason to develop 3MA (Micromagnetic-, Multiparameter-, Microstructure-, and stress-Analysis) by Fraunhofer-IZFP, starting in the late seventies in the German nuclear safety program, was to find microstructure sensitive NDT techniques to characterise the quality of heat treatments, for instance the stress relieve heat treatment of a weld. George Matzkanin (Matzkanin, 1979) just has had published a NTIC report in the USA to the magnetic Barkhausen noise, nowadays very often called magnetic Barkhausen emission (MBE). The technique was sensitive to microstructure changes as well as to load-induced and residual stresses. Therefore a second direction of research started in programs of the European steel industry and the objective was to determine residual stresses in big steel forgings, like turbine shafts. Beside the magnetic Barkhausen effect a magneto-acoustic-one became popular (Theiner & Waschkies, 1984). The technique has based on acoustic emission measurements during controlling the magnetisation in a hysteresis cycle but was – because of the high amplification – sensitive to electric interference noise. Therefore the acoustic Barkhausen noise technique has never found a wide-spread real industrial application. However, because influenced by only 90° Bloch-wall interactions in the laboratory the technique was an ideal sensor to enhance the basic understanding. Later, further micromagnetic techniques were developed: the incremental permeability measurement, the harmonic analysis of the magnetic tangential field and the

measurement of the dynamic or incremental magnetostriction by use of an EMAT (Altpeter, 2002). Basically the idea of Paul Höller, a former director of IZFP, was to develop the micromagnetic techniques in order to replace electron microscopy, i.e. to find algorithms to determine microstructural parameters like vacancy density and distribution, dislocation density and distribution, precipitation density and distribution, etc. However, because the need for calibration of the micromagnetic measuring parameters, the researchers very quickly understand to follow the more pragmatic direction, i.e. to directly correlate the micromagnetic properties with mechanical technological parameters like yield strength or hardness. The main argument for that decision was the large scatter in electron microscopy data and the fact of the strong propagation of errors in the calibration, when based on these data.

2.1 Micromagnetic basics

Ferromagnetic materials - even in a demagnetised or 'virgin' state - consist of small, finite regions called domains (Kneller, 1966; Seeger, 1966; Cullity, 1972; McClure & Schröder, 1976). Each domain is spontaneously magnetised to the saturation value of the material. The directions of magnetisation of the various domains, however, are such that the specimen as a whole shows no net magnetisation. The process of magnetisation is then one of converting the multi-domain state into a single domain magnetised in the same direction as the applied magnetic field H. The process is performed not continuously but stepwise by movement of the domain walls, named Bloch walls, and for stronger applied fields, by rotation of the magnetisation vectors in the domains into the direction of the applied field. In iron-based materials we find 180° and 90° Bloch walls. The indicated angle is the angle between the magnetisation vectors in two adjacent domains. Domains with directions parallel or nearly parallel to the magnetising field increase in their size while all others are annihilated. The Bloch wall movements as wall jumps take place discontinuously because the walls in a polycrystalline material are temporarily pinned by lattice defects as microstructural obstacles like dislocations, precipitates, phase- or grain-boundaries. The stepwise pull-out of the wall from the obstacle changes the magnetisation state locally and is called a Barkhausen event. The local magnetisation changes induce pulsed eddy currents in the vicinity of the events propagating in all spatial directions. The amplitudes of the eddy currents are damped according to a well known dispersion law, i.e. higher frequencies in the spectrum are damped more than lower frequencies. The eddy currents induce electrical voltage pulses, called Barkhausen noise, which may be detected by an induction coil surrounding the magnetised specimen. The time domain integral of the Barkhausen noise during a magnetisation reversal is the magnetic induction B, and B versus H is the hysteresis loop. Figure 1 documents the influence of different microstructures (ferrite / martensite) and Figure 2 presents the influence of load-induced or residual mechanical stresses. 'Magnetic hard' materials, like a martensitic steel microstructure, have larger coercivity (tangential magnetic field value H_t at B=0) and smaller remanence (B-value at H_t=0). In case of compressive stresses the hysteresis is sheared and tensile stress generates a slighter curve. Whereas the hysteresis - by definition - is measured by a coil surrounding the specimen under magnetisation, more suitable pick-up techniques for local, spatially resolved ND-testing have been developed.

Ferromagnetic materials show the property of magnetostriction (Cullity, 1972). When exposed to a magnetic field, its dimension changes. The effect can be measured as a function

of the applied magnetic field in profile-curves ($\lambda(H_t)$ or λ_L (H_t), L-lengthwise magnetised cylindrical specimen) by using strain gages (Figure 3 and Figure 4). The inverse effect, called Villari effect, is the spontaneous magnetisation of a virgin ferromagnetic material exposed to a mechanical stress field. This is the reason of the stress sensitivity of micromagnetic NDT (Sablik, Burkhart, Kwun, & Jiles, 1988; Burkhart & Kwun, 1989; Jiles 1988). In the lower part of Figure 4, according to a 4-domain model, the effect of increasing tensile stress on the domain wall movement is discussed. The domains with magnetisation direction in the tensile stress direction are preferred.

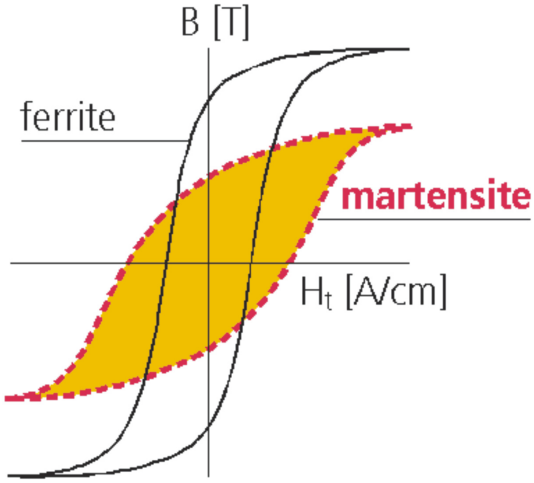


Fig. 1. Hysteresis loop and influence of microstructure

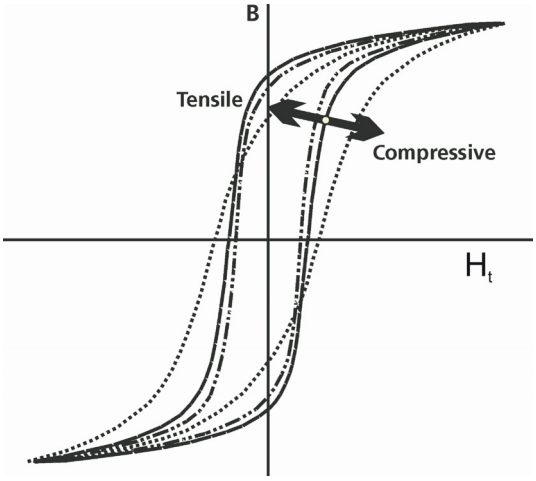


Fig. 2. Hysteresis loop and influence of mechanical stress

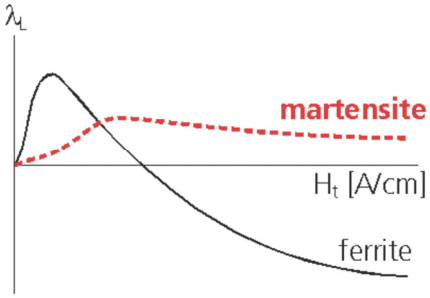


Fig. 3. Lengthwise magnetostriction and influence of microstructure

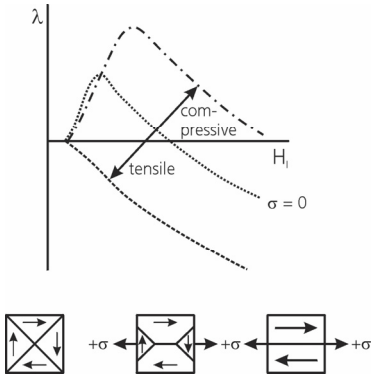


Fig. 4. Lengthwise magnetostriction and influence of mechanical stress

2.2 Micromagnetic techniques in detail

For the development of non-destructive (ND)-techniques, some special principles have been identified from the above mentioned basic physical facts, providing a selective interaction with the microstructure and stress fields. These are particularly micromagnetic techniques.
Depending on the magnetic process performance irreversible and reversible processes have to be distinguished. Techniques using irreversible processes take advantage of the non-linearities of the hysteresis which are the result of the above mentioned Bloch-wall jumps and rotational processes (Jiles, 1990).
A distortion factor K can be derived from a Fourier analysis of one period of the time-signal of the magnetic tangential field strength H_t which is measured by a Hall element. K is the geometrical average (root-mean-square-value) of the power of the upper harmonics, normalised to the power of the fundamental (Pitsch, 1989). The distortion factor K is defined as shown in Figure 5. By empirical investigations it was documented that from the time dependence of the sum of all harmonics above the fundamental a coercivity value H_{C0} may be derived. This value correlates with the hysteresis coercivity H_C. The investigations are confirmed by other authors (Fillon, Lord, &Bussiere, 1990). Thus, using this measurement of the tangential field strength, with only one sensor, two independent parameters (K and H_{C0}) may be derived. The depth that is analyzed is controlled by the penetration depth of the applied field H_t.

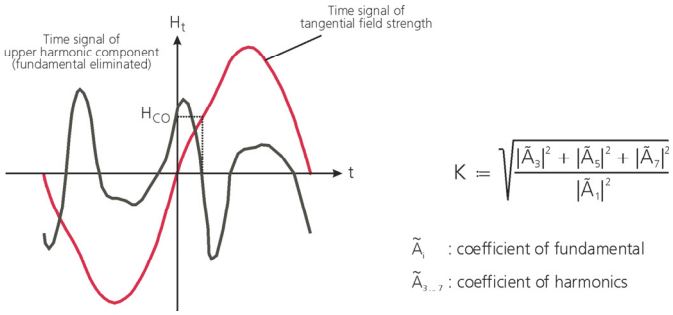


Fig. 5. Fourier analysis of the magnetic tangential field

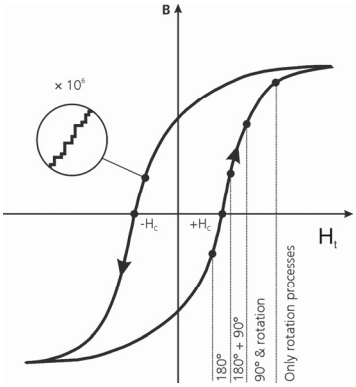


Fig. 6. Hysteresis, Barkhausen noise, micromagnetic events: Bloch wall jumps, rotation processes

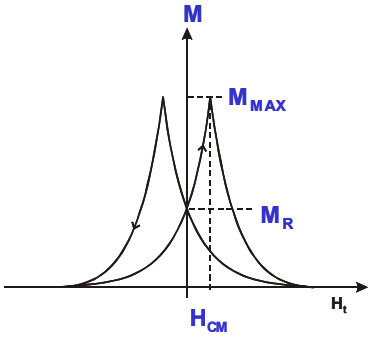


Fig. 7. Magnetic Barkhausen noise profile curve

The Barkhausen noise is magnetically received – as said before - directly by a pick-up air-coil, a ferrite-core-coil, or other inductive sensor as induced electric voltage pulses. After low-noise pre-amplification (60 dB fixed), band-pass filtering (cut-off frequencies to be adjusted), rectifying, low-pass-filtering in order to receive the envelope of the high-

frequency content of the noise, and final amplification (variable up to 60 dB), the magnetic Barkhausen noise profile-curve $M(H_t)$ is obtained (Theiner, Altpeter, & Reimringer, 1989). This type of measurement technique was proposed by Matzkanin.
The spatial resolution is restricted by the sensor geometry. The mm-range can be achieved by using standard pick-up air coils. The band pass filter is needed to suppress the large magnitude of the fundamental magnetising frequency in the received signal in addition to its higher harmonics. This effect is stronger for magnetic soft materials than for hard ones. Soft materials have an essential higher harmonic content. Harmonics up to the 7th order are observed. The lower cut-off frequency of the band pass filter therefore has to be adjustable, depending on the material under inspection. It should be mentioned here, that this measure to suppress disturbing effects and to only receive the random Barkhausen noise limits on one hand the analyzing depth for technical steels in general to a depth of ≈ 2 mm. Because on the other hand the analyzing depth of the above mentioned distortion factor measurement is limited only by the magnetising frequency, this parameter K completes the information of the Barkhausen noise to a larger analyzing depth. Because of the random character of the Barkhausen-noise a large natural scatter in the data is observed due to small magnetic fluctuations from the environment superimposed on the driving field. Time averaging (up to ten magnetising cycles) is a help to obtain reliable data. In process applications at higher inspection speeds, the effect can hinder the use of Barkhausen-noise.
Figure 6 shows a hysteresis loop as a continuous curve. Only with a higher amplification the fact of the discontinuous magnetisation in Bloch wall jumps is revealed. The Figure documents the fact of the interaction of the different Bloch-wall types at different magnetic field amplitudes. In the vicinity of the coercivity mainly the 180° walls contribute with their jumps to the magnetising process. The 90° walls need more energy for movement and therefore their contribution to magnetisation is in the knee region of the hysteresis followed by rotation processes.

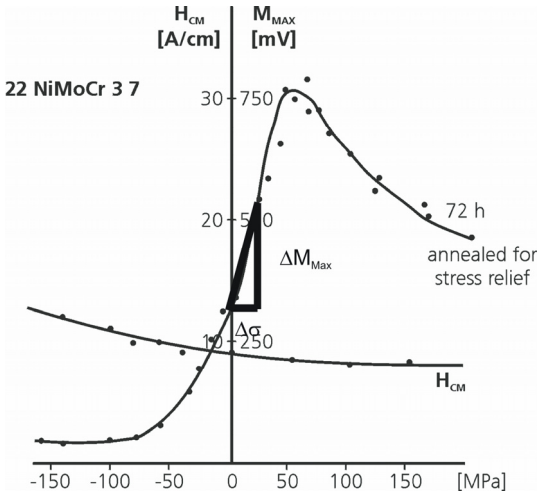


Fig. 8. Barkhausen noise peak maximum and peak separation as function of mechanical stress and stress sensitivity $\Delta M_{Max} / \Delta\sigma$ of the German pressure vessel steel 22 NiMoCr 37 after welding and 72h stress relief annealing

Fig. 7 shows a Barkhausen noise profile-curve ($M(H_t)$) and the measuring parameters are presented. For some steel grades we can observe profile-curves with up to three peaks per magnetising half-cycle, depending on the materials microstructure. Besides the different peak maxima (M_{Maxi}, i=1,…, 3) their separation (H_{CMi}) as well as different half-width-values (for instance ΔM50%) are measured.

For residual stress characterisation (see Figure 8), the peak values are measured as a function of compressive and tensile stress, M_{Max} versus stress σ. The magnetic stress sensitivity of a material microstructural state is characterised by the function $\Delta M_{Max}(\sigma) / \Delta\sigma$ i.e., the inclination of the curve $M_{Max}(\sigma)$ at $\sigma = 0$. The fact of the decreasing of the peak amplitude of the Barkhausen noise at higher tensile stress levels is observed at a mechanical stress value in the tensile regime at which the magnetostriction curve under magnetisation (see Figure 4) starts directly with negative values.

Due to magnetostrictive effects, Barkhausen-events excite, in their vicinity, pulsed magnetostrictive strains which result in magnetostrictively excited acoustic emission signals named acoustic Barkhausen noise (Theiner & Waschkies, 1984). This is named as MAE – magneto-mechanical-acoustic-emission in literature too (Buttle & Hutchings, 1992; Allen & Buttle, 1992). The acoustic emission (AE) is received by narrow band, sensitive AE-transducers. Nevertheless, a high amplification of ~ 80 - 100 dB is needed and the technique is sensitive to disturbing noise. Typical AE-signal analysis equipments like RMS-voltmeters or energy-modules are in use. Because of its sensitivity for disturbing noise, the technique can reliably be applied only in the laboratory. Nevertheless investigations with this technique are of interest for comprehensive interpretation studies because here only the information of 90° Bloch wall movement and their interaction with the microstructure is observed.

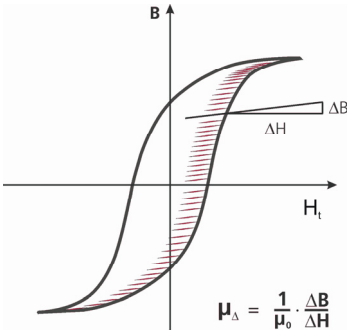


Fig. 9. Incremental permeability

Applying reversible magnetic techniques, the material is magnetised with a much smaller field strength amplitude ΔH compared to the coercivity (H_C) of the material. The magnetisation follows linearly the magnetic field. Therefore, all eddy current techniques using low electric currents in the pick-up coil, resulting in small magnetic fields ΔH ($\Delta H << H_C$), are reversible techniques. Parameters influencing the coil impedance are the operating frequency f_Δ, the electrical conductivity σ_{el} and the incremental permeability μ_Δ. The latter depends on the magnetic history of the specimen under inspection. An eddy current impedance spectroscopy is performed by tuning the frequency from high to low values. (σ_{el}, μ_Δ) - gradients in near surface zones are sensed by changing the field penetration. The spatial resolution is the same as that for eddy current coils.

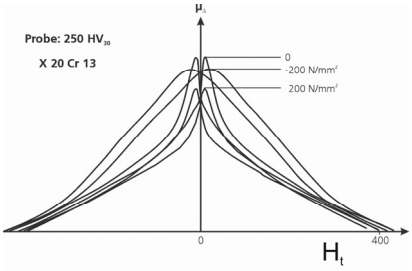


Fig. 10. Incremental permeability profile curves documenting the influence of mechanical stress, steel quality X20Cr13

The incremental permeability profile-curve, $\mu_\Delta(H_t)$, as a function of a controlled applied magnetic field H_t is a well defined property of the material and independent of the magnetic prehistory as long as $H_{Max} >> H_C$ and $\Delta H << H_C$. The frequency f_Δ of the incremental field ΔH is a parameter for selecting the depth of the analyzed near surface zone; f_Δ should be chosen such that $f_\Delta \geq 100 \times f$, where f is the frequency of the applied field H_t , controlling the hysteresis. $\mu_\Delta(H_t)$ is measured as eddy current impedance parallel to the hysteresis reversals. The hysteresis is modulated by the alternating field ΔH, excited by the eddy current coil. The spatial resolution is the same as that for eddy current coils. Figure 9, shows the hysteresis with the inner loops, performed by the above mentioned modulation. By definition $\mu_\Delta(H_t)$ is proportional to the inclination of each individual inner loop touching the hysteresis for magnetic field values H_t. In Figure 10 $\mu_\Delta(H_t)$ profile-curves are presented, indicating the characteristic measuring parameters as function of mechanical stresses.
The dynamic or incremental magnetostriction profile-curve $E_\lambda(H_t)$ is the intensity of ultrasound which is excited, and received by an EMAT (Electro-Magnetic-Acoustic-Transducer) (Salzburger, 2009) for instance by measuring a back-wall echo, caused by magnetostrictive excitation as a function of the applied field H_t, controlling the hysteresis. The incremental, alternating field ΔH in this case is excited by the EMA - transmitter using a pulsed current.
The magnetostriction is modulated. (Figure 11, upper part) The spatial resolution - depending on the transmitter design - is of the order of ~ 5 mm. In order to achieve such a spatial resolution, an EMA - receiver was designed to transform the ultrasonic signal into an electrical signal only using the Lorentz-mechanisms (Koch & Höller, 1989). Figure 11, lower part, presents a half-cycle of the dynamic magnetostriction profile-curve $E_\lambda(H_t)$ in the magnetic field range < 300 A/cm. The amplitude value of the first peak as well as the corresponding tangential magnetic field value as well as the H_t-field position of the minimum are sensitive quantities for materials characterisation.
The micromagnetic measurements are performed by an intelligent transducer consisting of a handheld magnetic yoke together with a Hall-probe for measuring the tangential magnetic field strength and a pick-up coil for detecting the magnetic Barkhausen noise or the incremental permeability. Normally a U-shaped magnetic yoke is used, which is set onto the surface of the material under inspection, i.e. the ferromagnetic material is the magnetic 'shunt' of the magnetic circuit. Therefore all the well known design rules for magnetic circuits have to be observed. The mathematical methodology of the Micromagnetic-, Multiparameter-, Microstructure-, and stress- Analysis (3MA) in detail is described in (Altpeter, 2002). However, a short explanation is given here according to Figure 12.

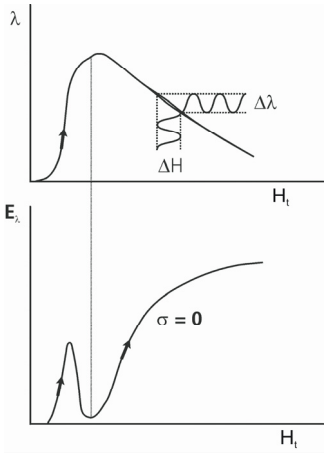


Fig. 11. Dynamic or incremental magnetostriction

With 3MA different micromagnetic quantities, let's say X_i, i = 1, 2, 3, … are measured at 'well defined' calibration specimens. These are derived by analysis of the magnetic Barkhausen noise $M(H_t)$ and the incremental permeability $\mu_\Delta(H_t)$ as function of a tangential magnetic field H_t which is analysed and by eddy current impedance measurements at different operating frequencies.

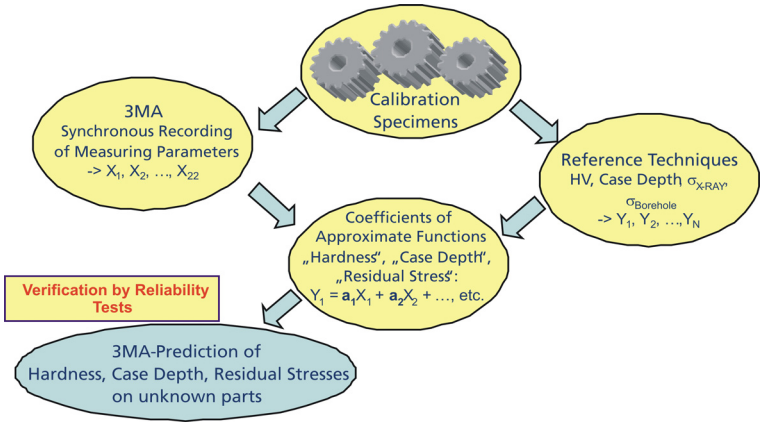


Fig. 12. The 3MA-calibration

'Well defined' here has the meaning that the calibration specimens are reliably described in reference values like mechanical hardness (according to Vickers or Brinell, etc.) or strength values like yield and/or tensile strength, or residual stress values measured, for instance, by X-ray diffraction. A model of the target function is assumed (for instance Vickers Hardness $HV(X_i)$, or strength value like $Rp0.2(X_i)$, or residual stress $\sigma_{res}(X_i)$). This model is based on the development of the target function by using a (mathematically) complete basis function system, which is a set of polynomials in the micromagnetic measurement parameters X_i. The unknown in the model are the development coefficients, in Figure 12 called a_i. These a_i are

determined in a least square or other algorithm, minimizing a norm of the residual function formed by the difference of the model function to the target reference values. In order to stochastically find a best approximation, only one part of the set of specimens is used for calibration of the model, the other independently selected part is applied to check the quality of the model (verification test). By using for instance the least square approach the unknown parameters are the solution of a system of linear equations.
3MA is especially sensitive to mechanical property determination as the relevant microstructure is governing the material behaviour under mechanical loads (strength and toughness) in a similar way as the magnetic behaviour under magnetic loads, i.e. during the magnetisation in a hysteresis loop. Because of the complexity of microstructures and the superimposed stress sensitivity there is an absolute need to develop the multiple parameter approach.

3. NDT characterisation of thermal ageing due to precipitation

Beginning in 1998 Fraunhofer-IZFP in co-operation with the Materials Testing Institute at the University Stuttgart (MPA) (Altpeter, Dobmann, Katerbau, Schick, Binkele, Kizler, & Schmauder, 1999) has investigated the low-alloy, heat-resistant steel 15 NiCuMoNb 5 (WB 36, material number 1.6368) which is used as piping and vessel material in boiling water reactor (BWR) and pressurized water reactor (PWR) nuclear power plants in Germany. One argument for its wide application is the improved 0.2% yield strength at elevated temperatures.
Conventional power plants use this material at operating temperatures of up to 450°C, whereas German nuclear power plants apply the material mainly for pipelines at operating temperatures below 300°C and in some rare cases in pressure vessels up to 340°C (e.g., a pressurizer in a PWR). Following long hours of operation (90,000 to 160,000 h) damage was seen in piping systems and in one pressure vessel of conventional power plants during the years 1987 to 1992 (Jansky, Andrä, & Albrecht, 1993) which occurred during operation and in one case during in service hydro-testing. In all damage situations, the operating temperature was between 320° and 350°C. Even though different factors played a role in causing the damage, an operation-induced hardening associated with a decrease in toughness (-20%) was seen in all cases. The latter is combined with a shift in the transition temperature of the notched-bar impact test to higher temperatures (+70°C) and in the 0.2% yield strength of about +140 MPa.
According its specification the steel has in between 0.45 and 0.85 mass% Cu (in average 0.65%) in its composition. The half part of the Cu is in precipitation because of annealing and stress relieve heat treatment during production, the other half still is in solid solution and can precipitate when the material is exposed at service temperatures. The material can obviously be recovery annealed when after the service exposures again is heated-up at the stress relieve heat treatment temperature and hold some time. The precipitates are dissolved again in solid solution obtaining a microstructure state comparable but not identical to the 'as delivered' state.
Micromagnetic investigations at first were performed at 'service exposed' (57,000h at 350°c) and 'recovery annealed' (service exposed + 3h 550°C) material using cylindric (diameter 8mm) test specimens. Whereas the hysteresis curves of the two microstructure states are nearly identical, differences were observed when the magnetic Barkhausen noise was

registered and when the lengthwise magnetostriction was measured. The specimens were measured in the stress-free state as well under variable tensile load (according to Fig. 8) in order to reveal the stress sensitivity of the microstructures.

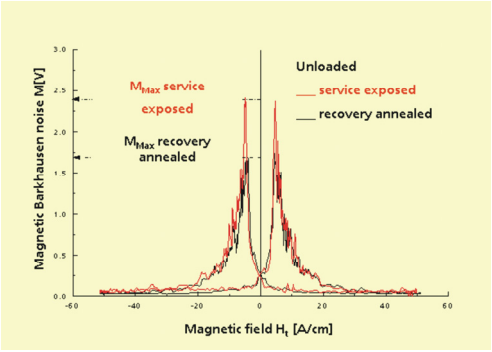


Fig. 13. Magnetic Barkhausen noise of service-exposed and recovery annealed WB36 microstructures in the stress-free state

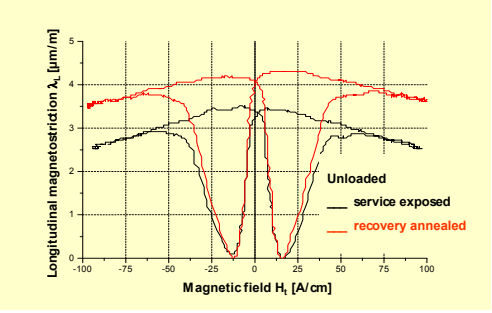


Fig. 14. The lengthwise magnetostriction of the microstructure states of Fig. 13

In Figure 13 the profile curves of the magnetic Barkhausen noise related to the two material states are shown and Figure 14 documents the behaviour of the magnetostriction in the stress-free state.

The service exposed microstructure has higher Barkhausen noise maximum and lower magnetostriction values. Both effects indicate the influence of tensile residual stresses induced by the Cu-rich precipitates in the iron matrix. In TEM and SANS investigations the precipitation state was studied. The particle size is in between 2nm – 20nm distributed. Particles < 6nm diameter have body centered cubic crystallographic structure like the iron matrix (coherent precipitates). As the atomic radius of Cu is larger compared with iron the Cu precipitate acts with compressive stresses which are balanced by tensile residual stress in the matrix. Particles with diameter > 20nm are face centered cubic and in between these two states a transition crystallographic structure exist. About 50% of the precipitates have this transition structure and especially contribute to micro residual stresses in the tensile stress regime in the matrix. Figure 15 shows the like coffee-beans shaped particles of the transition structure visible in the TEM and the diffraction pattern.

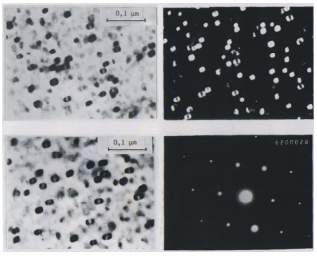


Fig. 15. TE micrographs and diffraction pattern of the Cu particles

Fraunhofer-IZFP has performed experiments under load-induced tensile stresses too. Figure 16 and Figure 17 show the result at the service exposed and recovery annealed microstructures. As discussed in Figure 8 the Barkhausen noise maximum Mmax (σ) as function of the tensile load σ increases with the load to an absolute maximum and then decreases again. The threshold load where this maximum occurs is exactly the load value where magnetostriction becomes directly negative in sign when the specimen additionally is magnetised.

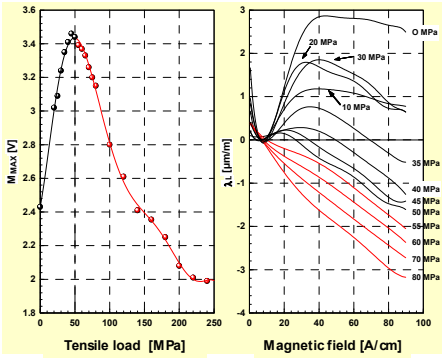


Fig. 16. The service exposed microstructure

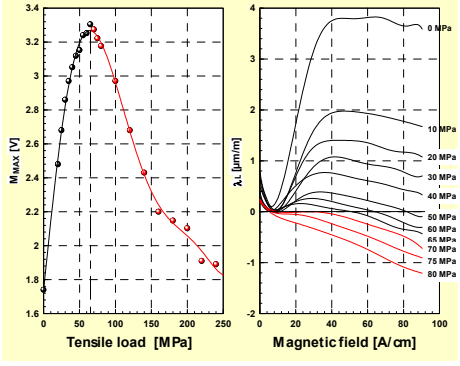


Fig. 17. The recovery annealed microstructure

Comparing the two microstructure states in its stress sensitivity the difference in the residual stress state due to the Cu precipitates can only be responsible to shift the maximum position about 17-20 MPa to smaller tensile loads in the case of the service exposed material. This value should be the amount of the average residual stress in the iron matrix which locally near the precipitate can be much higher but cannot be measured with another reference technique.
Further investigations in order to statistically confirm the results were performed at 400°C in order to speed-up the precipitation process.

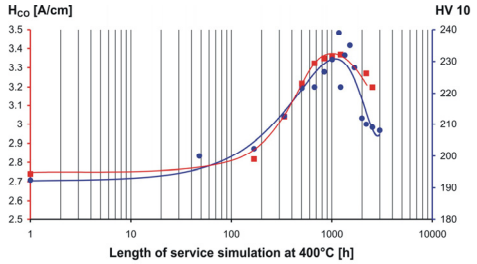


Fig. 18. Coercivity H_{C0} derived from the harmonic analysis of the tangential magnetic field strength and Vickers hardness 10

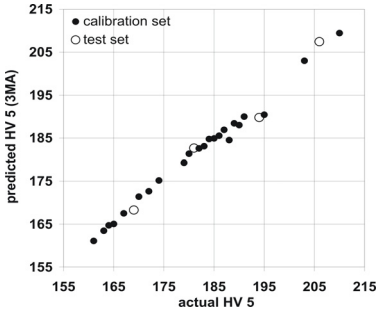


Fig. 19. 3MA approach to characterise the Cu precipitation microstructure state in terms of Vickers hardness 5

Comparing the coercivity (Figure 18.) derived from the harmonic analysis of the tangential magnetic field strength with the measured Vickers hardness 10 as reference to characterise the thermally aged microstructure both quantities are correlated and meet a typical hardness maximum which is the critical material state for possible failure of a component if the design has not taken into account the strengthening ageing effect. When the exposure times are further enlarged hardness is decreasing by precipitation coarsening. In order to obtain the good correlation in the 3MA-approach beside micromagnetic characteristics eddy current impedances were implemented. These are especially suitable as the Cu precipitates contribute to an enhanced electrical conductivity.
Parallel to the project activities in the German nuclear safety program a PhD thesis (Rabung, 2004) was performed in different projects of the German National Science Foundation (DFG). Fe-Cu-model-alloys were investigated mainly to study the effect of the Cu

precipitates without influences of the magnetic cementite-phase as in WB36. The Cu-content was varied in between 0.65% and 2.1% (Altpeter, I., Dobmann, G., Kröning, M., & Rabung, M., 2009).

There was always the supposition that any form of energy, other than only heat, put in WB36 components will contribute to enhanced precipitation of Cu particles. The effect of low cycle fatigue at service temperature was therefore studied in a further project in the German nuclear safety program the last 4 years (Altpeter, I., Szielasko, K., Dobmann, G., Ruoff, H., & Willer, D., 2010). As in literature (Solomon & De Lair, 2001) dynamic strain ageing (DSA) was expected in the lower temperature regime (200°C) to be additionally a driver for WB36 thermal degradation two different heats were selected which were different in the Al/N-ratio in the composition. Because of the higher N content (Al/N (E2)=0.92) the heat E2 was assumed to be more prone for DSA than the heat E59 (Al/N (E59)=3.87). E2 material came from a plate in the virginal condition ('as delivered state'), named E2A. The E59 material came from a used vessel which at 350°C for 57,000 h was in service. The material was investigated in the state 'recovery annealed' (600°C, 3h) named E59 EG. Furthermore, some material of E59 was especially heat treated, 'stepwise stabilised annealed' in order to stabilise the Cu precipitation distribution in coarse particles, named E59 S4. Compared with E59 EG, E59 S4 should be less prone for further precipitation development under service conditions.

Under LFF-conditions (mean strain-free, R_ε=-1, strain-controlled with $\Delta\varepsilon$=1.05% at 220°C and 300°C) specimen of the heat E2A were cycled in one-step fatigue tests with cycle period (24s, load cycles 350 at 220°C; 2400s, load cycles 200 at 220°C; 2400s, load cycles 200 at 300°C). The expected material behaviour was confirmed, i.e. degradation will be enhanced by accumulated elastic-plastic deformation; Figure 20 represents the result in terms of Charpy-test-energy versus test temperature. As documented, the 41J ductile-to-brittle transition temperature (DBTT, ΔT41J) shifts to higher temperature with plastic deformation energy input.

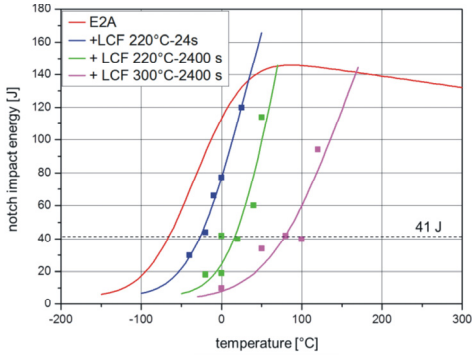


Fig. 20. Charpy tests at different thermally aged and LC fatigued material states, documentation of material degradation of the heat E2A

A maximum shift ΔT41J of 144.3°C can be observed. It should be mentioned here that the tests performed with the heat E59EG have shown much smaller effects in degradation, documenting the fact that the microstructure states in the state 'as delivered' and 'recovery annealed' are not identical when exposed to further thermal ageing and fatigue.

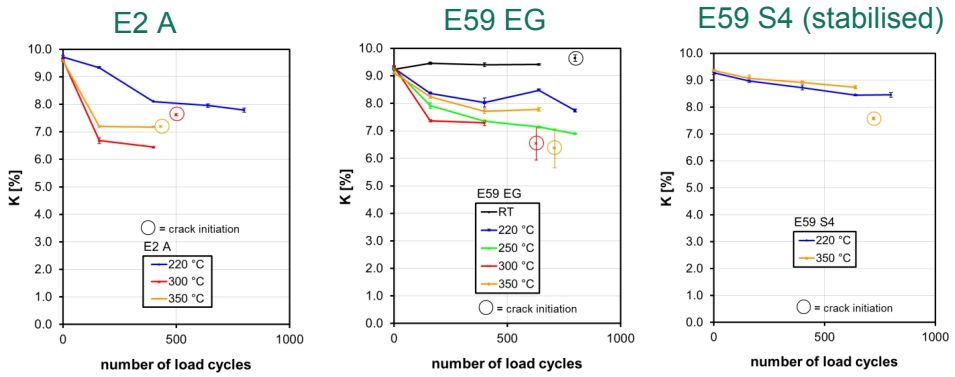


Fig. 21. Distortion factor K as function of cycle number measured after well defined fatigue intervals in test interruptions followed by a further fatiguing of the same specimen

A very wide space for investigations was addressed to interval tests where all of the 3 heats were fatigued mean strain-free with $\Delta\varepsilon$=1.05% and a cycle period of 2,400s at different elevated temperatures (E2A (220°C, 300°C, and 350°C; E59EG (220°C, 250°C, 300°C, and 350°C; E59S4 (220°C, 350°C). The specimen were fatigued to a certain load cycle number in terms of a fraction of the average live time (N_a-averaged cycle number to failure, N=0.2 N_a, N=0.5N_a, N=0.8N_a, and N=N_a=800 cycles). The test then was interrupted for non-destructive tests followed by further fatiguing, etc. The over all result can be presented in micromagnetic life-cycle diagrams as shown exemplarily for instance in case of the measuring quantity K (distortion factor of the tangential field strength, measured according to Fig. 5) in Figure 21.

Concerning the decreasing of K the material states of E2A show the strongest effect compared with the E59EG states in case of the fatigue test temperature of 300°C. The decrease here is stronger than for the test temperature of 350°C. Obviously most of the decrease is in the first fatigue time interval, followed only by a moderate further decreasing, what allows the interpretation that due to strain hardening and dislocation development local precipitation sources are generated enhancing the Cu precipitation. K seems more influenced by the dislocation strengthening effect than by the precipitation what is seen in the secondary fatigue interval. However, very rapidly critical material states are obtained which is documented by the fact that all specimens under these conditions early failed in the following fatigue intervals.

As the first decrease in E2A fatigued at 350°C is smaller compared to the 300°C test the strain hardening effect seems to be smaller, may be, due to recovery effects by transverse dislocation slipping. This is confirmed by measurement of the volume fraction of Cu precipitates performing SANS.

The smallest effects are observed with the stabilised annealed material. As due to this processing most of the Cu content is precipitated in coarse particles the decrease in K is mainly due to fatigue effects (dislocation cell development and cell size change and arrangement) and not due to thermal ageing, i.e. further pronounced precipitation.

The overall 3MA approach, by taking in addition other 3MA-quantities in account and combining these, has used a generic algorithm (Szielasko, 2001) for prediction of the Vickers

hardness 10 and the G- value (Figure 22, Figure 23) with very high confidence level and regression coefficient. The G-value is the electrical residual resistance ratio which is defined as the ratio of the specific resistance measured at ambient temperature to the specific resistance measured at nitrogen temperature. G is a measure of impurity (foreign atoms in the iron matrix) of a material and here therefore is a direct measure of the Cu content of the precipitates. The MPA measures G very carefully in the laboratory and has compared the results with SANS measurements. There is a linear correlation (Figure 24).

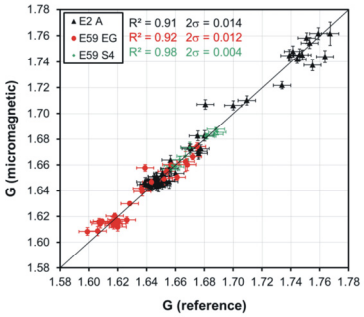


Fig. 22. 3MA prediction of the G-value

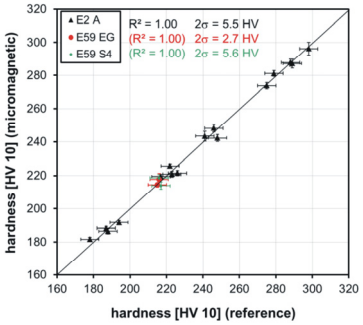


Fig. 23. 3MA prediction of the Vickers hardness 10

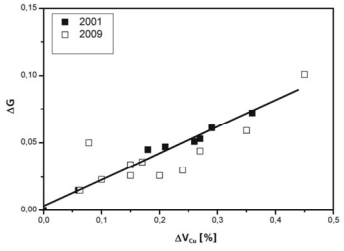


Fig. 24. Change in the G-value (ΔG) compared with the change in Vol.% Cu precipitation (ΔV_{Cu}) determination with SANS (measurements from the years 2001 and 2009)

With 3MA there is therefore a reliable ability to characterise the degradation in terms of the Cu precipitates volume fraction as well as in hardness.
Concerning the expected DSA-effects the investigations have shown serrations in the stress-strain diagrams only in the small temperature window 130-185°C. At service temperature it does not play a role.

4. NDT characterisation of fatigue at austenitic stainless steels

Activities to the non-destructive characterisation of fatigue phenomena at austenitic steels were performed in a co-operation with the Institute of Material Science and Engineering of the Technical University Kaiserslautern, Germany and started in 1999 with 2 PhD thesis's (Bassler, H.J., 1999; Lang M., 1999).
Austenitic steel of the grade AISI 321 (German grade 1.4541 - Ti-stabilised and AISI 347 German grade 1.4550 - Nb-stabilised) is often used in power station and plant constructions. The evaluation of early fatigue damage and thus the remaining lifetime of austenitic steels is a task of enormous practical relevance. Meta-stable austenitic steel forms ferromagnetic martensite due to quasi-static and cyclic loading. This presupposes the exceeding of a threshold value of accumulated plastic strain. The amount of martensite as well as its magnetic properties should provide information about the fatigue damage. Fatigue experiments were carried out at different stress and strain levels at room temperature (RT) and at T = 300°C. The characterisation methods included microscopic techniques such as light microscopy, REM, TEM and scanning acoustic microscopy (SAM) as well as magnetic methods, ultrasonic absorption, X-ray and neutron diffraction. Sufficient amounts of mechanical energy due to plastic deformation lead to phase transformation from fcc austenite without diffusion to tetragonal or bcc ferromagnetic α'-martensite. As the martensitic volume fractions are especially low for service-temperatures of about 300°C highly sensitive measuring systems are necessary. Besides systems on the basis of a HT_C-SQUID (High Temperature Super Conducting Quantum Interference Device) special emphasis was on the use of GMR-sensors (giant magnetoresistors) which have the strong advantage to be sensitive for DC-magnetic fields too without any need for cooling (Yashan, 2008). In combination with an eddy-current transmitting coil and universal eddy-current equipment as a receiver the GMR-sensors were used especially to on-line monitoring the fatigue experiments in the servo-hydraulic fatigue machine.

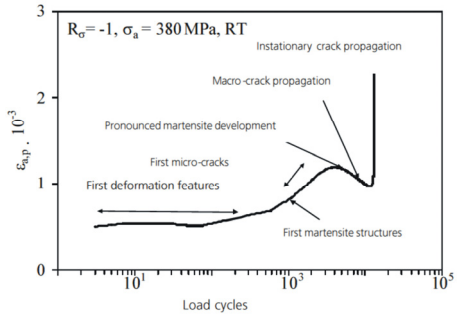


Fig. 25. Fatigue at RT

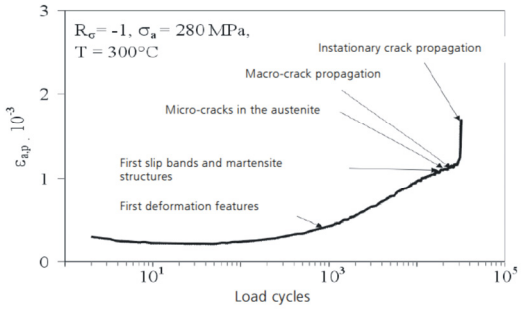


Fig. 26. Fatigue at 300°C

As function of fatigue these steels at room temperature show secondary hardening caused by continuously increasing martensite formation. Martensitic volume fractions created at service temperature T = 300°C are too small to cause cyclic hardening. Generally speaking, an accelerated martensite formation leads to shortened life times in cyclic deformation experiments. At room temperature crack initiation mainly takes place in martensitic regions (besides slip bands in the austenite phase) and often starting at carbonitrides. In martensitic regions a zigzag-shaped crack path is observed causing slower crack propagation. At T = 300°C crack initiation only occurs at slip bands. Increasing martensite formation is an indicator for increasing material damage subsequent to cyclic loading. The detection of martensite at austenitic components can be seen as a hint to local plastic deformation and thus local damage. Figure 25 and Figure 26 show the fatigue damage development and accumulation in the case of the 1.4541 material (Ti stabilised) at RT and at 300°C as can be revealed by optical and electron microscopy as function of the load cycles in cyclic deformation curves. The one-step fatigue test was performed stress controlled and mean stress-free.

By using the GMR as eddy current receiver online and in real time the fatigue experiment was monitored in the servo-hydraulic testing machine. Figure 27 (one-step fatigue tests) and Figure 28 (multiple step loading and load mix) document results obtained during online measurement in real time.

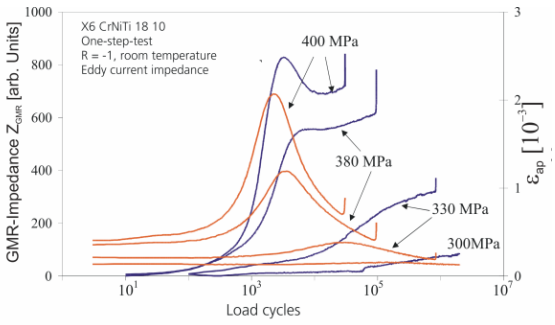


Fig. 27. One-step stress controlled fatigue tests at room temperature

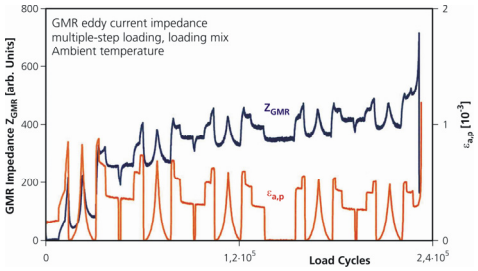


Fig. 28. Multiple step stress controlled fatigue test with time dependent load mix at room temperature

The NDT-quantity measured is the eddy current GMR-transfer impedance (Figure 27) which clearly indicates the fatigue behaviour and gives an early warning before failure. In the case where the secondary hardening effect due to the martensite formation is pronounced (stress > 380MPa) the impedance shows this secondary hardening effect too. In the multiple step experiment the impedance follows exactly the time function of the total strain but with an off-set indicating the martensite development.
It should be mentioned here that these monitoring technique was performed at plain carbon steel too. However, here the measuring effects are one order in magnitude smaller because not phase transformation to martensite takes place and only changes in the dislocation cell structure are to observe in the microstructure.
The online monitoring measuring technique was enhanced at the technical university Kaiserslautern and another type of sensor was integrated by IZFP (Altpeter, I., Tschunky, R., Hällen, K., Dobmann, G., Boller, Ch., Smaga, M., Sorich, A., & Eifler, D., 2011) into the servo-hydraulic machine. Because fatigue experiments should be monitored at service temperature of 300°C the idea was to integrate ultrasonic transducers in the clamping device of the fatigue specimen and to monitor the ultrasonic time-of-flight (tof) of a pulse propagating from the transmitter to the receiver transmitting the fatigue specimen (Figure 29). Because of the high temperature exposition coupling-free electromagnetic acoustic transducers (EMAT) were used based on a pan cake eddy current coil superimposing a normal magnetic field produced by a permanent magnet. By exciting Lorentz forces radially polarized shear waves are excited (Salzburger, 2009).

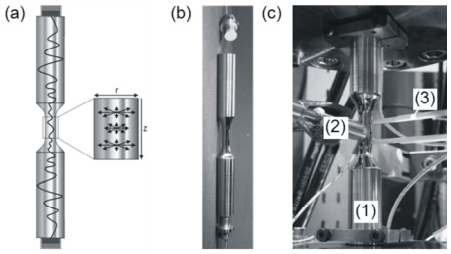


Fig. 29. Schematic diagram of wave propagation: wave propagation direction ‘z’ and particle displacement ‘r’ (a), fatigue specimen and EMAT probes with radial polarized wave type (b), clamped fatigue specimens (1) in grips, which enclose the transmitter at the one end and received at the other as well as Ferritescope (2) and an extensometer (3) (c)

The Ferritescope is used at room temperature experiments to measure the content of the developing martensite phase. The material investigated was the meta-stable AISI 347 with low C-content (0.04 weight %) and high Ni-content (10.64 weight %). Because of this fact a martensite phase transformation develops only at room temperature fatigue experiments. The fatigue tests were performed strain controlled, mean strain-free at a cycling frequency of 0.01 Hz with strain amplitudes 0.8, 1.0, 1.2 and 1.6 %. Figure 30 shows the measurement procedure to measure the tof which is determined as an average value between the maximum and minimum value obtained in each cycle (Figure 31).

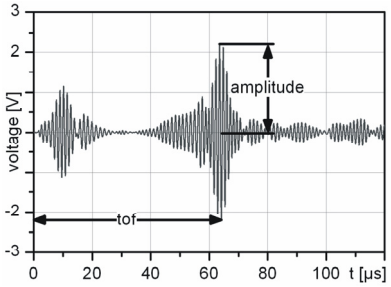


Fig. 30. Time-of-flight (tof) measurement procedure

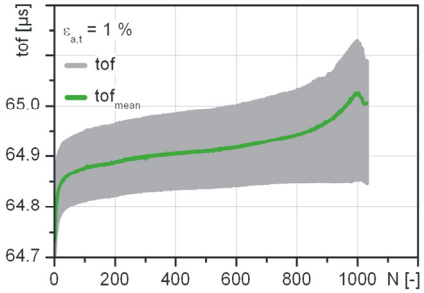


Fig. 31. Determination of the average (mean) tof-value at ambient temperature

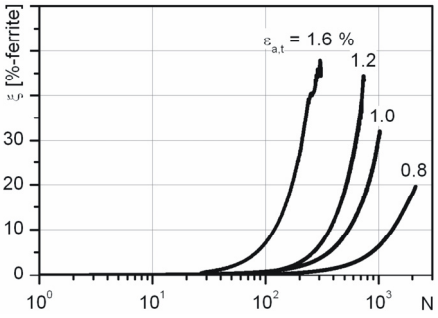


Fig. 32. Cyclic deformation curves at ambient temperature

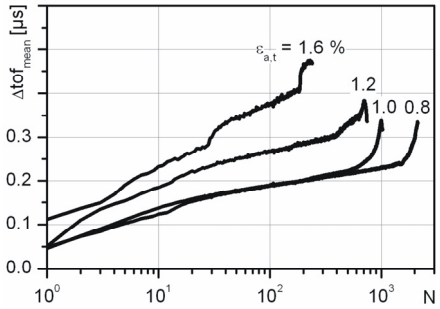


Fig. 33. Mean tof-curves at ambient temperature

The mean tof-value measured online shows a distinct behaviour as function of the fatiguing and is different in the case of ambient temperature and at 300°C. Figure 32 shows the cyclic deformation curves and Figure 33 the respective mean tof-curves where clearly the martensite development can be identified. The behaviour at 300°C is documented in the Figures 34 and Figure 35.

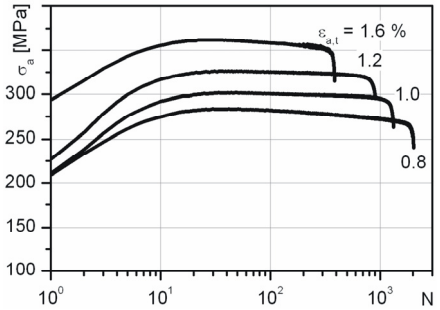


Fig. 34. Cyclic deformation curves at 300°C

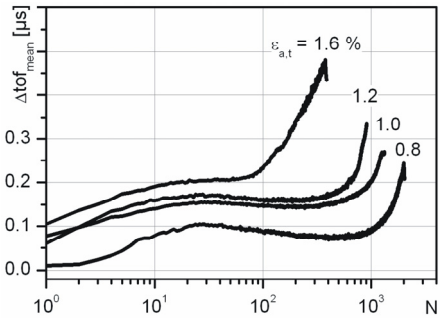


Fig. 35. Behaviour of the mean tof-values at 300°C

For further development of the tof-technique in order to be used for online monitoring of plant components Rayleigh surface waves or shear horizontal waves excited and received by EMATs will be applied.

5. NDT for characterisation of neutron degradation

In the case of power plant components, such as pressure vessels and pipes, the fitness for use under mechanical loads is characterised in terms of the determination of mechanical properties such as mechanical hardness, yield and tensile strength, toughness, shift of Ductile-to-Brittle Transition Temperature (DBTT), fatigue strength. With the exception of hardness tests which are weakly invasive, all of these parameters can be determined within surveillance programs by using destructive tests only on special standardized samples (Charpy V samples and standard tensile test specimens). The specimens are exposed in special radiation chambers near the core of the Nuclear Power Plant (NPP) to a higher neutron flow than at the surface of the pressure vessel wall in order to generate a worst case. From time to time these specimens are removed from the chambers and used for destructive tests. The number of the samples is limited and in the future it will be very important that reliable non-destructive methods are available to determine the mechanical material parameters on these samples without destruction of the specimens. Furthermore an in situ characterisation of the reactor pressure vessel inner wall through the cladding is of interest for inservice inspection, additionally to the measurements on samples.
To solve this task a combination testing technique based on 3MA and the dynamic magnetostriction measurement by using an EMAT (Electromagnetic Acoustic Transducers) was developed.

Sample set of neutron irradiated base material

Material	Short term	Neutron-fluence [n/cm²]	Range ΔT_{41} [K]
22 NiMoCr 3 7	P7	0 – 5.38 x 10^{18}	0 - 32
20 MnMoNi 5 5	P141	0 – 10.7 x 10^{18}	0 - 9
22 NiMoCr 3 7	P147	0 – 44.0 x 10^{18}	0 - 23
ASTM A508 C1.3 (22 NiMoCr 3 7)	JFL	0 – 86.0 x 10^{18}	0 - 78
ASTM A533B C1.1 (20 MnMoNi 5 5)	JRQ	0 – 98.0 x 10^{18}	0 - 221

Sample set of neutron irradiated weld material

Material	Short term	Neutron-fluence [n/cm²]	Range ΔT_{41} [K]
S3NiMo3-UP	P16	0 – 11.6 x 10^{18}	0 - 67
S3NiMo1-UP	P141	0 – 13.9 x 10^{18}	0 - 38
S3NiMo1	P140	0 – 10.4 x 10^{18}	0 - 21

Table 1. Material description of investigated Charpy samples, base and weld material

The neutron induced embrittlement results in microstructure changes. These microstructure changes are the generation of vacancies and precipitations of Cu-rich coherent particles (radius: 1-1.5 nm). This results in an increase of yield strength and tensile strength, a decrease of Charpy energy upper shelf value and an increase of DBTT.
The potential of micromagnetic testing methods for detection of Cu-precipitates was demonstrated in chapter 3. The interaction between dislocations and copper particles leads to an increase of mechanical hardness and the interaction of the copper particles and Bloch-walls leads to an increase of magnetic hardness. Since the dynamic magnetostriction is sensitive for lattice defects it was assumed that a magnetostrictively excited standing wave in the pressure vessel wall should reflect the neutron embrittlement too and first experiments were performed with a special designed magnetostrictive transducer at Charpy specimens in the hot cell in order to principally demonstrate the potential.
Using several electromagnetic measurements at the same time, a variety of measuring quantities is derived for each measurement cycle. When combined, they achieve the desired result (e.g. material property) more efficiently compared to individual measurement. By using a calibration function or pattern recognition the desired quantity of an unknown set of samples investigated by that method can be detected non-destructively.
Depending on the specific design of a pressure vessel –which varies in different countries – the pressure vessel material in nuclear power plants is exposed to neutron fluences in the range between 5.6×10^{18} n/cm² and 86.0×10^{18} n/cm². In order to characterise the neutron irradiation-induced embrittlement, Charpy samples exposed to neutron fluence in the above mentioned range have been investigated in a hot cell at AREVA NP whereby only the 3MA and EMAT sensors were arranged within the hot cell and the electronic equipment (3MA and EMAT device) was outside. These Charpy samples (base material and weld material) of Russian and western design have been provided by AREVA NP and the Research Centre Dresden-Rossendorf (see Table 1).
The result of the 3MA-approach based on pattern recognition algorithms are shown in Figure 36 (base material) and Figure 37 (weld material) in terms of changes in the DBTT evaluated at Charpy energy of 41J (ΔT_{41}).

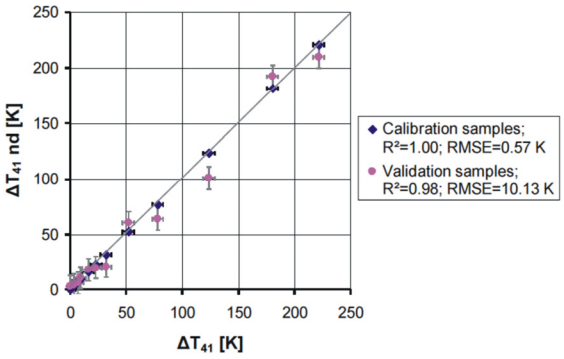


Fig. 36. 3MA prediction of ΔT_{41} in case of the base materials

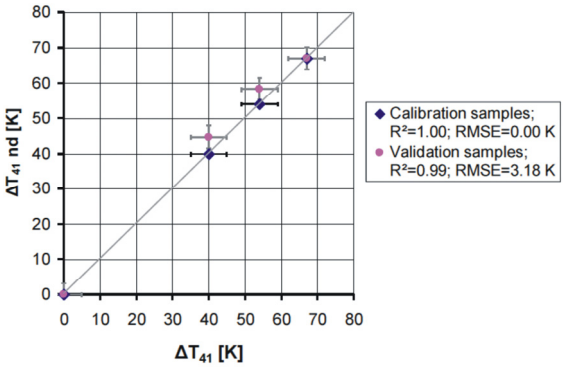


Fig. 37. 3MA prediction of ΔT_{41} in case of weld materials

As can be seen, excellent correlation coefficients on high confidence levels were obtained. Very important in this context is the contribution of the dynamic magnetostriction which is measured as the sensitivity to excite an ultrasonic standing wave in the base material under the cladding in the pressure vessel wall. Here the electrical conductivity and magnetic permeability in the ferritic-martensitic steel structure is higher than in the austenitic cladding and therefore the efficiency to excite eddy currents is enhanced but this region is also the most influenced microstructure by the irradiation.
The enhanced electrical conductivity and permeability facilitate the excitation and the large lift-off effect of the transducer due to the cladding can be compensated by the higher eddy-current density.

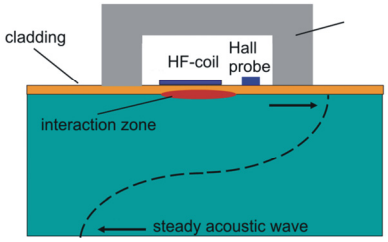


Fig. 38. Magnetostrictively excited standing wave in the PV-wall

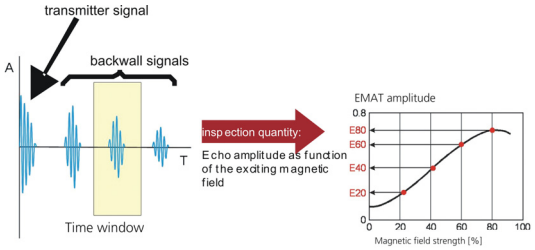


Fig. 39. Inspection quantity selected E60

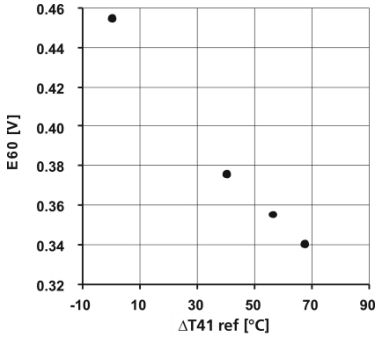


Fig. 40. Correlation of the quantity E60 with the DBTT shift at 41J Charpy energy

Figure 38 shows the measurement principle and Figure 39 defines the measurement quantity E60 which is the dynamic magnetostriction, i.e. the magnetostrictively excited ultrasound amplitude when the transducer is using a magnetic field magnitude which is 60% of the maximum field delivered by the current generator. The linear correlation of E60 with ΔT41 is documented in Figure 40.

6. Conclusion

By use of micromagnetic non-destructive techniques the ability to characterise materials ageing was demonstrated, in terms of hardness enhancement and Cu precipitation due to thermal degradation, in terms of supplying an early warning before fatigue life is elapsed due to Low Cycle Fatigue, and in terms of indicating the DBTT 41J shift when material degradation is due to neutron embrittlement. The demonstration was at well defined laboratory-type specimens but a high sensitivity and confidence in the results was obtained.

However, the next development step to perform is the demonstration of the techniques at real life components and the integration in inservice testing respectively in ageing management procedures of real plants. This will in addition include UT by using EMAT.

As the special application of EMAT sensors has demonstrated its reliable use at a service temperature of 300°C the integration of this sensor type into plant lifetime management systems should be an engineering problem which is to solve concerning the proper selection of cooling devices for the driving microelectronic systems and heat resistant wires for coils and cables especially isolated for high temperature access.

An EMAT itself is 'no more' than a magnet-inductive transducer of which the induced eddy current field is superimposed by a magnetic field. The last can be static as well as dynamic. When the superimposed magnetic field is changed dynamically in a hysteresis loop the EMAT can be applied to collect as well other micromagnetic quantities as a combination sensor for 3MA.

One strong critic comes always up when micromagnetic techniques are discussed for materials characterisation. This is the need of defined calibration specimens and the efforts for recalibration when transducers have to be adapted or are to replace after repair. Therefore the scientific community works hard to define and optimise robust calibration procedures to reduce or even to avoid the efforts and first success can be reported.

When material microstructure states are needed for reference to fix absolutely a time scale, for instance in fatigue life estimation, the procedure can be applied always when a component starts into new life, for instance after a replacement.

7. Acknowledgment

The author very much acknowledges the high valued contribution of his colleagues from Fraunhofer-IZFP which are Iris Altpeter, Klaus Szielasko, Madalina Rabung, Ralph Tschuncky, Gerhard Hübschen, and Karl Hällen. The special thank is to the Materials Testing Institute, MPA, at the University Stuttgart (Prof. E. Roos) and to the Institute of Material Science and Engineering, WKK, (Prof. D. Eifler) at the Technical University Kaiserslautern with their teams for the long year fruitful co-operation. Last but not least thank is to the ministry of economy and technology for the financial support in different projects beginning in 1979 up to now.

8. References

Allen, A.J. & Buttle, D.J. (1992). From Microstructural Assessment to Monitoring Component Performance - A Review Relating Different Non-Destructive Studies", *Nondestructive Characterisation of Materials V*, T. Kishi, T. Saito, C. Ruud, R. Green, Eds., Plenum Press, New York, pp. 9 – 30

Altpeter, I, Dobmann, G., Katerbau K.H., Schick, M., Binkele, P, Kizler, P., & Schmauder, S. (1999). Copper precipitates in the steel 15 NiCuMoNb 5 (WB 36): Material properties and microstructure, atomistic simulation, NDE by micromagnetic techniques, *Proceedings of the 25 MPA Seminar*, 7-8 October, Stuttgart

Altpeter I. (2002). Electromagnetic and Micro-Magnetic Non-Destructive Characterisation (NDC) for Material Mechanical Property Determination and Prediction in Steel Industry and in Lifetime Extension Strategies of NPP Steel Components, *Inverse Problems*,18, pp. 1907-1921

Altpeter, I., Dobmann, G., Kröning, M., & Rabung, M. (2009). Micro-Magnetic Evaluation of Micro Residual Stresses of the IInd and IIIrd Order, *NDT & E International*, 42, 4

Altpeter, I., K., Szielasko, K., Dobmann, G., Ruoff, H., & Willer, D. (2010), Influences of ageing processes on the fatigue life-time and toughness of the low alloyed steel WB 36, *Report No 090116-TW of the Fraunhofer-IZFP*

Altpeter, I., Tschuncky, R., Hällen, K., Dobmann, G., Boller, Ch., Smaga, M., Sorich, A., & Eifler, D. (2011). Early detection of damage in thermo-cyclically loaded austenitic materials, submitted for publication in the ENDE 2011 proceedings, *ENDE 2011 conference, March 10-12*, Chennai

Bassler, H.J. (1999). Cyclic deformation behaviour and plasticity-induced martensite formation of the austenitic stainless steel X6CrNiTi1810, PhD-thesis at the university, Kaiserslautern

Burkhart, G.L. & Kwun, H. (1989). Measurement of residual stresses around a circular patch weld using Barkhausen noise, *Review of Progress in Quantitative NDE, Vol. 8B*, D. O. Thompson, D. Chimenti, Eds. Plenum Press, New York, p. 2043

Buttle, D.J. & Hutchings, M.T. (1992). Residual stress measurements at NNDTC, *British Journal of NDT*, Vol. 34, No 4, 175

Cullity, B.D. (1972). *Introduction to magnetic materials* Addison- Wesley, London

Fillon, G., Lord, M., & Bussière, J.F. (1990). Coercivity Measurement from Analysis of the Tangential Magnetic Field, *Nondestructive Characterisation of Materials IV*, C. Ruud, J.F. Bussière, R. Green, Eds., Plenum Press, New York, pp. 223 - 230

Jansky, J., Andrä, T., & Albrecht, K. (1993). Feedwater piping guillotine breaks at 340°C operation temperature, *Transactions of the 12th Intern. Conf. on Structural Mechanics in Reactor Technology, ed. K.Kussmaul, North-Holland*, Vol F, pp. 207-214

Jiles, D.C. (1988). Review of magnetic methods for nondestructive evaluation, *NDT International*, Vol. 21, 311

Jiles, D.C. (1990). Microsttructure and stress dependence of the magnetic properties of steels, *Review of Progress in Quantitative NDE, Vol. 9*, D. O. Thompson, D. Chimenti, Eds. Plenum Press, New York, p. 1821

Kneller, E. (1966). *Ferromagnetismus*, Springer, Berlin

Koch R. & Höller, P. (1989). A modulus for the evaluation of the dynamic magnetostriction as a measured quantity for 3MA, *Nondestructive Characterisation of Materials III*, P.Höller, V. Hauk, G. Dobmann, C. Ruud, R. Green, Eds., Springer, Berlin, p. 644

Lang, M. (1999). Non-destructive characterisation of the cyclic deformation behaviour and plasticity-induced martensite formation of the austenitic stainless steel X6CrNiTi1810 by sensitive magnetic sensors, PhD-thesis at the Saarland University, Saarbrücken

Matzkanin, G.A., Beissner, R.E. & Teller, E.M. (1979). *The Barkhausen Effect and its Application to Nondestructive Evaluation, NTIAC report 79-2, pp 1-49*, Nondestructive Testing Information Analysis Center, San Antonio, Texas

McClure, J.C. & Schröder, K. (1976). The Barkhausen effect. *Critical Reviews in Solid State Sciences*, 6, 45

Pitsch, H. (1989). Die Entwicklung und Erprobung der Oberwellenanalyse der magnetischen Tangentialfeldstärke als neues Modul des 3MA-Ansatzes, PhD-Thesis, Saarland University, Saarbrücken

Rabung, M. (2004). Erarbeitung metallphysikalischer Grundlagen zur Anwendung der Mikromagnetik zum Nachweis der Werkstoffveränderungen infolge von Kupferausscheidungen. PhD Thesis, Saarland University, Saarbrücken

Sablik, M.J., Burkhart, G.L., Kwun, H., & Jiles, D.C. (1988). A model for the effect of stress on the low-frequency harmonic content of the magnetic induction in ferromagnetic materials. *J. Appl. Phys.* 63, 3930

Salzburger, H.J. (2009). H.J. EMATs and its Potential for Modern NDE - State of the Art and Latest Applications, *proceedings of the IEEE International Ultrasonics Symposium 1* , 621-628

Seeger, A. (1966), *Moderne Probleme der Metallphysik*, Springer, Berlin

Solomon, H.D. & De Lair, A.E. (2001). The influence of dynamic strain ageing on the low cycle fatigue behaviour of low alloyed and carbon steels in high temperature water, *General Electric Research and Development Centre*, Technical Information Series, CRD 134

Szielasko, K. (2000). Aufbau eines Modularen Messsystems auf Softwarebasis zur Zerstörungsfreien Charakterisierung des Versprödungszustandes von kupferhaltigen Stählen. Diplomarbeit. Hochschule für Technik und Wirtschaft des Saarlandes, Saarbrücken

Theiner W.A., & Waschkies E. (1984). Method for the non-destructive determination of material states by use of the Barkhausen-effect, Patent DE 2837733C2

Theiner, W.A., Altpeter, I., & Reimringer, B (1989), The 3MA-testing equipment, application possibilities and experiences, *Nondestructive Characterisation of Materials III*, P.Höller, V. Hauk, G. Dobmann, C. Ruud, R. Green, Eds., Springer, Berlin, p. 699

Yashan, A. (2008). Über die Wirbelstromprüfung und magnetische Streuflussprüfung mittels GMR-Sensoren, PhD-thesis at the Saarland University, Saarbrücken

Part 4

Plant Operation and Human Factors

18

Human Aspects of NPP Operator Teamwork

Márta Juhász and Juliánna Katalin Soós
Budapest University of Technology and Economics, Department of Ergonomics and Psychology, Budapest, Hungary

1. Introduction

The aim of this Chapter is to describe several important human aspects of NPP operator teams that have significant effect on safe and efficient operations. The first part of the Chapter provides an overview about the concept of high reliability organisations, safety culture, focusing on the question how to conceptualise and measure safety culture, presenting two distinct perspectives about background of human unsafe acts. Based on theoretical and empirical works made in high reliability organisations, the second part of the Chapter aims to detail the paradox of human factors, describing the main task, job and teamwork characteristics of first line personnel.

The first line personnel in the NPP control room works in team. There have been several attempts to describe the characteristics of efficient teamwork, although little is known about the antecedents of efficient teamwork in high reliability organisation. Based on the *Input-Process-Output* model the empirical works aim to understand those inputs and processes that determine safe and efficient operator teamwork.

After the theoretical considerations, the chapter synthesise different empirical works made in NPP control room analysing operator teamwork from different perspectives. Based on the theoretical works about specific task loads, the goal of Case study is to identify particular sources of task load, as *inputs* that influence operators well being. The revealed list of task load offers a practical guidance how to enhance operators' well-being, safe and efficient work performance. Another important input of operator teamwork is the team members' personality. Even though the NPP environment is strongly standardised, providing little room for individuals' personality, team members' characteristics influence how they behave and perform in this restricted environment. Based on Five Factor Model of personality the goal of Research 1 is to determine those personality traits that count for efficient teamwork, relating personality to team communication, to behavioural markers of team members, and performance. Operator team is a professional work team, requiring the interaction of team members representing different areas of speciality. It is important to understand how the operator team members having specific technical and professional knowledge are able to operate and manage jointly the plant system. *Communication as a key process* is used to share specific technical and human aspects of the plant parameters, operations, establishing the shared knowledge about the plant, environment, task and team members. This shared knowledge helps the operator team to develop joint strategies in order to manage the plant and to share different levels of task load during their operation. Research 2 aims to describe

characteristics of efficient team communication, to relate operator team communication to performance and to different levels of task load. The *output* of teamwork communication manifests in shared knowledge, the output of effective balance between demands created by high level of task load and operators resources is shown in the team well being. The most important output of teamwork is the team performance. In order to enhance team performance those input and process factors should be considered that determine efficient and effective performance. In sum, the present Chapter aims to provide theoretical background to understand the human factors of operator teamwork, and through the empirical works aims to reveal those factors that influence team performance.
All the presented empirical works were made in Hungarian NPP. The Hungarian NPP located in Paks, along river Danube, houses four nuclear reactor units, and covers more than 40% of national energy production.

2. Safe and efficient operations

Advances in technology have remarkably improved an organisation's ability to build and manage hazardous technologies. Although, hazardous technologies are not maintained for their own sake, but to improve the quality of life of human beings. Since the complexity of hazardous technologies has extensively spread, the balance between safe and efficient actions has been widely acknowledged. It is often claimed that the management should endeavour to find the right balance between safe and financially, economically efficient operations (Reason, 1997). This equilibrium becomes crucial in the case of high reliability organisations that can be described as organisations that are faced with high hazard situations, and for this reason they try to achieve and maintain high reliability, safe and efficient performance, while managing complex systems.
First in the 1960s, three organisations were considered as high reliability organisations: the US air traffic control system, organisations operating at nuclear power stations, and the US Navy nuclear aircraft carrier operations. The initial definitions were less precise saying that hazardous systems are organisations that should function almost without errors, accidents. This broad definition raises the question how to interpret "almost without" or "near error-accident free function". This unclear definition has been changed to a more precise interpretation, which instead of error, accident rate emphasises the effective management of inherently risky technologies or environment. Another expression, describing organisations in which there is more than normal chance for damage one's own life, the life of others or to material property is called high risk environments (Dietrich & Childress, 2004). The latter concept focuses on the inner characteristics of these organisations, high risk, hazards, while the former emphasizes the efficient management of high-hazard situations.
In the present work *high reliability organisation* and *high risk environment* concepts are used interchangeably to describe organisations such as a Nuclear Power Plant (NPP) where the idea of safety is not just a theoretical concept, but an eternal, conscious endeavour to maintain safety in the nature of the high hazard operation, environment.

2.1 Differences and similarities between high reliability organisations

Recent researches on high reliability organisations deal with domains in health, safety and environmental issues, including studies about the personnel of aircraft, air traffic control, nuclear power stations, operating rooms, medical team, intensive care units, fire service (Reason, 1997). Even though these organisations have some strong common characteristic in

their function, it is necessary to consider the differences between local features, such as the output of erroneous actions, types of damages stemmed from inefficient work, the degree of standardisation of action and of communication, the size and structure of team. In the case of a NPP the output of erroneous action is the most severe, demanding high number of fatalities, or worst case environmental catastrophe, while in medical field a committed error leads to relatively low number of fatalities. Despite these specific features, there is a list of joint characteristics of these organisations stemming from the duty of efficiently and reliably managing high reliability situations. One of the main characteristics of high-reliability organisations may be described by the eternal endeavour to collect, analyse information about errors, incidents, near misses, sources of potential accidents, such as mistakes, lapses, slips. The aim of *information collection* and analyses is to enhance the safety of the system. For example, in an NPP data and information gathered from reports about unsafe acts are used in order to improve the training procedures, or to refine rules, procedures that govern safe operations.

The process of establishing technical standards, norms, and procedures to implement guidelines for safe actions has led to a high level of standardisation. *Standardization* can be regarded as the key element in minimizing unsafe acts. The efficient and reliable management of high-hazard situations necessitates highly level of training, to have the fundamental professional knowledge about the function of the systems, about events, and about the correct actions. In NPP, the simulation centres provides the opportunity to establish, update and practice professional knowledge, at the same time to drill compliance with rules, procedures.

The obligation to manage efficiently and safely an innately hazard environment implies a *strong pressure* on the first line personnel to provide high-level performance under all possible circumstances. This high pressure is indispensable to maintain safe acts; however, it can have its own adverse effects. As a response to this strong pressure, the highly trained first line personnel have developed a strong sense of *invulnerability*. The recognition of effects of stress or the acknowledgement of vulnerability to error is an indispensable part of efficient stress and error management strategy. This does not imply that the organisations should decrease the pressure to high-level of performance, but rather to underline the importance of performing efficiently and reliably, without the pressure to cover up, or to hide the errors, and vulnerability. The personnel should be given more information about the effects of stress with guidance on how to manage stressful situations, such as the reallocation of human resources between team members.

Due to high-level standardisation the first line personnel need to work mainly under a low or moderate *level of task load*. The vast majority of operations in a NPP are highly automated, in this situation the personnel need to monitor, follow the processes and to react on the specific events. The personnel need to be aware of the eternal presence of external factors that threaten the safety and effectiveness of operations. This awareness of new unfamiliar event emergence causes continuous alertness, one of the main sources of high task load in a high- risk environment (Mumaw, 1994).

3. The paradox of human factors in NPP operations

3.1 Autonomy vs. control

High reliability organisations strive for minimizing hazardous, unsafe actions while improving operational efficiency by keeping maximum control and implementing strict

procedures that prescribe how to act, interact, and communicate under certain conditions (Gudela & Zala-Mező, 2004). On the other hand, the first line personnel is composed of individuals with high professional knowledge, highly skilled and experienced persons, with the strong need to manipulate and control their environment, striving for autonomy. *Autonomy* is a self determination regarding which goals, rule, and procedures to follow. Hackman and Oldham (1975, as cited in Byrne & Davis, 2006) defined job autonomy as the degree of freedom that employees have in order to schedule, choose and determine the method how to accomplish his/her tasks, responsibilities and work. *Control* is an influence to guide or regulate the activities or operation of a person, system or machine. In work-related settings, the constellation when a person has little autonomy and strong centrally determined control is called strong situation (Barrick & Mount, 1993). In high risk environment the management is committed to minimize uncertainty with strong centralized control; this in turn will not allow the local actors to be engaged in situations where they have the opportunity to use their professional knowledge. In strong situations, individual differences in knowledge, skills, abilities, experience, and personality has little opportunity to be manifested, because individual's actions are constrained by a variety of external factors including detailed rules, standard operation procedures, and supervision. In contrast, high degree of autonomy and low control situations, so called weak situations, permit or even elicit individual differences, where individual characteristics would play a major role in behaviour, actions to accomplish assigned tasks.

The paradox between the centralised control and the need for local autonomy could be dissolved by the implementation of right balance between centrally defined rules procedures and opportunities to use expert knowledge.

3.2 "The golden rule is that there are no golden rules" (George Bernard Shaw)

The need to reduce uncertainty is manifested by the need to foresee all the possible events, including normal situations and operations deviating from normal events. Uncertainty reduction is achieved by developing standardized procedures, rules for all the potential cases. In this sense standardization proves to be the key element of coping with uncertainty at organizational level. The paradox is caused by the fact that overregulated behaviour guided by strict rules may impede local actors to adapt to and manage situations when the rules are incomplete or inappropriate. During these uncertain events, especially during abnormal operations the appropriate action depends on *reasonable flexibility* between the use of procedures and professional knowledge (Gudela & Zala-Mező, 2004). In NPP industry 60% of human performance problems are associated with wrong or inapplicable rules, procedures (Reason, 1997). In sum, harmful effects of high standardisation could be demonstrated by the incapability of reaction to those uncertain situation which lack adequate rules, and by an over-reliance on inappropriate rules. The detrimental effects of high standardisation can be avoided by the use of deep professional knowledge and experience.

3.3 Certainty of uncertainty

Due to strong standardisation the work of first line personnel mainly consists of periods of low and moderate levels of workload. During a low workload period, the personnel need to monitor and detect different sources of information, or to scan and follow particular parameters. During a period of moderate workload, the personnel need to react on certain parameters, information from different sources of the control panel, based on

predefined rules, procedures. This period of the work is described by routine operations. During a period of low and moderate workload the aim is to keep the system in equilibrium, to sustain or to improve the function of the complex system (Waller et al., 2004). The task accomplishment is described by foreseeable actions and the steps, the decisions are mainly based on the application of prescribed rules, procedures. Even though, the "sharp end" personnel are exposed to the emergence of *unexpected events* (Mumaw, 1994). This constant alertness *evokes high workload on the personnel's capacity* to adapt from the routine to non-routine actions, where the appropriate reaction relies, not just on rules-based behaviour, but more on expert knowledge to manage the unfamiliar and uncertain problems. The occurrence of uncertain situations can hardly be predicted, thus the adaptability to dynamic task load, flexible change from predictable to unpredictable situations is considered to be the key characteristics of managing complex system functioning.

3.4 Centralisation vs. decentralisation

The management of complex technologies necessitates the implementation of a conventional hierarchy, where the *hierarchical structure of information and interaction flow*, chain of actions is kept under control by centralised persons. The effective handling of an unexpected, uncertain event, where the rules do not cover the situation or they are not appropriate for the tasks in hand, the reliance on decentralised actions of the professional experts can be the crucial aspect of reliable and efficient performance. In this way it could be stated that in efficient high reliable organisations a decentralised hierarchy should coexist with a centralised hierarchy, in the form of adaptability from a hierarchical centralisation structure to a decentralised professional mode. This flexibility is an important way of achieving common, joint awareness of the situation, and to share the high task load between the team members, increasing the teams' capacity to face high task demands (Reason, 1997).

3.5 Individual- vs. teamwork

Work teams consist of highly trained individuals with special expert knowledge. Each person is responsible for one specific area of the complex system. The intensive and constant pressure to perform efficiently and reliably, the high expectations from the organisation as well as from the society, stress out the importance of well established professional knowledge that guides the behaviour of these personnel under any circumstances. In this way the training of first line personnel is mainly focused on practicing professional technical knowledge, such as the management of the complex system. On the other hand, researches about safety functioning show that one of the major contributing factors of accidents, unsafe acts are not linked to the lack of professional knowledge related to technical aspects of the complex system, but rather to the *failures of efficient teamwork*, such as inappropriate *communication, coordination* (Flin et al., 2002). The duality lies in the fact that even though these operators are highly trained and strive for independence, the efficient system functioning depends on teamwork. The central point of this paradox is caused by the fact that the variety of the systems exceeds the variety of the persons who control them (Reason, 1997). In this way the nature of these systems requires cooperation between the first line personnel. Since the cooperation, communication and coordination within the team is efficient, the end of these processes could exceed the variability of individuals, which in turn could minimize the unbalance between the variability of the system and the variability of

the human controllers. All work teams can include team members who prefer to work alone, rather than to work in group, to share information, cooperate with other team members. Also demands of the tasks can influence the level of the necessity of cooperation in teamwork (Thompson 1967 as cited in Blyton et al., 1989; Gudela & Zala-Mező, 2004; Hellriegel & Slocum, 2007).

4. The relevance of teamwork in NPP

More and more organisations prefer to restructure the workflow around teams. This is particularly true in the case of high reliability organisations, where the actions are based on technically complex operations. The increased complexity of operations requires the knowledge, experience, skills and abilities of more than one person, as the management of complex problems is too demanding for one individual. Furthermore, technological developments have led to a high variety of the system. As mentioned above this system variety exceeds the variety of the human controller's characteristics. No wonder that professional work teams have started to play a crucial role in the management of complex operations, where the team members need to interact and integrate their individual capabilities to efficiently cope with the variety in the system they coordinate. Even though the education, training of individuals become more specialised, the problems that the teams need to face turns out to be more complex, requiring an inter- and multidisciplinary contribution. This line of reasoning underlines the demand of discipline specialisation, the justness of heterogeneous teams, where the team members posses specific roles knowledge, expertise (Ballard et al., 2008; Barry & Stewart, 1997; Cooke et al., 2001; Kiekel, & Cooke, 2004; Mathieu et al., 2000; Cannon-Bowers et al., 1993).

4.1 NPP control room operator teams

NPP control room is the central point for safe and reliable plant system coordination. Control room operators are responsible for maintaining safe and correct running of plant operations and optimising all of its parameters. The operators' primary task is to monitor important plant parameters and to coordinate the efficient functioning of the reactor and its support system. The operators also direct activities of the personnel in the outer fields of the plant (for example maintenance staff). For safe and reliable plant operations the synchronisation of different support systems is needed. The control room operator team requires the interaction of six members: Unit Shift Supervisor, Reactor Operator, Turbine Operator, Turbine Chief Mechanician, Unit Electrician, and Shift Leader. The professional supervisor of the operator team is the Engineer in Duty.

When taking into account the group's definition (Cooke & Gorman, 2006) saying that a group consists of two or more individuals assembled together for a special common purpose, operator teams can be considered as a group. Furthermore, the operators in the control room can be considered as a team, because they exceed the characteristics of the group, consisting of members with specific and varied roles and with some degree of interdependence among the members. A team is "a distinguishable set of two or more people who interact, dynamically, interdependently an adaptively toward a common and valued goal, each of whom has been assigned specific roles or functions to perform, and who has a limited life-span of membership" (Salas et al, 1992 as cited in Salas & Fiore, 2002). The assemble of the operators in the control room is considered to be a team, inasmuch as they all need to follow and achieve a common goal: to obtain and maintain the optimal plant parameters and to detect and react on

incidents deviating from the normal conditions. The output of efficient coordination manifests in the „syntonization" with the entire plant. A control room operator team is a heterogeneous professional work team. Heterogeneous because the roles, responsibilities, tasks within the team are distributed heterogeneously between the team members. Moreover, the operator team is distinguished from other types of teams falling into the category of work teams. Work teams are relatively stable in terms of time, team objectives, goals and team members, where the individuals have stable work roles (Cohen & Bailey, 1997). An operator team can be considered as a professional work team highly differentiated from other teams by means of the exclusive membership of experts, where the team members represent different areas of speciality. The teams execute operations under technologically complex conditions, which require extended training and preparation from the organization as well as from the individuals focused on the development of both professional and social skills (Dietrich & Childress, 2004; Cooke et al., 2000).

4.2 Team Input-Process-Output model

Despite the renaissance of teamwork, relatively little is known about how the individual contributes to the team processes and outcomes. The dominant way of thinking about the team is the Input-Process-Output (IPO) model. The model posits that a variety of inputs are combined to influence processes, which in turn affect team outputs. This model has a powerful influence on recent empirical research on team effectiveness, and on theories studying the influencing factors of team performance (Salas et al., 1992; Barrick et al., 1998; Essens et al., 2005). Hackman (1987) divided the inputs into three categories: 1) individual-level factors (team member attributes, personality, knowledge and skills), 2) team-level factors (structure and size) and 3) environmental-level factors (task characteristics, level of the autonomy). Intragroup processes refers to interactions that take place among the team members and include interaction patters such as conflict, efforts toward leadership and those communication patters that differentiate teams from each other. Each team has its own communication style depending on the environment they are working in. Team output refers to team outcomes associated with productivity, performance, as well as capability of team members to continue the work cooperatively.

Based on the IPO model we would like to present those inputs-process-output variables that could count as important factors in NPP.

Team inputs could be distinguished in three categories:

- Team members' characteristics, such as team members' knowledge, ability, skills, personality. In case of NPP operator teams, the emphasis is on the team members' professional knowledge, although the team members need to posses social skills and abilities for teamwork.
- Tasks characteristics: level of autonomy and control; level of task interdependence; different level of task load, task complexity, uncertainty. The control room operator team needs to face different levels of task load, necessitating a continuous behavioural adaptation from the team members.
- Organisational process: organisational culture employee selection, training, performance appraisal, reward system. In a NPP environment safety is the central key concept that appears in the practice of each organisational process.

Process variables include all the written and unspoken rules, norms, beliefs, and team processes such as communication, information exchange, coordination, cooperation, leadership, and stress management.

Output variables include the quantitative and qualitative aspects of team performance, effectiveness, efficiency, productivity, team members' satisfaction, well being, and commitment. The current and the future performance predict the capability whether the team continue to work together as a unit or not. The most important measure of team effectiveness is the current performance assessment of the team, which is based on either supervisor ratings of team productivity or objective indicators of team quantity and quality of productivity. Another critical measure of team effectiveness is the assessment of the team's capability to continue functioning as a unit.

5. Inputs of operator teamwork

5.1 Task load and workload

For those working in the control room of a NPP the level of task load continuously changes. In the work setting *task load* can be considered as an objective difficulty connected to the properties of a task and the term *workload*, is used when referring to how a situation is perceived by the people facing the task (Gudela et al., 2004). In our concept task load is considered as the sources of stressor that are related to the task environment and task fulfilment. Task load as a demand is inherent in the objective circumstances of the task and of the environment.

5.1.1 Sources of task load in a NPP

NPP operator teams need to work under various levels of task load, during normal operations there is always a chance for an unexpected, novel event that can have a strong effect on the team members' behaviour. High level of task load may be a consequence of sudden and unexpected high demands that disrupt the normal procedures of task accomplishment. Task environments are complex and often unpredictable; the causes of the events are sometimes unclear, even though they demand a quick and immediate response. The personnel must perform multiple tasks under high time pressure, noise, heat or other type of stressors. The consequences of poor performance are immediate and severe (Driskell et al, 2006).

Early stress approaches focus on the environment characteristics that have a direct impact on the operators. With the introduction of new technology the focus has shifted to task demands that operators are faced with. In the process of individual performance gradually increased the task complexity and the necessity of internal resource. In a NPP surrounding the following distinct categories as sources of task load can be distinguished:

- **Environmental factors**

The activity of control room operators and technical personnel are severely monitored by the management, the state, local government, the media and the citizens, causing a high pressure to perform without unsafe acts. Any slips, lapses, mistakes, delays, near misses or accidents should be analysed later and the operators, staff members are aware of their responsibilities in line with this. In a workplace such as a NPP the licensed operators need to be highly qualified persons who are requested to periodically renew their licences taking simulator-based requalification exams. Continuous professional trainings, exams can extensively load the personnel causing *performance anxiety*, which has been shown to be associated with performance decrements during trainings and exams.

The operator teams are exposed to facing several external and internal environmental task load factors which may have significant effects on their operations. Hockey (1986, as cited in Mumaw, 1994) classifies *external* environmental factors in two categories: *a)* The physical

environment includes high heat, poor lightening, protective clothing, noise, and vibration. The control room operators' concentrated activity may be disturbed by a high number of activated alarms. *b)* The social environment includes cooperation with each other, managing conflicts between operators, communication, increased demands of coordination with the personnel, and the requirement of keeping each other informed of event progress. The next three environmental stressors are considered as *internal* determinant of physiological state such as *c)* drug use such as caffeine, nicotine, depressants, alcohol. Under certain conditions some of these can facilitate task performance while others may impede it. *d)* Fatigue states caused by prolonged work, sleep deprivation or disruption and *e)* cyclical changes-regular, periodic changes in hormonal levels, alertness, body temperature due to sudden changes in work shift.

In general, in a NPP environment, control room personnel and support staff are well protected from the environmental changes mentioned above because the back-up system is well functioning and highly reliable.

- **Factors related to the characteristics of the task**

An occurrence of novel and uncertain event such as loss of critical information, failed implementation of a plan is a serious phenomenon that should be considered. *Novelty* refers to events that have not been experienced before and are perceived as a potential risk. *Uncertainty* generally refers to an inability to know how an event will progress or be resolved or the lack of exact information how to act properly. The role of technology as a source of stress such as unfriendly interfaces can increase information uncertainty. Novelty can be tied to uncertainty when a situation is novel, as there is no expectation about the outcomes. Both novelty and uncertainty are significant sources of task load for control room operators. To reduce the effects of these types of task load more information should be provided to the operators, which could make events more predictable, getting back the control over the event's outcomes.

In NPP settings the task demands are very high and the increased occurrence of unsafe acts is likely to occur due to greater requirements in task demands. Time pressure, increased monitoring of plant state and increased job complexity due to multiple task accomplishment are additional sources of task load contributing to higher level perceived stress, workload. Stress may be defined as a state of imbalance between environmental demands and the human's resources for dealing with the demands. The effects of *time pressure* impede the task performance in two ways. On one hand, under high time pressure people may perform the task more quickly at the expense of accuracy. On the other hand, performers may give an incomplete performance and the decision-making process can potentially produce significant errors. *Multiple task accomplishment* may have a negative effect when multiple sources of information need to be monitored or consulted simultaneously. Under these circumstances the shift in the focus of attention is needed for an effective task performance, although, this fast change, adaptation is impaired by the narrowed, focused attention. Complex multiple task environments strain the performer's cognitive resources. *Cognitive load* is provoked by stressful conditions where the performer's attention becomes more and more narrowly focused on cues of tasks and less sensitive to the more peripheral cues. Conditions described by a huge amount of information activate certain attention filters causing an increased selectivity of attention during perception of the tasks components. The filters serve as a protection from cognitive overload. A high amount of information process loads the working memory capacity that requires storing temporally relevant environmental cues, rules, procedures related to task accomplishment.

5.1.2 Reactions to task load

It is not enough to measure the task characteristics causing demands independently of individuals' ability because the difficulty in a stress situation is due to the degree of mismatch between task demands and human resources. A well-experienced operator possesses more abilities, skills, resources helping to cope with high demands in an overwhelming situation. For this reason, confronting with the same task a very experienced operator perceives less workload compared to an inexperienced one. Workload for individuals depends on the relationship between the cognitive resources of the individual and the demands of the situation. Experience is positively related to decision quality under high stress. Well established professional knowledge stimulates the person to analyse systematically the situation, to seek optimal solution, loading the cognitive resources (Fiedler, 1995). During high task load and under time pressure there is no room for systematic elaboration, in this way professional knowledge may impede a fast and efficient reaction. Professional knowledge by itself, without experience may impede the optimal contribution during high task load situations, due to the strong need to seek rational solutions which may not be available. Although, experience enriches the person with higher perceived control in the vast majority of the situations, and provide the feeling of comfort and stability during managing events. Experience enables the operators to react in an appropriate way without the need to think systematically.

All the stress theories emphasise the interaction between a person and the environment, looking at stress as a misfit between them. Cooper (1998) provides an approach to describe why one person seems to flourish while another suffers in the same situation. Individuals try to maintain *equilibrium* between environmental demands and their own resources. The person's physical and emotional state has a "range of stability" in other words "*comfort zone*" in which the individual feels stable, comfortably maintaining the control over the situation. The individual strives to cope with the external and internal sources of task load in order to restore the feeling of control and comfort. The balance between demands and resources should be kept by the persons' endeavour to mobilize his/her own resources. The level of stress depends on the individual perception of the mismatch that can be considered as workload.

Resources play an important role in the stress process. Skills, knowledge and ability are important resources to manage the task in hand and to cooperate with the team members. In a control room the operator team members can share the high level of workload by exchanging information via communication and asking each other to provide direct support. In order to support operators in keeping the balance between task demands and their resources, information should be provided to personnel about different sources of task load and potential limits and strength in resources, enlightening the personnel about the certain effects of task loads.

5.1.3 Consequences of task and workload

While there are some positive effects associated with high level of task load, such as the increased level of arousal, the high level of vigilance, wide range of cognitive skills may be affected at individual and team level, leading to various psychological, physical, behavioural problems.

What price do we pay for imbalance between resources and demands? Scientifics have identified the physical and behavioural symptoms of stress that affect individuals' well being. *Physical symptoms of stress include:* insomnia, constant tiredness, headaches, cramps and muscle

spasms, high blood pressure. *Behavioural symptoms of stress include* counterproductive behaviour such as absenteeism, aggressive behaviour, swear words, frequent drug use, smoking, loss of interest in other people, loss of sense of humour, difficulty in concentrating. *Psychological* problems: include constant irritability with people, feeling unable to cope with stress, lack of interest in life, feeling of ugliness. All these symptoms cause not only human suffering but they also imply economical costs.
In order to avoid the negative consequences of uncertainty, task load, there have been strong efforts to foresee as many non routine situations as possible deviating from normal operations and to develop standardized procedures. High level of standardisation has been developed in order to reduce the influence of individual differences in the perception of imbalance between demands and resources.

5.1.4 Case study

Some years ago in the Hungarian Nuclear Power Plant a new model for monitoring and assessing the psychological state of the front line employees was worked out. A new model, called **P**sychological **S**tate **A**ssessment (PSA) was developed in order to capture whether adaptation to the task load endangers employees' health and safe, effective work. The model is based on previously revealed sources of task load relevant in the work of first line personnel. The goal of the model is to provide guidance to assess employees' psychological state, and identify symptoms that could endanger safe work behaviour. The application of the model during several years promoted the establishment of preventive attitude in the organisation, providing counselling and training system, and various health promotion programs for employees.
First of all, job analysis was carried out to identify the sources of the main task load for the first line personnel. Based on these results a 41-item questionnaire was compiled and sent to 380 employees. 61% of the persons sent back the questionnaire so our sample consists of 231 workers' answers.
Analysing the fulfilled questionnaires by means of factor analyses the sources of task load were categorised in three groups: 1) Task, 2) Environment and 3) Organisation.

1. **Task**: *Complexity of job* (high amount of information to be provided and to be received, high level of attention and concentration, great amount of cooperation); Constant *alertness, readiness, decision* (decision making and working under time pressure, unexpected events, continuous alertness, responsibility for decisions consequences); *Work shift* (multi-shifts, overtime); *Continuous learning* (requalification exams, following technological developments).
2. **Environment**: *Working conditions* (working equipments and devices, the materials, the equipments, the protective outfits, hygienic conditions, changing room, restroom, dining room); *Physical environment* (climate control, noise, lighting, potentially dangerous circumstances).
3. **Organization:** *Organizational operation* (roles and responsibilities, over-regulated work process, information flow in the organisation); *Atmosphere at work* (work climate, work conflicts); *Organizational instability* (organizational changes influencing the work, employment uncertainty).

In the following, factors are summarised that decrease or increase employees' well being:
Factor influencing well being negatively: shift work, overregulatedness, responsibility and decision making, increased attention and concentration, work overload, time pressure, permanent learning, and exams.

Factor influencing well being positively: experience of success, problem-solving, good community and atmosphere at work, opportunity to develop, wider knowledge, interesting, various exercises, human relation, communication, professional challenges, which require creativity, "correct" salary.

5.2 Team members' personality

The job characteristics of the operator teams of a Nuclear Power Plant are complex and highly controlled in which there are considerable demands and pressures to behaviour conformity and a person is restricted in the range of his/her own behaviour. Thus, individual differences in personality characteristics are more likely to influence the specific behaviour a person adopts. This type of environment determines and regulates the team members' communication flow that consists of team and task-oriented utterances. The role of personality in team process and team performance is unarguable. All these circumstances lead our focus on analysing the relationship between the employees' communication and observable behaviour and their personality traits.

Personality is an important factor in accounting for how employees behave in teams and in the organisation. The interest in identifying personality predictors of job performance has led researchers to use the Five Factor Personality Model as an important conceptual framework. The development of the Five-Factor Model (FFM) is an important event in the history of personality psychology because provides taxonomy for measuring personality traits. It describes personality traits based on five basic dimensions (Costa & McCrae, 1992).

i. **Neuroticism (N):** The tendency to experience nervousness, tension, anxiety, emotional instability, hostility and sadness.
ii. **Extraversion (E):** An energetic approach to the external world, including sociability, assertiveness and positive emotionality.
iii. **Openness to experience (O):** Describes the breadth, depth, originality and complexity of an individual's mental and experiential life.
iv. **Agreeableness (A):** The quality of one's interpersonal interactions along a continuum from compassion and altruism to antagonism.
v. **Conscientiousness (C):** Persistence, organization, and motivation in goal-directed behaviours, and socially prescribed impulse control.

The predictive power of the model within the employment context has often been demonstrated (Barrick & Mount, 1991; Tett et al., 1991; Piedmont & Weinstein, 1994; Salgado, 2001; Gellatly & Irving, 2001). In a review of Moynihan (2004) three basic theoretical perspectives explain the nature of personality effects on team performance. *Universal* approach: certain traits always predict teamwork process and team performance. *Contingent* approach: certain traits predict team performance depending on the task and organisational culture. *Configurational* approach: the mix of traits within a team and the fit of individual members with each other predict team performance.

Universal approach: Conscientiousness (C) has been examined in team performance because it is a reliable predictor of individual and team performance in field and laboratory settings (Neuman & Wright, 1999; Lepine et al., 1997; Barry & Stewart, 1997; Waung & Brice, 1998). Conscientiousness has consistently been found to be positively related to task focus and team performance, but only when both the team level and the leaders' conscientiousness are high. But it seems that in creative tasks, for example, a brainstorming study found that when team members are allowed to discuss strategies, teams composed of highly conscientious people produce better-quality performance (in terms of feasibility), whereas teams composed of low-

conscientiousness members produce a greater quantity of potential solutions. Tasks that require creativity may moderate the relationship between team conscientiousness and task performance. Therefore, Conscientiousness may be broadly applicable across numerous types of tasks, but may not predict specific types of tasks that require a high degree of creativity. The level of Conscientiousness in a team influences team functioning and outcomes. High level of Conscientiousness facilitates cooperation and creates an atmosphere in which individual team members are willing to learn from each other resulting in satisfied team-mates. If the level of Conscientiousness is low, no one feels responsible for a task, and team members do not stick to agreements or decision. All this can cause intragroup conflicts, stress and thus dissatisfaction. Conscientiousness relates to satisfaction and learning if the team is autonomous. A high level of autonomy is necessary to make decisions concerning any kind of work issues increasingly intensive intra-team communication and the mutual adjustment of efforts. If the team members are conscientious, they actively participate in decision making, and there is an opportunity to learn. So by sharing work-related attitudes and cooperating with each other, teamwork improves, contributes to satisfaction (Molleman et al., 2004).

The trait of Extraversion (E) has been shown to have positive effects on individual job performance for jobs requiring a high degree of social interaction (Barrick & Mount, 1991; Mount & Barrick, 1995; Littlepage et al., 1995). Teams higher in mean levels of Extraversion receive higher supervisor ratings of team performance than teams low on Extraversion. Teams with more extraverted members tend to be more socially cohesive and more highly evaluated by their supervisors. The degree of variance of Extraversion has a curvilinear relationship to task performance suggesting that too many or too few extraverts in a team can be inefficient. In general, Extraversion appears to facilitate cohesive team process, but only at moderate levels.

Teams with high mean levels of Agreeableness (A) have higher team viability, because Agreeableness is characterized by the concern for the team over desires and interests. In teams of management students working on a case study analysis and presentation task, individuals high on Agreeableness were more likely to be rated as cooperative team members by their peers. Low levels of Agreeableness (high individualism) are associated with reduced individual effort or social loafing in teams. Individuals low on Agreeableness tended to be unresponsive to teammates and tended to focus on their own task performance (Wagner, 1995; Comer, 1995).

Neuroticism (N) has been identified as a detrimental variable for team performance, and productivity. Teams with negative affective tone (negative affectivity or neuroticism) experienced higher rates of absenteeism. In sum Neuroticism is negatively associated with cohesive team process and effective decision making.

Contingent approach: According to this perspective the optimal team performance depends on the nature of the work, task and the organizational culture. These situational variables have moderating effects on the relationship between personality and team process or performance. Some studies consider the role of moderators in the relationship between personality traits and job performance (Barrick & Mount, 1993; Gellatly & Irving, 2001; Bono & Vey, 2007). The most important moderator is the situation in which the job performance takes place. The level of *task autonomy* moderates the relationship between personality and job performance: personality-performance correlations are founded to be higher in highly autonomous work situation than in less autonomous work situations (Beaty et al., 2001). The Agreeableness and performance relation is positive when the autonomy is low. When the autonomy is low, high level of agreeableness can helps the team member to achieve a higher

level of performance, while in high autonomy situations agreeableness can impede a high level of performance. This result indicates that personality-contextual performance correlations vary across situations with different expectations for performance. Personality and contextual performance behaviour is most strongly correlated when there are only weak cues and less correlated when there are strong cues.

Configurational approach: Certain personality traits may interact with others to result in desirable, as well as undesirable workplace behaviours depending on the pattern and interactions of other traits. Studies on team composition attributes have highlighted the relationship between team composition characteristics and team outcomes, but the results are inconsistent. Most of researchers have found a positive relationship between the mean level of Conscientiousness in a team and performance (Barrick & Mount, 1991; Hogan & Ones, 1997). Using the supervisory rating as a reliable measurement of workplace behaviour and performance, the evaluations show that highly conscientiousness workers (C) being low in Agreeableness received lower ratings of job performance than highly conscientiousness workers being high on Agreeableness. Highly conscientious workers who lack interpersonal sensitivity may be ineffective, particularly in jobs requiring cooperative interchange with others (Witt et al., 2002; Barrick & Mount, 1993; Molleman, 2004). If all team members are highly conscientious, each member contributes to the team task, and this will lead to many opportunities of learning from each others, facilitating cooperation. However, if the level of Conscientiousness is low, no one will feel responsible for a task, and team members will not stick to agreements or decisions resulting an atmosphere in which members are blaming each other for social loafing. This will cause intragroup conflicts, stress, and thus dissatisfaction.

A team that consists of stable members (N) is more effective. Stable individuals are more confident and less insecure while collaborating with others, and therefore they will more easily bring in their own knowledge and opinions and be more receptive to the inputs of others. This will enhance the opportunities of learning and lead to a more relaxed atmosphere. As Barrick (1998) argued, teams with unstable people tend to demonstrate more anxiety and negative feelings, which lessen the satisfaction of the individual team members.

Individuals who are open to experience (O) will prefer tasks that demand creativity, and they will enjoy experimenting with new problem-solving strategies; hence, they will be motivated to learn. They will prefer work that challenges them to utilize and develop their cognitive abilities. Persons low in Openness to experience will easily bear a cognitive overload and avoid new and ambiguous situations that demand creativity and offer opportunities of learning (Molleman, 2004).

5.2.1 Research 1

Our research aim was to focus on NPP operator team members' personality traits and to relate personality traits to communication patterns, to behavioural markers of non-technical skills, and to teams' performance.

5.2.1.1 Methods

The data collection was based on 16 operator teams' (N=96) interactions analysis in the Simulator Centre of the Hungarian Nuclear Power Plant (NPP). The NPP Simulator Centre is a realistic, high-fidelity tool that is widely used in training and exams creating the required level of face-validity to be relevant for real life situations.

Each of the 16 operator teams had to follow the same scenario. In order to provide a complete picture of simulation the scenario "*Failure of one turbine unit*" will be described

briefly: according to the annual schedule used by instructors, a live Switchover Test needs to be performed, while an unjustified operation of the turbine protection occurs resulting in the failure of one turbine unit. The failure of the equipment is followed by the malfunction of the primary circuit pressure control, creating a condition that also needs to be managed. The mean duration of the scenario is about 35 minutes.

Video records of operators' activity during the selected scenario have been used for collecting and analyzing data. In order to keep the operators' real life behaviour at the beginning of the simulator study they were informed about video recordings during the ongoing training session, but they did not know exactly which of the programmed scenarios would be videotaped. Video recordings were made with the operators' joint consent.

The operator team consists of the following team members: 1) Unit Shift Supervisor (USS), 2) Reactor Operator (ROP), 3) Turbine Operator (TOP), 4) Turbine Chief Mechanician (TCH), 5) Unit Electrician (UE), and 6) Shift Leader (SL).

- **Personality measurement**

Each team member (N=96) was asked to fill in the NEO-PI-R personality questionnaire. The NEO-PI-R focuses on five major domains of personality, as well as the six traits or facets that define each domain (Costa & McCrae, 1992). (Table 1.)

Neuroticism **N**	Anxiety **NAN**; Angry hostility **NAH**; Depression **NDE**; Self consciousness **NSC**; Impulsiveness **NIM**; Vulnerability NVU.
Extroversion **E**	Warmth **EWA;** Gregariousness **EGR;** Assertiveness **EAS;** Activity **EAC;** Excitement seeking **EEX;** Positive emotions **EPE**
Openness to experience **O**	Fantasy **OFA;** Aesthetics **OAE;** Feeling **OFE;** Actions **OAC;** Ideas **OID;** Values **OVA.**
Agreeableness **A**	Trust **ATR;** Straightforwardness **AST;** Altruism **AAL;** Compliance **ACO;** Modesty **AMO;** Tender mindedness **ATM.**
Conscientiousness **C**	Competence **CCO;** Order **COR;** Dutifulness **CDU;** Striving for achievement **CAS;** Self discipline **CSD;** Deliberation **CDL.**

Table 1. NEO-PI-R factors and scales

- **Communication measurement: team-oriented utterances**

All the video recorded conversation during the selected scenario was transcribed word by word, identifying the operators' verbal utterances by two independent expert evaluators. Difficulties occurred in transcribing videotapes due to communication density during some periods of the interaction, much simultaneous conversation flow between members, additionally we were faced with a noisy control room environment. For all these reasons we have few blind points in the transcribed videotapes, where the speaker of some utterances cannot be identified properly.

Our aim was to capture some relevant *team and task-oriented communication utterances.* Research 1 focuses exclusively on team-oriented communication utterances that are likely to be related to team processes, on the team atmosphere stemmed from the individuals' personality. Team-oriented communication refers to the activities required to coordinate the workflow among team members. Task-oriented communication utterances and their analyses will be described in Research 2. During the task accomplishment specific team-oriented communication utterances were identified that were not strongly related to task accomplishment but rather to team process and interactions during the operation. **Communication utterances:** ***R****elation* (R) - Relation-related utterances, maintenance of contact, relationship, and vigilance in sentences ("*Hold the line please!*", naming the addressee). ***P****oliteness* (P) - The speaker gives a command,

information, question or affirmation formulated politely. The speaker determines the team atmosphere, and indicates the mutual respect among team members (*"Thank you", "Would you be so kind…", "Do it, please"*). ***M**otivation* (M) - Encouragement, formulated as reinforcement, completed with motivation, stimulation (*"It's perfect, just go on like this!"*). *First person plural (**We**)* - The speaker uses first person plural (*"We, our, us, let's"*). ***A**ffection* (A) - Words describing emotions, someone's emotional status, indicating astonishment, exasperation, frustration, excitement, relieve happiness or contentment (*"I regret it"*, "*I'm quite happy*" or laughing). ***T**hinking, cognitive* (T) - Words indicating cognitive process. It may suggest a problem-solving mechanism and can increase especially in facing with technical problems (*„I think…", „Attention!", „If… than…", "Check it!"*).

- **Team performance measurement**

The team performance was assessed by the instructors' impression about the teams' efficiency using a 3-point Likert scale (1: poor, 2: medium, 3: excellent) according to how fast and punctual they accomplished the task and in what degree they distorted from the optimal solution. 17% of the examined 16 teams were assessed as poor, 40 % as a medium and 35 % as an excellent performance teams.

- **Non-technical skills measurement**

Non-technical skills are defined as the cognitive, "hard" and social "soft" skills of team members (Flin et al., 2003). The cognitive so called "**hard**" **skills** are related to task-solving processes: *Professional knowledge* (appropriate knowledge about technology, equipment, environment, and ability to transfer and use this knowledge during operations); *Problem solving* (the skill to recognize and define the sources of task difficulties, and to be active in providing and implementing solutions); *Standard compliance* (following technical norms, rules, procedures, and stimulating other team members to comply with standards).

The social "**soft**" **skills** are team relevant skills: *Task load management* (efficient coping mechanism with unexpected and novel events and with difficulties during team processes); *Cooperation* (the ability to work effectively in team, to consider and support other team members' needs); *Communication* (the ability to exchange information briefly and clearly, acknowledging the received information). After each scenario accomplishment the instructors were asked to evaluate each nontechnical skill using a 4-point Likert scale (1: weak, 2: acceptable, 3: good, 4: excellent).

5.2.1.2 Results

- **Team-oriented communication utterances**

Analysing team-oriented communication utterances, the results reveal that the most frequently used communication utterances are *Thinking (T),* indicating the team members' cognitive, mental effort during the scenario. In the case of work teams, such as the operator team where the team's goal is mainly task-oriented, the frequent use of cognition related utterances is inevitable, although these elements of the communication contribute to the establishment and maintenance of team processes. The second most frequently used communication utterance is the *first person plural pronoun (We)* that indicates that the team members apply team perspectives in their point of view, emphasizing a high level of identification with the team. *Motivation,* as a communication utterance is relatively rarely used by the team.

Analysing the occurrence of communication utterances among different roles, the findings suggest that the Unit Shift Supervisor (USS) is the most active member in the communication process, often using team-oriented communication utterances such as Relation (R), the first person plural pronoun (We), Thinking (T). (Figure 1.)

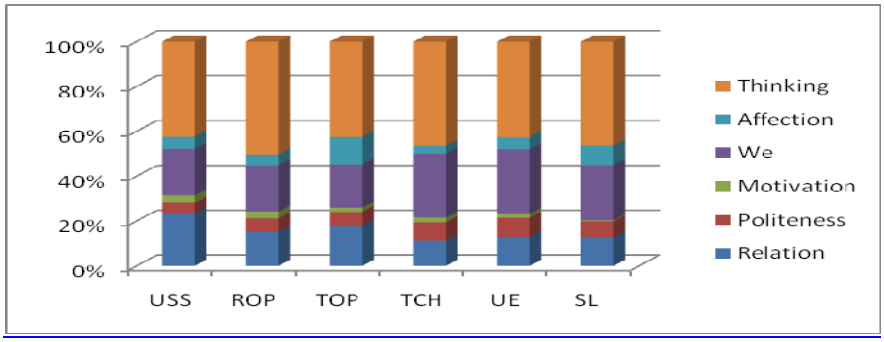


Fig. 1. Descriptive statistic of the non-technical-related communication utterances according to observed teams' roles

- **Relationship between team-oriented communication and personality**

The significant correlations between the frequency of different types of communication utterances and the NEO-PI-R factors and scales are presented. Correlation coefficients between personality and communication utterances organize around the Extroversion, Agreeableness, and Openness to experience personality factors and their scales. (*p<0,05; **p<0,00)

These analysed operator teams' communication refers to maintain relationship (*Relationship*) shows a significant correlation with Assertiveness (EAS) personality scale (,23**). The Extraversion personality factor and their scales such as Activity, Excitement seeking have significant correlations with *Politeness* communication style (E, EAC, EEX – Politeness: ,34**; 32**; 34**). The polite and acceptable communication style also has significant correlations with Openness personality factor and openness to fantasy and feeling scales (O, OFA, PFE – Politeness: ,26*; 32**; 32**) and Achievement striving scale (CAS – Politeness: ,27*). Behind a polite communication there is a positive and open personality, who is able to create an open and sincere relationship with other people and has the power to form acceptable team ambience in which everybody respects and tolerates each other without exaggeration.

To our surprise the Agreeableness personality factor and their scales indicate negative correlations with most of these team-oriented communication utterances (A, AAL, AMO, AST - Relation: -,40**; -,29*; -,40**; -,38**; A, AMO, AST, ACO – Politeness: -,31**;-,27*;-,40**;-,35**; AST – We: -,24*; A, AMO, AST, ACO – *Thinking*: -,31**; -,26*; -,29*; -,27*. It seems that the higher score on the Agreeable factor and its diverse scales, the lower is the possibility of using communication utterances related to maintaining interaction in this highly task-oriented team. For maintaining good relationship and a strong cohesion in these types of work teams for the team members it is important to be assertive (EAS) and it seems to be less agreeable (A) or compliant (ACO). An agreeable character is less fitting to teams operating in a high risk and strongly standardised environment. Highly modest (AMO), altruist (AAL), compliant (ACO) operators are less willing to initiate a new social action and easily become pressed by others in the team. Less agreeable people (A) more frequently apply expressions related to problem-solving procedures like 'think', 'attention', 'if…than' than those high score on Agreeableness.

- **Team performance and personality**

The results of regression analysis are presented in Table 2. As shown, the relevant personality traits are significantly related to team performance as a dependent variable:

Extraversion (E) and Conscientiousness (C). The standardized Beta Coefficients give a measure of the contribution of each variable to the model. ΔR^2 value tells that the Order scale (COR) model accounts for 9,8% of variance in the scores. Seeing that *t* value in this case is almost 3, it suggests that the Order scale as a predictor variable has a moderate impact on the criterion variable, on team performance. These findings underline and reinforce the relevant role of Conscientiousness (C) in the wok-setting performance.

	Team performance rating (as dependent, criterion variables)			
Personality factors and scales (as predictor variables)	**ΔR^2**	**ß**	**t**	**p**
Extraversion: Assertiveness (EAS)	.048*	.248	2,156	.035
Extraversion: Activity (EAC)	.050*	.252	2,190	.032
C_Conscientiousness	.071*	.290	2,552	.013
CCO_Competence	.050*	.252	2,195	.031
COR_Order	.098**	.332	2,966	.004
CAS_Achievement striving	.076*	.298	2,633	.010
CSD_Self discipline	.036*	.223	1,923	.058

Table 2. Regression results for testing Team performance and various personality factors and scales. Note: *p<0,05; **p<0,00 (one-tailed), for *t* values (for unstandardized regression coefficients) or *F* values (for overall model). *ß* = Standardized Coefficients.

Furthermore, it has also been analyzed how the homogeneity and heterogeneity of a certain personality factor alter team performance. The previously used Levene test rejects the homogeneity of variances, the Welsch D test on Agreeableness shows a significant main effect on standard deviation (SD) (d2=6,218; p<0,05). So, highly performing teams have a greater standard deviation of Agreeableness than poor or average performing teams. (Figure 2.)

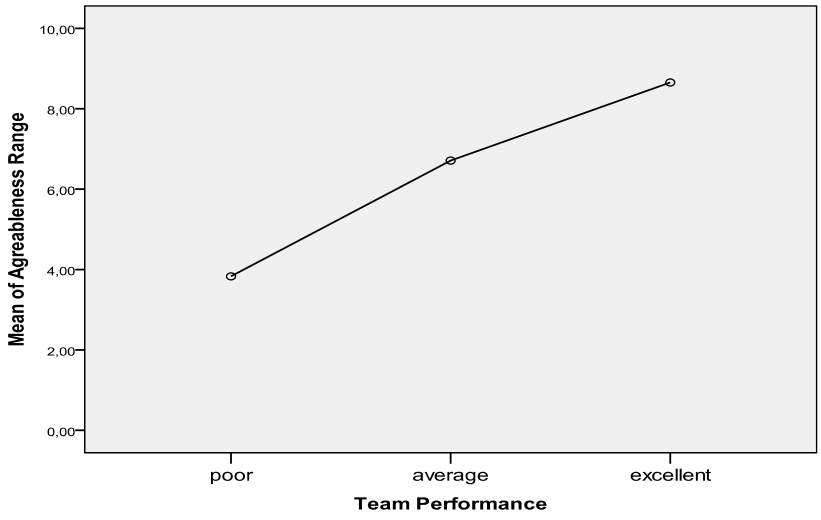


Fig. 2. Team performance and Standard Deviance of Agreeableness

- **Non-technical skills measurements and personality**

Using the stepwise linear regression analysis from all the predictors only the Anxiety (NAN) personality trait predicts significantly *Professional knowledge* as a dependent variable (*ß*=0,34; t=3,07; p<0,00) along the supervisor ratings. The NAN has been left in the model even if the dependent variable has been changed: *Comply with standard* (keeping rules) (*ß*=0,3; t=2,59; p<0,05), *Communication* (*ß*=0,38; t=3,38; p<0,00) or *Cooperation* (*ß*=0,37; t=3,24; p<0,00).

The other determinative personality trait that plays an important role in the instructors' judgment is Conscientiousness (C) factor, precisely Dutifulness (CDU) and the Order (COR) scales. When the *Comply with standard* factor has been evaluated the Dutifulness (CDU) factor has emerged from all personality traits (*ß*=0,251; t=2,16; p<0,05), and when the *Cooperation* factor is the dependent variable the Order (COR) personality trait (*ß*=0,31; t=2,78; p<0,00) influences mostly the instructors' rating. Furthermore, from the instructors' point of view an operator's *communication* skills mainly depend on his/her Assertively personality (EAS) type (*ß*=0,30; t=2,62; p<0,05).

Whilst the communication utterances have a strong relationship with the Extraversion, the Openness to experience and the Agreeableness factors, the "soft" and "hard" skills only show a significant correlation with the Anxiety scale (NAN) and the Order and Dutifulness (COR, CDU) scales.

Regarding our findings the Neuroticism factor associated with the Consciousness factor and their scales indicate their beneficial impact on the *Professional knowledge, Comply with standard* and on the interactive behaviour forms such as *Communication* and *Cooperation*. A moderate level of anxiety interacting with Conscientiousness can help persons to form a good impression about their own skills and behaviour. These persons endeavour to be accepted by others and strive to mark out from their environment with their remarkable performance. These people adapt to the changing environment in a very sensitive way.

5.2.2 Discussion

The extent to which the external environment constrains the individuals' personality varies in weak or strong situation. In strong situations, the organization exerts considerable pressure or demands to induce conformity. These controlling forces press the individual to behave in a specific way or exhibit a very narrow range of behaviours. Controversially, in weak situations the individual determines which behaviours to display, leaving bigger space for personality. In a NPP environment the individuals are placed in strong situations due to the high level of standardisation. In spite of this fact, personality has a key role in coping with these constrains.

The role of team members' personality in team communication has been analysed, in occurrence of observable non-technical skills and in team performance during specific task accomplishment. Our study reveals that team-oriented communication utterances highly correlate with Extroversion and Openness to experience personality traits, and to our surprise, in a negative direction with Agreeableness. Similar findings have been found in the Barrick & Mount (1993) study, in which the predictive validity of Agreeableness is investigated introducing autonomy as a moderator variable. The validity of Agreeableness is also higher in high-autonomy jobs compared with low-autonomy ones, but the correlation is negative. These findings suggest that the degree of the job autonomy influences the validity of personality dimensions. It means that in NPP operator teams when the members

work in a high autonomy, so-called strong situation the Agreeableness softly impedes the effective team functioning.
During the team-process the operators' "soft" and "hard" skills have a remarkable relationship with personality traits. First of all, Professional knowledge and Coordination behaviour markers show significant correlations with the Neuroticism and the Conscientiousness personality factors. The stable role of Conscientiousness has been reinforced, precisely Dutifulness and Order that mainly influence the operators' *Keeping rules* and *Cooperation* skills that largely determine their behaviour in this type of work settings. It seems that *Team-performance* as a team process output is directly influenced by the Conscientiousness and the Extraversion personality factors based on the instructors' evaluations.

6. Process of operator teamwork

The main question of studying teamwork in high risk environments is how the team members having specific knowledge, cognition and representing different fields are able to operate and manage a technically complex system. Cooke et al. (2004) emphasize that team cognition emerges from the interplay of the individual cognition of each team member and team process behaviour, thus team cognition is more than the sum of the individual team members' cognition. According to theoretical approaches of team cognition each individual has two different models: an *individual mental model*, which is long term knowledge (professional knowledge related to task, and team members) and an *individual situation model* describing a momentary, transient understanding of the current situation. In order to run a complex system it is needed to integrate the information and knowledge of the individual team members. The integration of long term knowledge, as well as the harmonisation of all the continuously changing environmental technical information may be attained through *team process* behaviour such as communication, coordination, cooperation and decision making, etc. The interaction of team members is remarkably important, since individual knowledge is transferred to team knowledge through these team processes. The output of this process will be two kinds of team level cognitive constructs: the *team mental model*, referring to the collective task and team knowledge (roles and responsibilities, knowledge of team mates, skills, abilities, beliefs), and the *team situation model*, describing collective team understanding of the specific situation. This team situation model guides the team in assessing and interpreting cues and patterns of the current situation (Cooke et al., 2000). In our view *shared knowledge* includes two of the above mentioned knowledge: team situation and mental model.
When analysing deeper the current literature of team cognition, two different complementary views of this construct can be found. The collective view of team cognition approaches this cognitive construct as aggregated individual knowledge. According to the other view the team knowledge may be assessed at a holistic level too, by focusing on the individuals' actions, and behaviour, not only on their knowledge. Team knowledge at a holistic level is the team members' knowledge that has been processed or integrated through team behaviours such as communication, coordination or cooperation (Cooke et al., 2004). On one hand, the collective view proved to be useful when knowledge is distributed homogenously among individuals, on the other hand, the holistic view is more appropriate when knowledge is distributed heterogeneously among team members (Kiekel & Cooke, 2004).

In spite of the fact that the individual knowledge is clear and accurate, the inefficient team processes (such as communication, coordination) may impede the integration of these knowledge structures, leading to inaccurate team knowledge, and inappropriate team action. This line of reasoning points out the importance of a holistic approach of team cognition. Thus, our view of team cognition describes this construct as the collection of individual situation and mental models, as well as those team processes that help the establishment and modification of team situation and mental models.
Team cognition guides the team in assessing the cues of situation, determining strategies, taking appropriate actions. Team performance will be maximized to the extent that team knowledge is accurate, appropriately apportioned among members, structured in a way that supports the development of effective strategies (Cooke et al., 2000). In turn, team performance may influence team process. An unsuccessful performance may urge the team to change their communication, coordination or decision strategies.
Team cognition is shaped by those team processes -such as interaction of the team members, communication - that helps integrate the team members' knowledge creating and continuously sustaining shared knowledge. In this way one of the critical aspects of team cognition is the team process that helps team members to create and share their individual knowledge.
Another question raised by researchers and practitioners is whether team knowledge exists, since team knowledge cannot be captured in one members' mind, brain. It exists within the context of team actions, interactions and within dynamic environment.

6.1 Communication as a crucial means to establish shared knowledge

Communication as a key team process is used in the team to share information, individual knowledge, to establish and to maintain current shared knowledge. Communication defines the way how team members execute complex tasks, and the way how a team handles and manages difficulties, and high task load situations.
There have been several attempts to help team communication in order to create and sustain shared knowledge under different circumstances. Waller et al. (2004) aimed to identify the adaptive communicative behaviours that help the NPP operator team to flexibly adapt to a dynamic task load environment. According to their studies adaptive behaviour such as information collection, task prioritization, and task distribution helps the team to create shared knowledge, which in turn helps the team to describe, explain, and make predictions and decide which action to be taken in a dynamically changing environment. It is also stated that information is collected and shared by the team members in order to identify tasks they need to perform, and receive, collect and screen information about these tasks. Appropriate information collection allows the team to better understand the situation, the system, which will help to build a shared conceptualization of the faced problems, leading to the effective establishment of team cognition (Waller et al., 2004). All these results suggest that teams attempting to collect more information will have an opportunity to gain, analyse, and understand the relevant cues from the environment resulting in higher level performance. While in low performing teams the members do not aim to acquire information reducing their ability to perceive the relevant environmental cues and act accordingly. Furthermore it has been also found that the use of long words is negatively related to performance and positively related to rates of errors. Similarly, studies claim that the use of more complex questions loaded the working memory, which in turn increased the risk of sending and receiving

erroneous messages (Sexton & Helmreich, 2000). Closed, yes/no questions are verifications, they are easy and quick to answer, in contrast with open questions ("what, why, how") that are incomplete and force the addressee to use the cognitive resources, to think and reflect. It has been found in the existing literature that the increases in communication volume, in particular communication about coordination (number of coordination requests), are inversely correlated with team performance, (Diedrich et al., 2005). However, it may be concluded that it is not just the communication quantity that affects team performance but also the characteristics of communication such as stability, focus, object of communication and timing. For the sake of an efficient information flow between team members it is also important to answer the question, to provide the information in timely manner. We tend to assume that team communication, has to be focused on the task itself trying to catch the relevant environmental cues from the present, and use this information to project future situations in accordance with the team's goals in order to facilitate the establishment of shared knowledge and performance. Furthermore, if the team's communication is consistently engaged in the past, they may fail to perceive and share relevant environmental cues from the present moment.

In the process of the formation of shared knowledge it is not sufficient to gather and to share the information, but it is also necessary to confirm the received information. It is not only the information collection behaviour that counts, but also the acknowledgement of the received information. Besides shared knowledge, the importance of its accuracy is also emphasized, since creating a shared cognition by itself does not lead to high performance only if the shared knowledge is accurate (Mathieu et al., 2000; Mohammed et al., 2000; Banks & Millward, 2007). In NPP operation, the tasks are allocated to several operators, and what is even more important is that each operator has a different information source. Communication is the only way of sharing information with each other, in this way it is crucial to clearly perceive the information and develop shared knowledge. One of the major characteristics of effective communication is verbal reaction, affirmation signing that the addressee perceived the information (Sträter & Fokuda, 2004). The lack of verbal feedback may suggest that the recipient overlooked the information (that may be relevant), in this way the speaker does not know whether the information has been perceived or not. At the same time the verbal reaffirmation of information may have some important side effects, the repetition of information may increase redundancy and what is even more important it strains the linguistic and cognitive resources of team members (Krifka, 2004). Individuals who expand their cognitive resources to speak more elaborately, to acknowledge the received information in detailed manner do so at the expense of decreased situational awareness (Sexton & Helmreich, 2000). Krifka (2004) advises "Make your contribution as informative as is required, BUT do not make your contribution more informative than required". The use of simple affirmation will help the team to clarify and acknowledge the received information, in this way to establish an accurate shared understanding of the situation. Conversely the affirmation with information will overload the cognitive resources of both the information provider and receiver, creating interference, impeding the team in creating a clear shared picture of the relevant aspect of situation. This criteria of efficient communication is in line with the Grice maxims, namely with the maxim of quantity prescribing that during efficient information transfer the speaker needs to give as much information as necessary but not more and with the maxim of manner describing the need to be brief and clear and avoiding long-winded information transfer (1957, as cited in Pléh, 1997).

The complete information flow between team members is particularly important in the joint establishment and fine tuning of shared knowledge. Coherent communication can be viewed as communication that responds to a previously initiated thought. These thoughts must be recognized, responded and new thoughts related to the previous one must be developed by the speakers, interlocutors. This goal can be achieved only if the members of the conversation are aware of each other's needs. The coherent conversation can be viewed as a continuum, as there is a strong semantic connection, relation between the parts of conversation, such as cause, condition, affirmation, and summary. In other words, the conversation is hierarchically structured, each part is semantically related to other parts (Krifka, 2004). Analyzing the coherence of conversation, Grommes (2007) states that coherence can be connected to mental processes. The operating room team members share broad common professional knowledge which constitutes the basis to be engaged in a coherent conversation. In turn, the coherent flow of information facilitates the creation of shared knowledge, common ground, which is essential for efficient joint activities (Grommes, 2007).

Communication is the most appropriate means of preparing for a coordinated action during routine operations and becomes more emphasised during non-routine situations, when the shared knowledge of the current situation is the key factor of efficient team actions. Shared knowledge can constitute the basis of an economical form of communication, namely implicit communication. Implicit communication is based on the knowledge of each others' personality, competencies, needs, task and responsibilities allowing voluntary task relevant information exchange, listening and offering assistance, unsolicited help. This form of communication allows team members to reduce the costs of explicit communication. Explicit communication includes information exchange as a response to a specific request verifying and acknowledging information, giving orders and assigning tasks (Swain & Mills, 2003; Gudela et al., 2004). During high task load the individuals' cognitive resources are overwhelmed with the management of a novel and complex situation, therefore it is important to save resources by means of implicit communication.

6.1.1 Research 2

The present paper aims to describe data from empirical researches about Nuclear Power Plant (NPP) operator team's communication, and its application to efficient teamwork. The research aims to analyse and describe team communication, to identify those specific communication dimensions that help to create shared knowledge, supporting the joint assessment of the current situation and developing adequate team strategies to face it.

The study focused on the NPP operator teams' communication, firstly in order to identify and understand those key communicative utterances that could be linked to higher team performance, secondly to identify how the teams adapt to high task load situations.

6.1.1.1 Methods

The data collection was based on the analysis of 16 operator team interactions in the Simulator Centre of the Hungarian Nuclear Power Plant. Since communication is the central factor of our research, the empirical studies of a "lively" interaction can best be carried out by analysing carefully chosen simulator sessions.

Each team had to follow the same scenario "Failure of one turbine unit" described under Research 1. Choosing the simulation, it was taken into consideration that the scenario had to

be oriented toward communication and in this way all team members had to be involved in solving the control task. Possessing complementary knowledge they had to share information with each other to manage the problems occurring during the simulated malfunctions.

As described under Research 1, Methods, video recordings of operators' activity during the selected scenario have been used for collecting and analyzing data. All the 16 teams' conversation has been transcribed in order to analyse team communication utterances.

- **Task load evaluation**

The scenario was divided into 3 phases by the instructors, according to the level of task load.

1. Phase of scenario *Moderate* level of task load: Executing a live switchover test
2. Phase of scenario *High* level of task load: Identifying, announcing and managing sudden, unjustified turbine operation
3. Phase of scenario *Moderate* level of task load: Indicating pressure control failure as well as the drop of one safety shutdown, requesting support service, resolving the situation.

- **Team performance measurement**

The performance scores were made by the instructors' evaluation, based on their impression about the teams' efficiency under the different phases of the scenario using the same 3-point Likert scale (1 – poor, 2 – medium, 3 - excellent).

By eliciting data from performance assessments four team performance categories were developed:

1. Excellent team: the whole team performance is evaluated excellent, through all the phases of the scenario (No. = 4 teams).
2. Average team: the team performance is medium continuously through all the phases of the scenario (No. = 5 teams).
3. Unbalanced team: the team performance varies from excellent to poor through the scenario (No. = 3 teams).
4. Poor team: the team performance is evaluated steadily poor through the complete scenario (No. = 4 teams).

- **Communication measurement: task-oriented utterances**

In Research 2 our aim was to capture some relevant task specific static and sequential analyses of the operator team's communication. Static measurements consider the team's communication only at a given point of time (e.g. every 10 seconds), or as an aggregate of the information flow over a period of time (e.g. during a complete task accomplishment). Sequential analyses take into consideration the ongoing stream of information exchange, interaction (Kiekel et al., 2001; Kiekel et al., 2002; Cooke & Gorman, 2006).

In order to capture the most relevant content static aspect of team communication *task-oriented communication utterances* have been developed expanding and specifying the communication dimensions used in similar environments (Conversation Analysis by Sacks, 1992; Speech Act Type-inventory for the Analyses of Cockpit Communication, STACK by Diegritz & Fürst, 1999; Krifka, 2004).

Task-oriented communication refers to the activities strongly linked to task strategies, task accomplishment, and it refers to the technical aspects of a task that must be performed.

Categories of communication (static measurements):

- *Information collecting question:* The aim of the question is information acquisition, for example asking about certain indicators or resources.
- *Open Question Information*: The question is addressed in order to complete the proposition with certain information, therefore, it is likely to receive a long answer (Questions that usually starts with words like *what, when, who*, etc.).
- *Closed Question Information*: The aim of this question is verification, to judge the truth of a position; therefore, the answer is either a single word (yes, or no) or a short phrase (*"Can we start the program*?").
- *Information Providing:* The team members inform each other about some relevant aspect of the mission related to human or technical indicators.
- *Information Providing Past*: The speaker informs the addressee about technological information, certain indicators that happened in the past, or about the crew's past status, personnel resources in the past (*"The error sign was caused by the failure of the pressure regulator."*).
- *Information Providing Present*: The speaker informs the addressee about some actual, present technological information, certain indicators, or about the crew's present status, personnel resources (*"I am preparing the necessary condition for the switchover test."*).
- *Information Providing Future*: The speaker informs the addressee about some technological information that may change in the future, foretells about certain indicators, or about his intentions and future actions (*"We will continue the test as soon as we have managed this unjustified reaction."*).
- *Affirmation:* It is the manifestation of two-way communications.
- *Simple Affirmation*: Answers to closed questions or commands, acknowledges the received information (Affirmations, acknowledgements, acceptances, answer: *'yes', 'no', 'ok', 'good'*).
- *Affirmation with Information*: A feedback, reinforcement on a status report or information, or command completed with additional information (*"Please switch on the 1 cb001 circuit breaker!" "1cb001 circuit breaker is on, we have the necessary differential current."*).

The team communication sequential analyses focused on the coherence of conversations. The anchored point of the *coherence analyses* was the new thought (that can be a question, information, etc.) initiated by one of the team members. The main condition of the coherent conversation is *turn-taking*, following this thought; the interlocutor develops a new question, information or command related to the previous information. Otherwise, if an initiated thought is not followed by any of the team members, it will be considered as a *thought without turn taking*.

6.1.1.2 Results

- **Open information question and performance**

Appropriate information collection and distribution allows the team to better understand the situation helping to build a shared conceptualization of the faced problems. According to our analyses, several specific communication utterances were related to performance. Particular forms of questions proved to be the best way to dispel uncertainties and to realize efficient communication. The results revealed that fewer open information collecting questions are used by the excellent performing teams than the lower performing teams ($F= 4{,}690$, $p<0{,}05$). (Figure 3.)

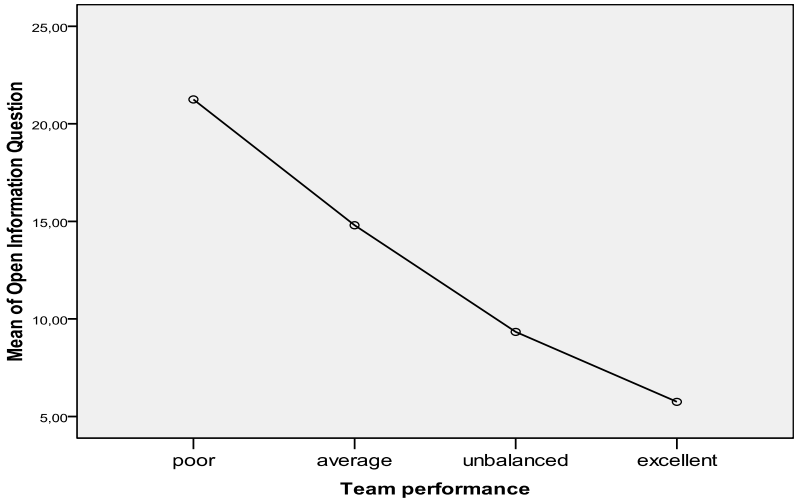


Fig. 3. The mean frequency of Open Information Question according to the team performance

The frequent use of open information collecting questions suggests that when lower performing teams formulate their questions, they have less information, knowledge about the environmental cues, so they formulate the questions in a less complete form. In turn, the excellent performing teams do not use open information questions so frequently, since they have more stable professional knowledge about the ongoing events being able to face the challenge of the situation. The open questions are incomplete and force the addressee to use the cognitive resources to complete the proposition. However, it is necessary to emphasize the usefulness of open questions in establishing team knowledge, but only during low task load, when the cognitive resources are not overloaded, so postulating an open question will not lead to any negative consequences. In this way effective communication that helps to establish team knowledge, and improve performance implies the ability of applying a simple and succinct vocabulary.

- **Affirmation and performance**

For the efficient information flow between team members it is also important to answer the formulated question, to acknowledge the received information. Although the differences are not statistically significant, the results can be regarded as a tendency that describes excellent performing teams using more simple affirmations and fewer affirmations with information, conversely with the low performing teams, where team members exchange more affirmations with information (Figure 4.). The result indicates the need for a clear information exchange that helps to establish accurate team knowledge, instead of creating an interference with additional, not so relevant, information. This result is line with Krifka's advice (2004) and Gricean quantity and manner maxim (1957, as cited in Pléh, 1997) to apply a simple brief vocabulary and to avoid providing information that is above the needed quantity.

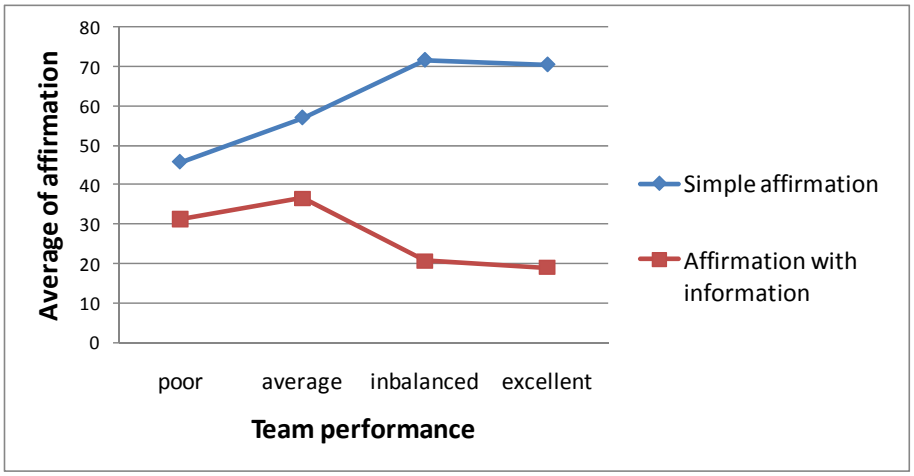


Fig. 4. The mean frequency of affirmation according to team performance

- **Information providing activity and performance**

Focusing closely on teams' information providing activity (Figure 5.) it is possible to describe the team's general tendency of focusing on the present, and less orientating about the past and future; at the same time there is a significant difference between the use of these communication utterances among excellent and poor performing teams. The poor teams' information flow contains more information about the past events ($p<0,05$), less information about the presently ongoing events ($p=0,005$) and about the future than the excellent performing teams' communication. The results suggest that excellent teams succeed to perceive the environmental elements in the present, to project the elements of the present status to the future, and focus less on the past.

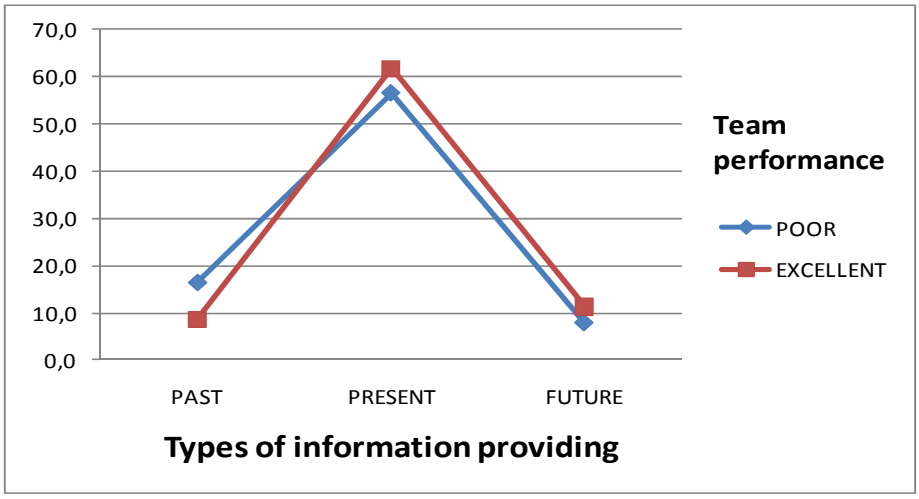


Fig. 5. The frequency of providing information about past, present, future according to team performance

- **Coherence of information flow and performance**

The coherent information flow between team members proved to be an efficient communication strategy to attain high performance. Comparing the coherence indicators of the excellent and the poor performing teams' conversations, the results show that the poor teams' conversations include more often thoughts without turn-taking ($t=5{,}506$, $p<0{,}05$) and fewer thoughts with turn taking ($t=4{,}069$, $p=0{,}05$) indicating an incomplete flow of information. Coherent communication means that the team members are aware of the information distributed by others, and react to the received information (either with a simple affirmation, or with a question, or providing additional information), creating a semantic connection in the information sharing activity. In this way coherent communication is one of the key elements of effective establishment or modification, fine tuning of accurate and complete team knowledge. The conceptual chain in the conversation helps the team to focus and maintain the attention on the exchange of information avoiding the loss of relevant information. (Figure 6.)

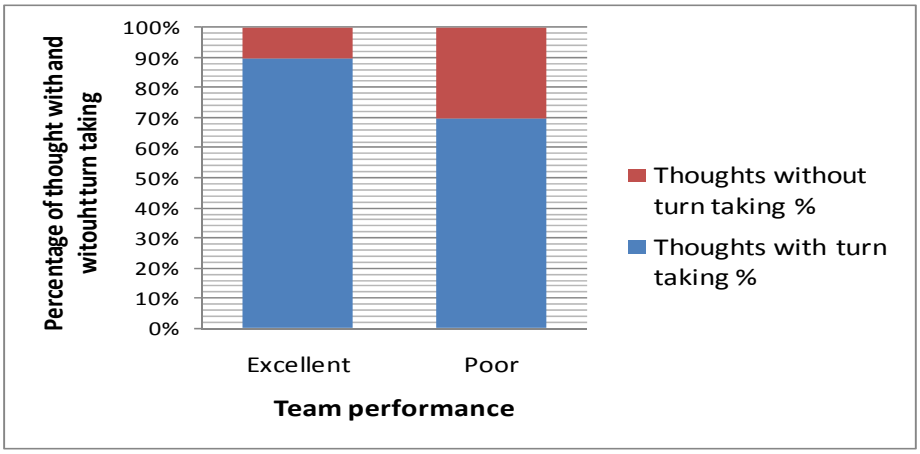


Fig. 6. The percentage of thoughts with and without turn-taking according to team performance

- **Communication under different level of task load**

Generally it can be concluded that as the task load increases, the frequency of communication utterances decreases. During high task load the communication is severely impeded, which can be explained by the operators' overloaded cognitive resources. The unexpected problems, failures intensively load the team members' cognitive capacity being unable to share their attention between the accomplishment of the task and communication. Furthermore, as the allocated resources disengage the collective need to process the causes and the consequences of unexpected event results in more frequent communication. (Figure 7)

6.1.2 Discussion

Some specific task-oriented communicative utterances prove to be crucial factors in the team processes that create and modify shared knowledge. Research 2 considers some specific aspects of communication that could be linked to establishing shared knowledge, such as using open information questions, affirmations, information providing activity, and

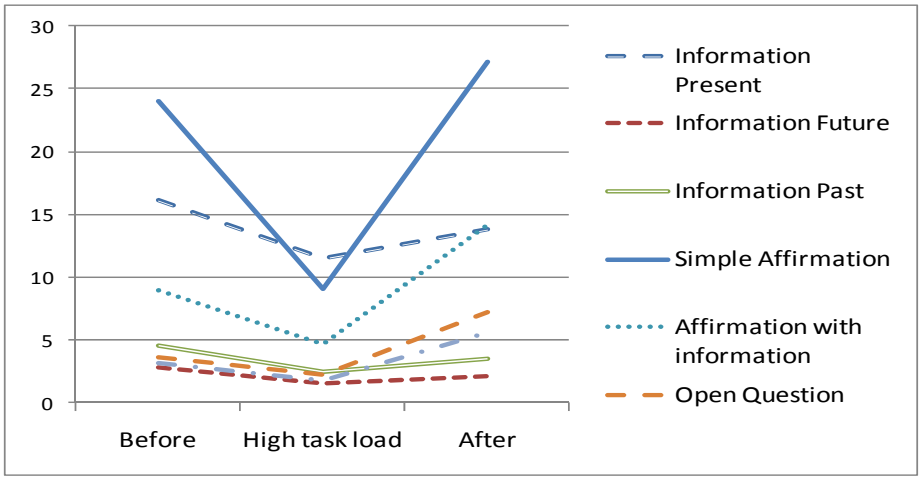


Fig. 7. The mean frequency of task-oriented communication utterances according to different levels of task load

coherence of information flow in a NPP environment. The use of effectively formulated information collection questions, the development of a well established effective communication strategy that focuses on the ongoing events and projecting the environmental cues to the future, affirming the received information could all help the team to build, and modify accurate shared knowledge and to improve team performance. The coherent information flow between team members proves to be an efficient communication strategy to attain high performance. Coherent communication means that the team members are aware of the information distributed by others, and react to the received information (either by means of a simple affirmation, or by means of a question, or providing additional information), creating a semantic connection in the information sharing activity.

Finally, at the beginning of the scenario, moderate task load necessitates less mental effort leaving more room for communication, although during simulation scenarios operators have some anticipation regarding unexpected events causing high task load (Antalovits & Izsó, 2003, Izsó, 2001). Under high task load the personnel's and the operators' cognitive resources are overloaded paying attention to the occurring problems, failures, and are unable to share their cognitive resources between problem management and communication. The communication density increases significantly in the third phase, under moderate task load as the cognitive resources disengage a collective need of elaborating the sources and consequences of high task load.

7. Output of operator teamwork

Taking into account that team performance is a multidimensional construct, measurements of team performance should focus not just on team efficiency but also on team effectiveness, because teams can be effective but inefficient at the same time. *Team effectiveness* measures the team output focusing on whether the team has reached all the specified goals, produced the intended results or not. While *team efficiency* refers to the way, how the team has reached the intended goals. The efficiency measures if the team was able to accomplish all of the

necessary tasks for a job in the appropriate amount of time, with the appropriate level of efforts or not. The aim is not just to reach the specified goals, but also get the maximum output for a long term, without consuming all the accessible resources. Effectiveness focuses on the end of the activity, efficiency focuses on the process or "means", if the team has reached the desired goals in an economical manner (Robbins, 1998 as cited in Horrowitz, 2005).

Measuring team performance two types of indicators should be considered:-*quantitative* and *qualitative* aspects of performance. In NPP control room the quantitative aspects of operator team performance includes incident, error, and accident data obtained from official reports, results of requalification exams. Qualitative aspects of performance consists of peer and supervisor ratings (Shift Leader) evaluating specific professional knowledge as well as cognitive and social skills required for safe and efficient plant functioning.

In order to improve team performance, the measurements and development efforts should focus near the outputs, on processes as well. Firstly, because the team performance output could be often influenced by external factors, on which the team has little impact, in this way, focusing on the output may be incomplete and misleading for developmental guidelines. Moreover, team improvement recommendations based only on outputs may lack some substantial information, since the output measures do not specify what aspects of performance are insufficient, deficient. Thirdly, the exclusive focus on team output may be misleading for team training recommendations, because in practice we can easily meet cases when team outputs has reached the specified level, even though the processes were impaired. In this case if we base our team training principles on these scenarios we could erroneously reinforce some defective processes (Cannon-Bowers & Salas, 2000).

Processes are regarded as a window to those strategies, knowledge, skills application that are used to accomplish specified goals, aims. Team processes refers to the application of team members' cognitive and social skills during task accomplishment. In NPP environment simulation training programs aim to develop and maintain these cognitive and social skills. During the program instructors evaluate and provide continuous feedback about task and team-oriented skills. Task-oriented skills are described as skills strongly required for task accomplishment such as decision making, problem solving, situation awareness. Task-oriented skills and knowledge are not sufficient when accomplishing tasks in a teamwork setting. Interpersonal, self-management skills and knowledge are regarded to be essential for performing well in teamwork setting. Team-oriented skills contains skills oriented toward team interaction, information exchange, those processes that help to maintain team as a whole unit, including cooperation, coordination, and communication between team members. The goal of training programs - besides developing professional, technical knowledge - is to reinforce efficient team interaction processes, to work better as a team. As a result shared knowledge will be formed, shaped about team including each other's characteristics, roles, needs, and about task referring to equipment, environment, rules and procedures. This shared knowledge serves as crucial aspect of the adjustment to novel and complex, dynamic task environment.

In recent years, the NPP industry has recognized besides the technical training the importance of team and task-oriented skill training, namely team interaction training. The concept of team interaction training stemmed from Crew Resource Management (CRM) training that has been developed as a special training in order to reduce error and increase the effectiveness of flight crew. This program focus on team and task-oriented skills critical for operational performance such as leadership, coordination, situation awareness, decision-

making, teamwork and communication. CRM in NPP field consists of the following topics: operational conduct, health at work, decision taking, situation awareness, choosing behaviour, feedback, communication, and team skills (Flin et al., 2002).
Training programs attempts to achieve change and development in different level: team members' attitude, skills, behaviour, knowledge, and finally to improve safety functioning at organizational level. To measure training efficiency accurately is a difficult task. Objective, direct measurements such as accident rates are incomplete, not always reliable indicator of training efficiency, especially as the high reliability organizations tend to have low accident rates. Subjective, indirect measurements including expert evaluators determine whether there has been any improvement in operators' knowledge, skill, behaviour, and attitude. For this evaluation experienced and trained observers use specially developed behavioral marker system such as in aviation Line/LOSA checklist (Line Operations Safety Audit, Klinect et al., 2003), NOTECHS (NOn-TECHnical Skills category, Flin et al., 1998), in medical field the behavioural marker system based on ANTS (Anaesthetists' Non-Technical Skills, Fletcher et al., 2003), NOTSS (Non-Technical Skills for Surgeons, Yule et al., 2006).

8. Conclusion

To conclude the results of the case studied and researches presented above, specific *sources of task load* were identified *related to the task* such as complexity of the task, constant alertness, continuous learning, and *related to the environment*, such as working conditions, *related to the organisation*, such as the amount of rules, responsibilities, the overregulated work process.
Analysing team members' personality based on Five Factor models and identifying the relationship between individual traits and performance, Extraversion and Conscientiousness were proved to be important characteristics influencing positively team members' behaviour and performance, while Agreeableness had a negative relationship with behaviour and performance.
Separating characteristics of team communication in two categories, the frequent use of *team-oriented communication* utterances were linked to Extroversion positively and Agreeableness negatively. Low level of Agreeableness could be associated with individualistic characters, attributes of experts striving for independence. Studying *task-oriented communication utterances*, features of the well established communication strategy has been described. Information gathering questions formulated in open way, affirmation also proved to be efficient way of communication, but what is more important to apply brief and clear vocabulary. The information flow between team members should focus more on the ongoing events and projecting the information to the future and less about the past. Coherent information flow between team members proves to be an efficient strategy for establishing and updating shared knowledge, achieving high performance. Describing team communication under different levels of task load, the results how that team adapts to high task load with increased communication, paying more attention to the management of the occurred failures.
The results of the researches and case studies can be used as directions for human resource management practices, especially in employee selection and development procedures, health promotion program. The revealed sources of task load and features of workload could help to develop and design health promotion programs. Operator teams' selection methodology should take into consideration personality trait patterns of operator teams, as well as the key competencies. Our results may have some important applications in

developing training interventions based on well established competency list, providing greater emphasize on communication. Feedback about team communication should focus on specific aspects of team communication that help to establish and modify accurate shared knowledge and to improve team performance. Instructors and operators are all responsible for developing efficient communication strategy. Although, it is difficult to generalize the results about team communication to all operator teamwork, as long as the presented researches are based on the analyses of communication in simulation environment following a particular scenario.

The presented researches emphasize one of the team processes, namely communication. Future research should go beyond communication, studying other team processes, such as coordination, cooperation, decision making. Furthermore future works could also reveal hidden complex patterns of team processes, in communication, in cooperation related to safe and efficient team performance.

9. Acknowledgements

The authors wish to thank the Management of Paks NPP for permission to carry out the simulator studies. We thank to István Kiss, Head of Training Division of Paks NPP, furthermore to the control-room teams and instructors of Simulator Centre, for the whole simulator staff for their cooperation in the research. We also wish to thank the help and support of professors Izsó Lajos and Antalovits Miklós from Budapest University of Technology and Economics, Department of Ergonomics and Psychology.

10. References

Antalovits, M. & Izsó, L. (2003). Assessment of Crew Performance and Measurement of Mental Task Effort in Cognitively Demanding Task Environment. In: G.R.J. Hockey (ed.): *Operator Functional State*, IOS Press and Kluwer Academic Publishers, UK.

Ballard, D.; Tschan, F. & Waller, M. (2008). All in the Timing Considering Time at Multiple Stages of Group Research. *Small Group Research*, Vol. 39 No. 3 pp. 328-351 ISSN 10464964

Banks, A. & Millward, L. (2007). Differentiating knowledge in teams: The effect of shared declarative and procedural knowledge on performance. *Group Dynamics: Theory, Research & Practice, 11,* pp. 95-106, ISSN 10892699

Barry, B. & Stewart, G. (1997). Composition, process, and performance in self-managed groups: The role of personality. *Journal of Applied Psychology*, Vol. 82, pp. 62–78 ISSN 0021-9010

Barrick, M. & Mount, M. (1993). Autonomy as a Moderator of the Relationships Between the Big Five Personality Dimensions and Job Performance. *Journal of Applied Psychology,* Vol. 78. No. 1. pp. 111-118, ISSN 0021-9010

Barrick, M. & Mount, M. (1991). The Big Five Personality Dimensions and Job Performance: A Meta Analysis. *Personnel Psychology*, Vol. 44. pp. 1-26, ISSN 0031-5826

Barrick, M.; Stewart, G.; Neubert, M. & Mount, M. (1998). Relating Member Ability and Personality to Work-Team Processes and Team Effectiveness. *Journal of Applied Psychology*. Vol. 83. No. 3, pp. 377-391 ISSN 0021-9010

Barry, B & Stewart, G. (1997). Composition, Process, and Performance in Self-Managed Groups: The Role of Personality. *Journal of Applied Psychology*, Vol. 82, pp. 62–78, ISSN 0021-9010

Bono, J. & Vey, M. (2007). Personality and Emotional Performance: Extraversion, Neuroticism, and Self-Monitoring. *Journal of Occupational Health Psychology*. Vol. 12, No. 2. pp. 177-192, ISSN 1076-8998

Beaty, J.; Cleveland, J. & Murphy, K. (2001). The Relation Between Personality and Contextual Performance in „Strong" Versus „Weak" Situation. *Human Performance*, Vol. 14 No.2., pp. 125-148, ISSN 0895-9285

Byrne M. & Davis E. (2006). Task structure and Postcompletion Error in the Execution of a Routine Procedure, *Human Factors* Vol 48., No.4, pp. 627-638, ISSN 0301-7397

Blyton P. et al. (1989). Time, work, and organisation, Routlege, ISBN 0-415-00418-7, London, New York.

Cannon-Bowers, J.;, Salas, E. & Converse, S. (1993). Shared Mental Models in Expert Team Decision Making. In Castellan, N. (Ed.), *Individual and Group Decision Making: Current Issues*. pp. 221-246 Hillsdale, ISBN 0-8058-1090-0, NJ: LEA.

Cannon-Bowers J. & Salas E (2000). *Making Decision Under Stress, Implications for Individual and Team Training,* American Psychological Association, ISBN 978-1-55798-767-9 Washington

Cohen, S. & Bailey, D. (1997). What Makes Team Work: Group Effectiveness Research from the Shop Floor to the Executive Suite, *Journal of Management*, Vol. 2 No.3, pp. 230-290, ISSN 0149-2063.

Cooke, N.; Salas, E.; Cannon-Bowers, J. & Stout, R. (2000). Measuring Team Knowledge. *Human Factors*, Vol. 42, pp.151-173, ISSN 0301-7397

Cooke, N. Stout, R. & Salas, E. (2001). A Knowledge Elicitation Approach to the Measurement of Team Situation Awareness. In McNeese, M.; Endsley, M. & Salas, E. (Eds.), *New Trends in Cooperative Activities: System Dynamics in Complex Settings*, pp. 114-139, Human Factors. Westport, Conn.: Praeger Security International, ISSN 0301-7397 Santa Monica, CA

Cooke, N.; Gorman, J. (2006). Assessment of Team Cognition. In P. Karwowski (Ed.), 2nd *Edition- International Encyclopedia of Ergonomics and Human Factors*, pp. 270-275. Taylor & Francis Ltd., ISBN 041530430X UK

Cooke, N.; Salas, E.; Kiekel, P. & Bell, B. (2004). Advances in Measuring Team Cognition. IN Salas, E. & Fiore, S. (Eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance*, pp. 83-106, American Psychological Association, ISBN 978-1-59147-103-5, Washington, DC

Cooper. C. (1998). Theories of Organisational Stress. Oxford University Press., ISBN: 9780198297055

Comer, D. (1995). A Model of Social Loafing in Real Work Groups. Human Relations, Vol 48 No. 6, pp. 647-667, ISSN 0018-7267

Costa, P. & McCrae, R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): *Professional Manual*. Odessa, FL: Psychological Assessment Resources

Diedrich, F.; Freeman, J.; Entin, E. & Macmillan, J. (2005). Modeling, Measuring, and Improving Cognition at the Team Level. IN Schmorrow D. (2005). *Proceedings of Augmented Cognition, proceedings; International Conference on Augmented Cognition (1st) and on Human-Computer Interaction* Vol. 11, Las Vegas, NV.

Dietrich, R. & Childress, T. (2004). Group interaction under Threat and High Workload IN Dietrich, R. & Childress, T. (2004). *Group Interaction in High Risk Environments*. The GIHRE Project. Ashgate, ISBN 978-0754640110, Burlington

Driskell, J.; Salas, E. & Johnstone, J. (2006). Decision Making and Performance Under Stress. IN Britt, T.; Adler , A.; Castro, C. Military life. The Psychology of Serving in Peace and Combat. Vol I. Military Performance, pp 128-54, Praeger Security International, ISBN 978-0-275-98300-0, Westport, CT

Diegritz, T. & Fürst, C. (1999). Empirische Sprechhandlungsforschung. Ansätze zur Analyse and Typisierung authentischer Äußerungen. Erlangen: Univ.Bilbiothek.

Essens P.; Ad Vogelaar; Jacques M.; Carol B.; Carol P.; Dr. Stanley H.; Joe B. (2005). Military Command Team Effectiveness: Model and Instrument for Assessment and Improvement, chapter 4: Review Of Team Effectiveness Models, NATO Research and Technology Organisation, *The RTO Human Factors and Medicine Panel* (HFM-087/RTG-023) Project on "Team Effectiveness" April 2005.

Fiedler, F. (1995). Cognitive Resources and Leadership Performance. *Applied Psychology* An international review, Vol. 44. pp. 5-28, ISSN 0269-994X

Flin, R., Goeters, K.; Hörmann, M. &Martin, L. (1998). A generic structure of Non-Technical skills for Training and Assessment, *23rd Conference of the European Association for Aviation Psychology*, Vienna, 1418 September 1998.

Flin, R.; O'Connor , P. & Mearns, K. (2002). Crew Resource Management: Improving Safety in High Reliability Industries *Team Performance Management,* Vol. 8, pp. 68-78, ISSN 1352-7592

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N. and Patey, R. (2003). Anaesthetists' Non-Technical Skills (ANTS): Evaluation of a Behavioural Marker System, British *Journal of Anaesthesia*, Vol. 90, No. 5 pp. 580-588, ISSN 0007-0912

Gellatly, I.; Irving, P.(2001). Personality, Autonomy, and Contextual Performance of Managers. *Human Performance*,Vol. 14 No.3, pp. 231-245, ISSN 0895-9285

Grommes P. (2007). Contributing to Coherence. An Empirical Study of OR Team Communication. In: Minnick-Fox, M.; Williams, A. & Kaser, E. *Proceedings of the 24th Penn Linguistics Colloquium*. Univ Penn Working Papers Linguistics 7:1, pp. 87–98.

Gudela, G.; Zala-Mező, E. & Grommes, P. (2004). The Effects of Different Forms of Coordination on Coping with Workload. In: Dietrich, R., Childress, T. (2004). *Group Interaction in High Risk Environments*. The GIHRE Project. Ashgate, ISBN 978-0754640110, Burlington

Gudela, G. & Zala Mező, E. (2004). Group Interaction in High Risk Environment of the Daimler-benz-Foundation. *Report on the Psychological Part of the Project*. Swiss Federal Institute of Technology Zurich

Hackman, J.R. (1987). The design of work teams. In Lorsch, J. (Ed) *Handbook of organizational behaviour*.pp. 315-342, Englewood Cliffs, ISBN 978-0133806502 NJ: Prentice-Hall

Hellriegel D. & Slocum W. (2007). *Organisational Behaviour*, Cengage Learning, ISBN 978-0324156843,

Hogan, J. & Ones, D. (1997). Conscientiousness and Integrity at Work. In: Wiggins, J. *The Five-Factor Model of Personality*. The Guilford Press. ISBN 978-1572300682, New York

Hopkins A. (2007). The Problem of Defining High Reliability Organisations, Working Paper 51, *National Research Centre for OHS Regulation*, The Australian National University, January 2007

Horwitz S. (2005). The Compositional Impact of Team Diversity on Performance: Theoretical Considerations, *Human Resource Development Review*, Vol. 4, pp. 219, ISSN 1534-4843

Izsó L. (2001). Developing evaluation methodologies for Human-Computer Interaction. Delft University Press ISBN 90-407-2171-8, The Netherlands

Kiekel, P.; Cooke, N.; Foltz, P.; Gorman, J. & Martin, M. (2002). Some Promising Results of Communication-Based Automatic Measures of Team Cognition. *Proceedings of the Human Factors and Ergonomics Society's Annual Meeting*, USA, Vol. 46, pp. 298-302.

Kiekel, P.; Cooke, N.; Foltz, P.; & Shope, S. (2001). Automating Measurement of Team Cognition through Analysis of Communication Data. In Smith, M.; Salvendy, G.; Harris, D. & Koubek R. (Eds.), *Usability Evaluation and Interface Design*, Mahwah, NJ: Lawrence Erlbaum Associates pp. 1382-1386, ISBN 0-8058-3607-1

Kiekel, P. & Cooke, N. (2004). Human Factors Aspects of Team Cognition. In Vu K. & Proctor R. (Eds.), *The Handbook of Human Factors in Web Design*, Mahwah, NJ: Lawrence Erlbaum Associates. pp. 90-103, ISBN 978-0805846119

Klinect, J.; Murray, P.; Merritt, A. & Helmreich, R. (2003). Line Operations Safety Audit (LOSA): Definition and operating characteristics. *In Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 663-668). Dayton, OH: The Ohio State University.

Krifka, M. (2004). Structural Features of Language and Language Use, In: Dietrich, R., Childress, T. *Group Interaction in High Risk Environments*, The GIHRE Project. Ashgate, ISBN 978-0754640110, Burlington

Lepine, J.; Hollenbeck, J.; Ilgen, D. & Hedlund, J. (1997). Effects of Individual Differences on the Performance of Hierarchical Decision-Making Teams: Much More Than G. *Journal of Applied Psychology*, Vol. 82, pp. 803–811, ISSN 0021-9010

Littlepage, G.; Schmidt, G.; Whisler, E. & Frost, A. (1995). An Input-Process-Output Analysis of Influence and Performance in Problem-Solving Groups. *Journal of Personality and Social Psychology*, Vol. 69, pp. 877–889, ISSN 0022-3514.

Mathieu, J.; Heffner, T.; Goodwin, G.; Salas, E. & Cannon-Bowers, J. A. (2000). The Influence of Shared Mental Models on Team Process and Performance. *Journal of Applied Psychology*, Vol. 85, pp. 273–283, ISSN 0021-9010

Mohammed S.; Klimoski R. & Rentsch J. (2000). The Measurement of Team Mental Models: We Have No Shared Schema, *Organizational Research Methods*, Vol. 3, No. 2, pp. 123-165, ISSN 1094-4281

Molleman, E.; Nauta, A. & Jehn, K.(2004). Person-Job Fit Applied to Teamwork: a Multilevel Approach. *Small Group Research*, Vol. 35., pp. 515- 539, ISSN. 10464964

Moynihan, L. & Peterson, R. (2004). The Role of Personality in Group Process. In: Scheinder, B. & Smith, D. *Personality and Organization*. Lawrence Erlbaum Associates. ISBN 978-0805837582

Mount, M. & Barrick, M. (1995). The big five personality dimensions: Implications for research and practice in human resources management. *Research in Personnel and Human Resource Management*, Vol. 13, pp. 153-200, ISSN . 0742-7301

Mumaw, R. (1994). The Effects of Stress on Nuclear Power Plant Operational Decision Making and Training Approaches to Reduce Stress Effects (NUREG/CR-6127). *US Nuclear Regulatory Commission.* Manuscript. Westinghouse Electric Corporation.

Neuman & Wright, J. (1999). Work-groups; Ability; Intelligence; Personality-traits; Performance-evaluation. *Journal of Applied Psychology* Vol. 84, No.3, pp. 376 – 89, ISSN 0021-9010

Piedmont, R. & Weinstein, H. (1994). Predicting Supervisor Rating of Job Performance Using the NEO Personality Inventory. *Journal of Psychology*, Vol. 128. pp. 255-265, ISSN 128255-265

Pléh Cs.; Síklaki I. & Terestyéni T. (1997). *Nyelv – kommunikáció – cselekvés*. Osiris. pp. 188-197 ISBN 9789633793046, Budapest

Reason J.(1997). *Managing the Risk of Organisational Accidents*, Ashgate Publishing Limited, ISBN 978-1840141054, England

Sorensen J. (2002). Safety Culture: A Survey of the State of the art, *Reliability Engineering System Safety* Vol.76, pp.189-204, ISSN 0951-8320.

Sexton J. & Helmreich R. (2000). Analyzing Cockpit Communications: The Links Between Language, Performance, Error, and Workload. *Human Perforance Extrem Environment*, Vol. 5 No.1, pp.63-68, ISSN 1529-5168

Salas, Dickinson, Converse, and Tannenbaum (1992). Toward an Understanding of Team Performance and Training, IN Swezey & Salas *Teams: Their Training and Performance*, Ablex Publishing Corporation Norwood, ISBN 978-0893919429, New Jersey

Salas E. & Fiore, S. (2002). *Team Cognition: Understanding the Factors that Drive Process and Performance*, American Psychological Association, pp. 83-106 ISBN 978-1591471035, Washington, DC

Salgado, J. (2001). Personnel Selection Methods. In: Robertson, I. & Cooper, C. *Personnel Psychology and Human Resource Management*. University of Manchester. John Wiley and Sons Ltd. ISBN 978-0471495574

Swain, K. & Mills, V. (2003). Implicit Communication in Novice and Expert Teams Land Operations Division Systems Sciences Laboratory DSTO-TN-0474

Sträter, O. & Fokuda, R. (2004). Communication in Nuclear Power Plants (NPP). In: Dietrich, R., Childress, T. M. *Group Interaction in High Risk Environments*. The GIHRE Project. Ashgate, ISBN 978-0754640110, Burlington

Sacks, H. (1992). *Lectures on Conversation.* Oxford: Basil Blackwell ISBN 1557862192

Tett, R.; Jackson, D. & Rothstein, M. (1991). Personality Measures as Predictors of Job Performance: A Meta-Analytic Review. *Personnel Psychology*, Vol. 44. pp. 703-742, ISSN 0031-5826

Wagner, J. (1995). Studies of Individualism-Collectivism Effects on Cooperation in Groups *Academy of Management Journal*, Vol. 38, No. 1, pp. 152-172, ISSN 0001-4273

Waller M.; Gupta, N. & Giambatista R. (2004). Effects of Adaptive Behaviors and Shared Mental Models on Control Crew Performance, *Management Science*, Vol. 50, No. 11, pp. 1534–1544, ISSN 0025-1909

Waung, M. & Brice, T. (1998). The Effects of Conscientiousness and Opportunity to Caucus on Group Performance. *Small Group Research*, Vol. 29, pp. 624–634. ISSN 10464964

Witt, L.; Barrick, M.; Burke, L. & Mount, M. (2002). Research Reports. The Interactive Effects of Conscientiousness and Agreeableness on Job Performance. *Journal of Applied Psychology*, Vol. 87. No. 1, pp. 164-169, ISSN 0021-9010

Yule, S.; Flin, R.; Paterson-Brown, S.; Maran, S. & Rowley, D. (2006). Development of a rating system for surgeons' non-technical skills. *Medical Education*, Vol. 40, pp. 1098-1104 ISSN 0308-0110

19

The Human Factors Approaches to Reduce Human Errors in Nuclear Power Plants

Yong-Hee Lee[1], Jaekyu Park[2] and Tong-Il Jang[1]
[1]*I & C - Human Factors Research Division, Korea Atomic Energy Research Institute,*
[2]*Department of Industrial Management Engineering, Korea University,*
Republic of Korea

1. Introduction

After the Three Mile Island accident, people have been showing a growing interest in human errors in the nuclear field. Human errors in nuclear power plants have been an important factor in the human factors researches. As a part of human factors practices, nuclear power plants are conducting safety assessments such as Periodic Safety Review (PSR) and Probabilistic Safety Assessment (PSA) in order to reduce any possibility which might cause major accidents or damage. Especially, to reduce the human errors recently not only some efforts to eliminate human related causes have been attempted, but also a means to widely manage the human errors such as Human Factors Management Program (HFMP) has been developed. Therefore, this chapter concerns the properties of complex systems and addresses the various practices of human factors.

Traditionally, approaches to reduce human errors were to classify error types and to protect operators from Performance Shaping Factor (PSF). Classification of human errors were conducted in previous studies and applied to find cause factors in various ways. These analyses of human errors were gradually systematized as a safety reporting system. Especially, short-term and long-term countermeasures were attempted to minimize possible human errors in a complex system such as aircraft, nuclear power plant. The purpose of these countermeasures is to reduce human errors and to improve performance of operators and systems at the same time. These countermeasures previously focused on satisfaction assessment of Man-Machine Interface (MMI), but these focused on broad considerations which are operator related such as safety culture, communication, aptitude test, etc. in recent. With regulation, the will of employer and the safety consciousness of operators are necessary to manage them efficiently. This chapter concerns various trials of reducing human errors and discusses requirements to perform an identification of them.

2. Complex systems

In general industry, insuring tranquility of individuals and systems is acknowledging a prerequisite because that is the first requirement to avoid accidents. Especially, accidents

in complex system do not simply lead to worker's individual disaster or pecuniary damage, but create unexpected damage or develop managerial risks or social issues. Therefore, when safety requirements are not satisfied in complex systems, the issues can bring serious problems in not only individuals and managements but also a longer society (Lee, 2006).

Process industry contributes to society by mass production based on cost-effectiveness through large system of high-reliability. However, if an accident occurs in this industry, it can be face with rejection rather than contribution because of a great deal of damage.

Recently, various safety-related efforts are being focused on large systems from workplace management for preventing individual's injuries to risk management for maintenance of systems, customer protection like Product Liability (PL) law, and preparing an expansion of social damage.

Among these efforts, safety related hardware is improving rapidly with the high-level reliability, but safety insurances for preparing human errors are not satisfied relatively. Especially, because relative importance of human errors related to accidents increases in large systems of high-reliability, it is difficult to determine correctly a direction of effective prevention. So, if efforts of accident prevention are effective, we should be more careful of which precautions are premised on relatively important considerations related to human error. That is, even if process industry is a high-tech industry of high-reliability, if it has not enough reliability, it can face unexpected new dimension of social antipathy, in spite of the achievement of industrial effect and role.

Therefore, we have to investigate items which consider improving the efforts for accident prevention in complex systems such as nuclear power plants. According to Lee's study (2006), he examined ten empirical reviews related to human error in nuclear fields and suggested basic considerations that need to prevent accidents. These reviews are misunderstandings about human error and suggestions which be found by trial and error;

1. Human error in an accident occurs by accident
2. Human error can be captured by the statistics
3. Human error is to blame to human
4. Human error can be reduced by enforcements
5. Human error can be reduced by voluntary efforts
6. Human error never recurs to the same human
7. Human error can be prevented by eliminating causes
8. Performance also means safety
9. The same cause, the same accident
10. Keep the basic principles against human errors

Also, according to the previous studies, accidents related to the complex system's reliability have three major the properties as follows (Lee, 2003; Park et al., 2008);

1. Dependency and inherence of an accident
2. Representativeness and latency of an accident
3. Chaining and structural properties of accidents

Not only nuclear power plants but other various major industrial accidents have these properties. Thus, the properties as stated above have to be considered to improve safety in a procedure for an accident analysis.

3. Human factors practices in NPPs (Nuclear Power Plants)

In the following, we introduce human factors practices which include human factors assessment and management in NPPs.

3.1 PSR

PSRs were adopted in order to guarantee the continued safe operation of nuclear power plants. PSRs are focused on considering various aging effects and are generally conducted approximately every ten years, and for this, analysis procedures are required such as an inspection, structure analysis, failure assessment and a combination of them (IAEA, 2010; Ko et al., 2006).

Through PSRs in Korean NPPs, the status of various human factors in operating NPPs has been reviewed by human factors experts and independent operation experts. Many points that are not suitable in a human factors sense have been revealed and remedies for these have also been discussed between the reviewers and plant personnel (Lee et al., 2004a, 2004b, 2004c; Lee et al., 2006a, 2006b).

In the process of PSRs, two different types of responses from plant personnel have been identified. One is to encourage our reviews and admit the findings as valuable information for upgrading human factors in their plant. Another is to refuse to assist in the reviews and to insist that they do not have any human factors problems.

We will describe here in detail about a PSR of human factors since we think that our PSR activities contribute considerably to an enhancement of the human factors in NPPs.

Our PSR of human factors complies with the IAEA (International Atomic Energy Agency) safety guide (IAEA, 2003). The following items are defined in the IAEA guide;

a. Staffing levels for the operation of a nuclear power plant with due recognition of absences, shift working and overtime restrictions
b. Availability of qualified staff on duty at all times
c. Policy to maintain the know-how of the plant staff
d. Systematic and validated staff selection methods (e.g. testing for aptitude, knowledge and skills)
e. Programs for initial training, refresher training and upgrading training, including the use of simulators
f. Training in safety culture, particularly for management staff
g. Programs for the feedback of operating experience for failures and/or errors in human performances that have contributed to safety significant events and of their causes and corrective actions and/or safety improvements
h. Fitness for duty guidelines relating to hours of work, good health and substance abuse
i. Competence requirements for operating, maintenance, technical and managerial staff
j. Human-machine interface: design of the control room and other work stations; analysis of human information requirements and task workload; linkage to PSA and deterministic analyses
k. Style and clarity of procedures.

These broad areas were grouped into five categories; (1) procedures, (2) Human Machine Interface (HMI), (3) human resources, (4) human information requirements and workload, and (5) use of experience. The relationship between the five areas and the IAEA assessment items is shown in Table 1.

Our assessment areas	Items defined in IAEA safety guide
(1) Procedures	(k)
(2) HMI	(j)
(3) Human resources (4) human information requirements and workload	(a), (b), (c), (d), (e), (f), (h), (i)
(5) Use of experience : incorporated into (1), (2), and (3)	(g)

Table 1. Relations between our assessment area and the PSR items defined in IAEA safety guide (IAEA, 2003)

For the five assessment areas, the details of these assessments are described as following.

1. Procedures

Class	Detail
a. Check Points	• The availability of the procedures; to evaluate if the plant provides procedures that explicitly identify the tasks related to plant safety • The appropriateness of the style and structure; to assure that procedures do not result in an excessive load to operators and cause them to become confused during their task performance • The suitability of the detailed elements; to evaluate if the structure and properties of the procedures satisfy the requirements in NUREG-0899, NUREG-1358, NUREG/CR-1999, other relevant NRC documents, IAEA TECDOC-1058, and various plant procedure management and guideline documents
b. Methods	• Procedural document reviews • Interviews with plant personnel • On-site reviews • Expert panel reviews
c. Scope	• Operation procedures: EOPs (emergency operation procedures), GOPs (general operation procedures), SOPs (system operation procedures), AOPs (abnormal operation procedures), and alarm procedures • Many departmental procedures

Table 2. Procedures assessment

2. HMI

Class	Detail
a. Check Points	• The availability of HMI; to evaluate if all HMI elements are provided as required for a performance of the tasks. Comparison between a list of HMI elements made from the analyses of the operation procedures and operator interviews and the list of HMI elements on the control boards • The suitability of HMI; to verify if the HMI properties are suitable for human factors guidelines NUREG-0700 or NUREG-0700 Rev. 2 and KINS-G-001 chapter 18.
b. Methods	• The effectiveness of HMI; to assure that HMI supports task performance so that operators can achieve the intended task objectives through the HMI. Experiments in a plant simulator were performed to evaluate the effectiveness of HMI • The suitability of the work environmental conditions; to check by measurement if illumination, noise, vibration, etc. on selected spots are within required limits
c. Scope	• MCRs (main control rooms), RSPs (remote shutdown panels), local control panels, SPDS (safety parameter display systems), and main computer systems

Table 3. HMI assessment

3. Human Resource

Class		Detail
a.	Check Points	*Work Management* • working hour management (e.g. adequate work hour, overtime) • shift management (e.g. rules of shift work, shift rotation schedule) • job substitute management (e.g. job substitute considering qualification, authority, and human factors) • work management during an O/H (overhaul) period *Health Management* • medical examination (e.g. epidemiology) • mental health and alcohol, substance abuse • health promotion activity (e.g. musculoskeletal disorders) • job satisfaction and devotion • health promotion activity • staff morale *Recruit and Qualification* • recruit (e.g. criteria for recruiting) • qualification and requirements for NPP personnel • maintaining a specialty of plant personnel *Training Program* • execution of SAT (Systematic Approach to Training) • assessment of instructors (academic career, job career) *Safety Culture* To check if the plants make an effort to enhance the awareness of plant safety through education; • plan and contents of safety culture education Operator Training using Simulators; To check if the plants provide adequate simulator training to operators for them to operate the plant safely and manage emergency state well; • training program for operators using simulators • suitability of simulators (e.g. facility status, maintenance and management status)
b.	Methods	• document reviews • structured interviews with plant personnel • on-site reviews • expert panel reviews

Table 4. Human resource assessment

4. Human Information Requirements and Workloads

Class		Detail
a.	Check Points	To determine if explicit task information requirements are satisfied and if a job operation by a department, a plant person, and an individual task is appropriate
b.	Methods	• selection and reviews of a total of 80 departmental procedures • structured interviews with plant personnel • on-site reviews • expert panel reviews
c.	Scope	• mental workload related information requirements • other factors related information requirements - personal requirements; expertness, experience, job characteristics, levels of knowledge - organizational requirements (among individuals or departments) ; work orders, training and education factors • environmental factors related information requirements (e.g. illumination, noise, vibration) • Workload (e.g. objective and subjective workload, physiological workload)

Table 5. Human information requirements and workloads assessment

5. Use of Experience

Class	Detail
a. Check Points	To review various operational experiences and to incorporate the findings of the above elements into human factors
b. Scope	• the issues and recommendations raised by the regulatory body • operation and maintenance experience • trip and event reports • human error reports • minor deficiency reports, • the implementation of the TMI action plan • FSAR changes

Table 6. Assessment for use of experience

Conclusively, reviews of human factors in NPPs by external experts have revealed many human factor problems which have remained hidden. Through PSRs, practical methods to assess the factors other than HMIs and the procedures have been established.

3.2 HFMP (Human Factors Management Program)

From the results of our PSRs, it has been found that human factors in NPPs need to be managed continuously by an organization inside the plant. For this reason, we are developing a prototype of the HFMP. We introduce the HFMP here as a proposition for a human error management in NPPs.

It will have a top level general human factors management procedure document, and detail documents for practice procedures, checklists, and technical criteria. The top level procedural document contains a general procedure and other information such as purpose, scope of application, references, definition, responsibility, and basic articles including the organization, committee, training and education for the operation of the HFMP. Plant personnel who are exclusively in charge of human factors are newly assigned and a committee for the HFMP operation is formed in the plant. General HFMP procedure has the form of a Plan-Do-Check or Study-Act which is a basic process in a BPM (business process management). It describes procedures for planning, execution and operation, assessments, reviews by the HFMP committee and decision making. Attachments of detailed procedures are provided for the management of individual human factors such as plant procedures, work management, qualification, training and education, workload management, HMI, and human error management. These items are considered in the HFMP based on the requirements for a PSR of human factors in NPPs. HFMP will have a complete form this year and many discussions with plant personnel and many cases of a real application will be attempted to establish the system. Figure 1 shows a structure of documents which include procedures and guides for HFMP.

4. Human error analysis

4.1 Human error taxonomy

When designing installations for safety-related complex systems it is important to be able to analyse the effect of human errors on essential tasks. For this reason the sensitivity and reliability of these systems to errors must be judged from some kind of EMEA (Error Mode and Effect Analysis) based on a classification of types of human error. To be useful also for adopting new technology in the HMI, taxonomy through psychological mechanisms is

necessary rather than taxonomy derived from behaviouristic classification (Rasmussen, 1988). However, there is no generally approved and used taxonomy for human errors. Taxonomy for human errors is just made for specific purpose.

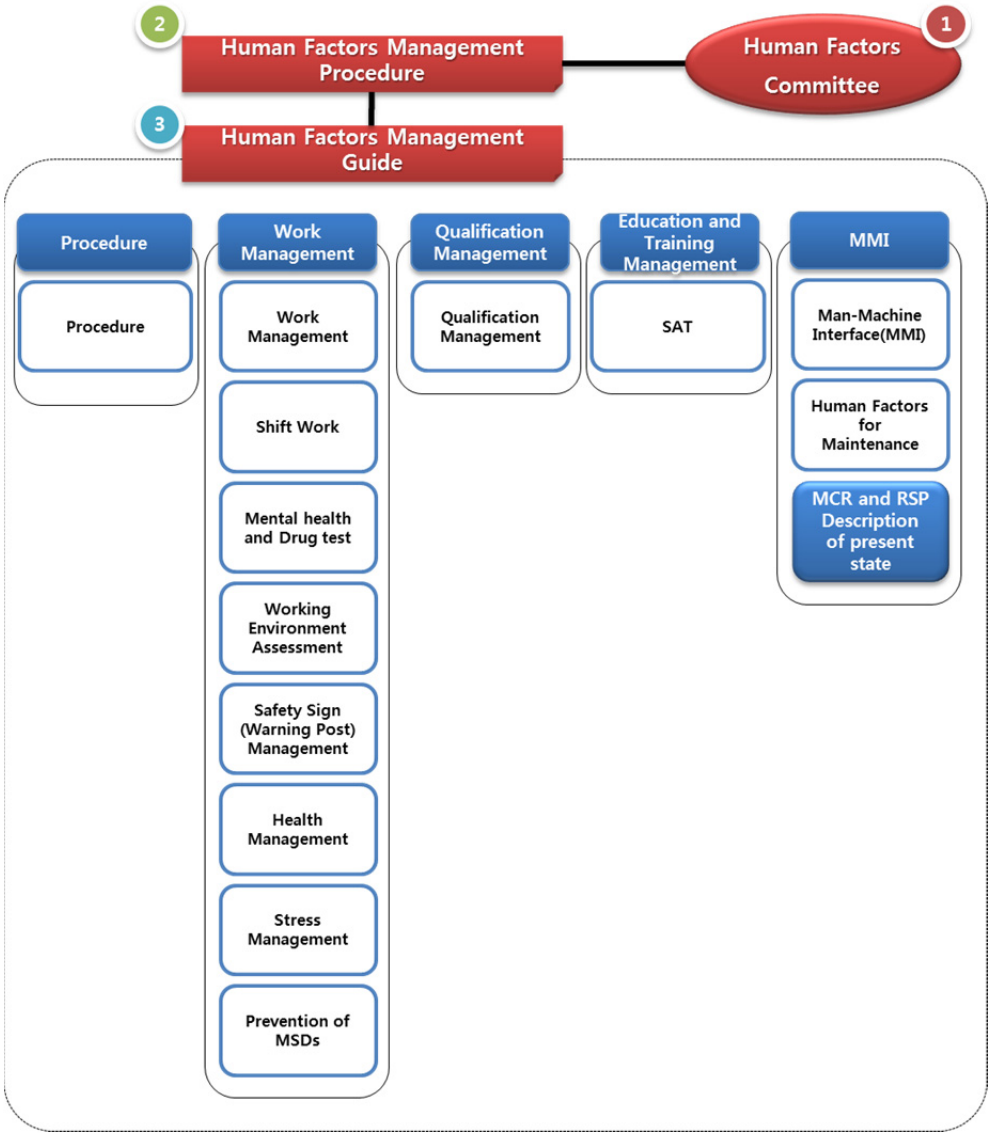


Fig. 1. Construction of HFMP

Swain (1982) suggested task-based taxonomies which state what happened. These taxonomies of human errors are classified as either error of commission or error of omission. Error of omission is defined as slip or lapse in performing a task, while error of commission

is defined as erroneous action while executing a task. Many studies of human error taxonomy focused on symbolic processing models. These approaches are more cognitive in their direction, and consider the human as having reference mental models and how things work, and how to perform. Rasmussen (1982) suggested SRK model. Reason (1990) suggested 4 types of error modes slips, lapses, mistakes and violations. Hollnagel (1993; 1996) suggested Cognitive Reliability and Error Analysis Method (CREAM) which based on a set of principles for cognitive modelling, Simple Model of Cognition (SMoC), Contextual Control Model (COCOM).

4.2 IAD (Industrial Accident Dynamics)

The dependency and the potentiality of the hazards in a NPP are defined by estimating the relative factors of the events using the IAD diagram as shown in Figure 2.

IAD matrix has usually been applied to the industrial safety domain. This technique is arranged as seven accident occurrence stages (background factors, background and initiating factors, initiating factors, intermediate factors, immediate factors, near accident, and accident) and two management stages (measurable results and countermeasures) for the column of a matrix. A row consists of four general classes: (1) machine, material and object of work, (2) human, (3) environment, and (4) others (management, supervision, education, etc.). In nuclear field study, these factors were modified to make the best use of the Frank Bird's accident theory (lack of control, fundamental factors, immediate factors, accident and injury) for ensuring an easiness of analyses. And finally, the IAD matrix consisted of the managerial and influencing factors, the fundamental causes and factors, unsafe conditions, unsafe actions, accident inducing factors, and the result and loss, as well as the 4M (Lee et al., 2007; Hwang et al., 2007; Hwang et al., 2008).

4.3 HPES

Human Performance Enhancement System (HPES) developed originally by INPO has been used in many countries, including Korea. In the case of our country, the Korean utility company modified the original HPES to become K-HPES similar to the J-HPES in Japan, which is a Japanese version of HPES and was developed by the Central Research Institute of Electric Power Industry (CRIEPI). The development and application of K-HPES was led by the top management of the Korean utility company in the early 1990s. The top management compelled plant personnel to generate K-HPES reports to the pre-assigned number of cases during the early years of its application. This enforcement hindered the advantages of voluntary reporting and brought about adverse effects in the use of the system. Workers felt stress by this reporting assignment, additional to their normal work, and sometimes reported artificial data, and hesitated to use the reports in their work practice.

Another feature of the initial version of K-HPES that caused its failure was the difficulty of plant personnel to produce a report by using K-HPES. It used many cognitive terms that are not understandable to plant personnel and required a high level of skill in the analysis of human error cases.

Many revisions have been performed. The system has become more simple and a web-based version has been developed (Jung et al., 2006). Also the compulsive attitude of management in the operation of K-HPES was mitigated. An analysis and report generation can be done with the web-based K-HPES. New K-HPES without the disadvantages that the initial version had may help plant personnel to reduce the number of human errors.

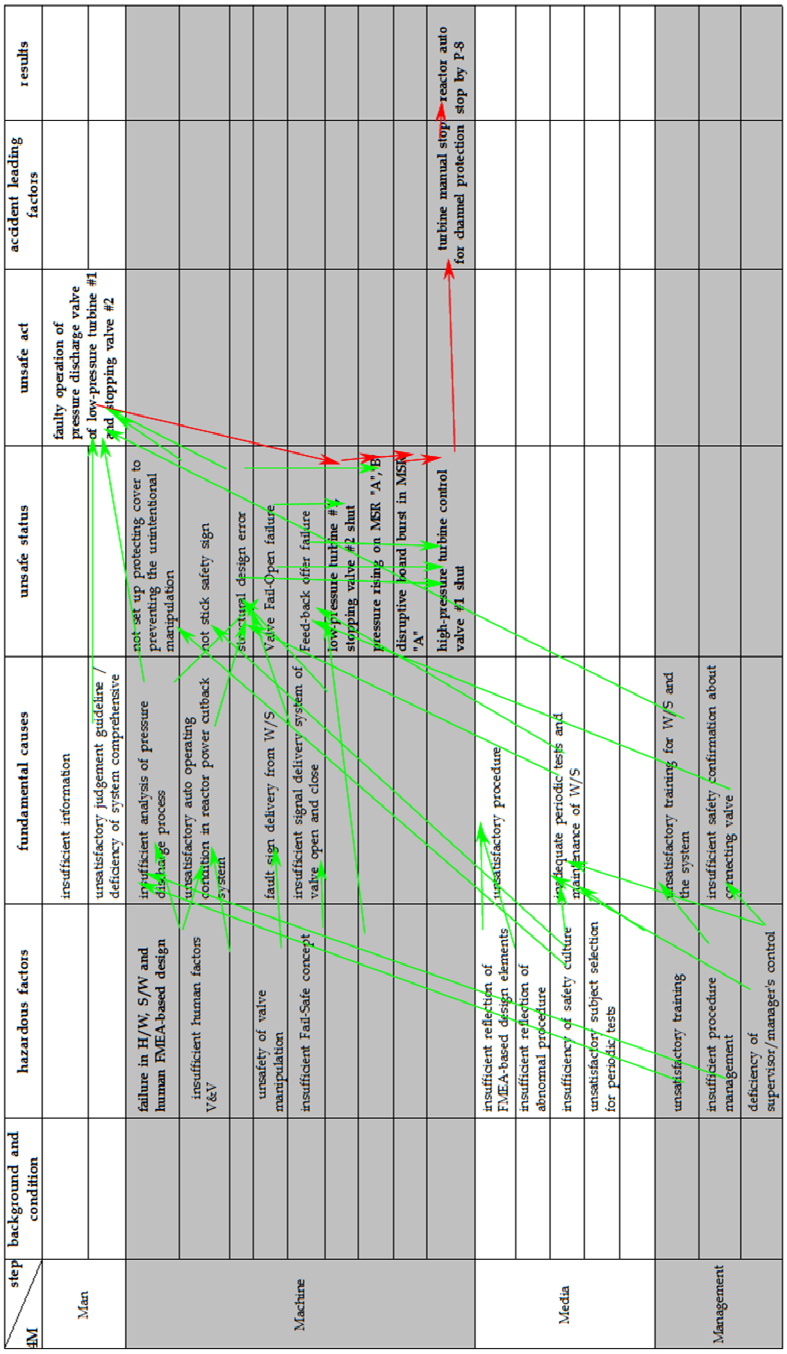


Fig. 2. Hazard factors using the IAD diagram – case study

5. Countermeasures of reducing human error in Korea

During the period 2004 to 2005, the Nuclear Safety Commission has suggest the importance of short and long term countermeasure as trip events of NPPs by human error has grown in Korea (KINS, 2006). Thereby, as part of countermeasure of reducing human error, human factors and nuclear power experts established a basic plan for reducing human error. These long-term countermeasures, the three directions and ten practical tasks have been selected and promoted. Also, experts suggested implementation plan for reducing human errors based on these practical tasks (Table 7).

Plan	Main execution items
Development of system and program for individual job analysis	- Management of individual job list of departments - Personality/aptitude tests and psychology tests - Establishment of job fitness
Task analysis of procedures related in safety	- A state-of-the-art review of task analysis methods - Development of task analysis methods
Improvement of method for EOP (Emergency Operating Procedure) presentation	
Grasp and improve communication types	- Analysis communication types among operators - Analysis communication types between operator and local - Analysis communication types between operator and support group - Improvement of communication channel and offer of communication tool
Development of teamwork enhancement technique	- Development of teamwork enhancement technique and reflection to training - Development of teamwork enhancement index
Simulator construction and application using web virtual technology	
Korea human error program development based on behaviour	- FMS (Fundamental Monitoring System), examination, human error tracking - Compensation for behaviour
Human factors assessment support	- MCR environment assessment - Human factors review support of automatic facility
Job support system development using mobile	

Table 7. Implementation plan suggested by experts group

Recently, the Korea Atomic Energy Research Institute (KAERI) is developing several technologies for human error reduction and suggests plans as countermeasure. The following sections are main activities or assessment for human factors management (Lee et al., 2011).

5.1 A suitability evaluation for human resources

A suitability assessment of department assignment intends to prevent human errors through job assignment considering employees' ability. Also, a purpose of this assessment is establishing an effective suitability assessment and developing an application plan in Korea NPPs. KAERI utilize the Organizational Personality Type Indicator (OPTI) which is developed to identify relationship with validity, immersion and satisfaction, based on

relationship and propensity correlation between personality types of individual and organization in organization diagnosis, development, personnel administration and psychology (O'Reilly et al., 1991; Yoo, 1999). Especially, the assessment guaranteed applicative possibility in business for a suitability assessment of department assignment through analyzing factors needed preliminary application after investigating relationship among propensities of organization, team administrator and individual.

5.2 A development of job suitability criteria

A Fitness for Duty (FFD), decision criteria of job suitability in human factors aspects, is developed to prevent human errors related in job of employee and improve job efficiency. The FFD derived factors which are necessary to manage human resources of employees in Korea NPPs using analysis for 10CFR26 (U.S. standard), ILO standard, employee characteristic and present state of suitability management. The reduced management factors are health diagnosis, mental health, drug management, job stress management, behavior observation, fatigue management, employee support and so on.

5.3 A human error analysis method for digital devices

In order to introduce advanced digital devices, KAERI analyzed types of human errors which occur on processes when user of digital devices use and developed plans which evaluate occasion possibility. Even if the digital devices are the same controller, the properties of devices can differ with the results through control methods. So, considering this point of view, they defined the Interaction Segment (IS) and the Error Segment (ES) which combined external physical units and control methods, and derived the types of human errors which are possible to rise up superposition of ES. If developed assessment applies job analysis, we can derive possible types of human errors and risk factors every types.

5.4 A communication analysis

Communication can help to harmonize job performance of employees in NPPs, but the communication can become causes of creating human errors as well as means of preventing human errors. Therefore, various studies which related in communication protocol and types between employees and interaction types with interface facilities are necessary in order to analyze communication types and improve communication tools. Especially, these studies can help to prevent hazard of human errors caused by communication.

5.5 Human error reduction campaign posters

The Korea Hydro and Nuclear Power (KHNP) bench marked the excellent foreign nuclear power plants and introduced human error prevention tools. The KHNP produced 40 posters for human performance improvement as shown in figure 3. The preceding posters which KHNP developed in 2006 give a message about specific information related to human errors events. However it is not enough to arouse interest in the effectiveness of posters because most people are favorably disposed toward a simple poster which has much of illustration. Therefore, KAERI developed new types of 30 posters for human error tools as shown in figure 4 (Lee, 2009). The developed posters illustrated the HE precursors to express effectively the primary intention and to make up for discrepancies in the current posters. The error precursors listed in table 8 were compiled from a study of the INPO's event

database as well as reputable sources on human performance, ergonomics, and human factors (INPO, 2002). These posters put the accent on worker's receptiveness than notification of information and lay also emphasis on visual characteristics.

Except for these technologies, the others propel various methods for reducing human errors. These contain a teamwork evaluation of Main Control Room (MCR) crews, a behavior based safety program, an enhancement of the procedures and a human error hazard analysis.

Category	HE precursors
Task Demands	Time pressure, High workload, Simultaneous tasks, Repetitive actions, Irrecoverable acts, Interpretation requirements, Unclear goals & responsibilities, Unclear standards
Work Environment	Distractions/Interruptions, Changes/Departure from routine, Confusing displays, Work-arounds instrumentation, Hidden system response, Unexpected equipment condition, Lack of alternative indication, Personality conflict
Individual Capabilities	Unfamiliarity with task, Lack of knowledge, New technique not used before, Imprecise communication habits, Lack of proficiency, Indistinct problem-solving skills, "Unsafe" attitude for critical tasks
Human Nature	Stress, Habit patterns, Assumptions, Complacency, Mind-set, Inaccurate risk perception, Mental shortcuts (biases)

Table 8. HE precursors

Fig. 3. An example of the preceding posters (Title : Reconfirmation of communication by habit)

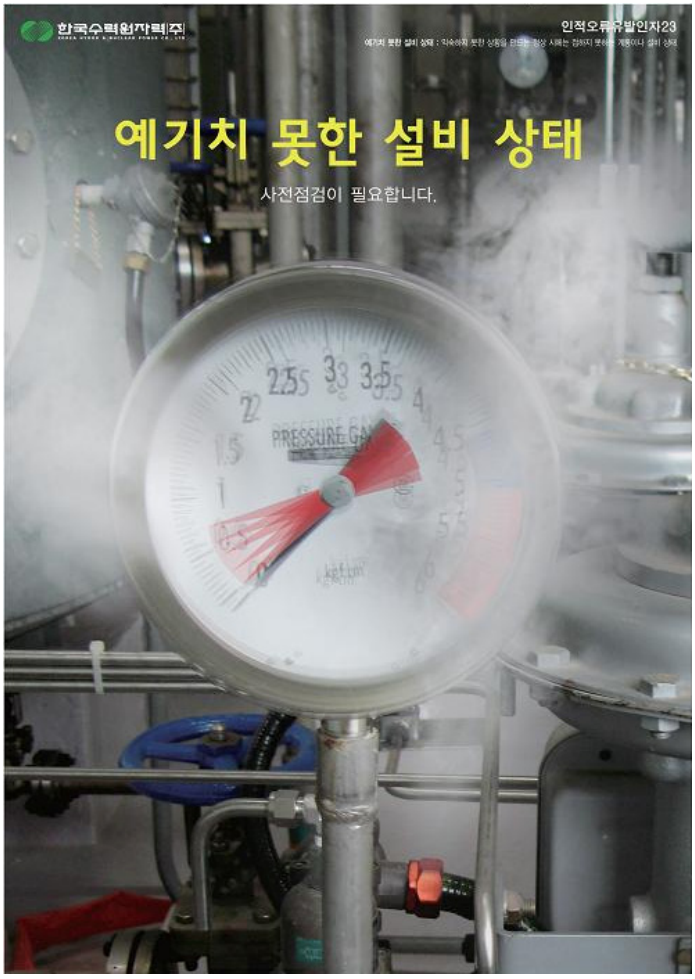


Fig. 4. An example of the developed posters (Title : Unexpected equipment condition)

6. Discussion

In this chapter, we introduce various human factors activities for reducing human errors in NPPs. Previous human factors activities were focused on regulation according to nuclear power laws. But these activities are going to expand an enterprise management as mentioned section 4-5 in recent years. The HFMP is an example of representative human factors activity in fragments. These management programs are necessary for complex systems, because many jobs interfered. That is, NPPs need integrated management systems with the parts working in coordination.

Several technologies and assessments, as mentioned section 5, are developed, and the others are going to improve still methods for preventing and reducing human errors. New

methods for reducing human errors have to identify and verify application effectiveness in on-site. These can help to offer methods to be considered for reducing human error in NPPs as well as other fields of industry.

7. References

Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control*, Academic Press, ISBN 978-0123526588, London.

Hollnagel, E. (1996). Reliability Analysis and Operator Modeling, *Reliability Engineering and System Safety*, Vol. 52, No. 3, pp. 327-337, ISSN 0951-8320.

Hwang, S. H., Kim, D. H., Oh, I. S., & Lee, Y. H. (2007). A case study for a human error analysis in nuclear power plants, *Proceedings of the Conference on Ergonomics Society of Korea*, Busan, Korea, November, 2007 (in Korean).

Hwang, S. H., Park, J., Jang, T. I., Lee, Y. H., & Lee, J. W. (2008). A research for the instrumented-related countermeasures from the precedents due to human errors in nuclear power plants, *Proceedings of the Ergonomics Society of Korea*, Gumi, Korea, May, 2008 (in Korean).

IAEA (International Atomic Energy Agency) (2010). *Periodic Safety Review of Nuclear Power Plants: Experience of Member States, IAEA TECDOC Series No. 1643*, IAEA, ISBN 978-92-0-100710-0, Vienna, Austria.

IAEA (International Atomic Energy Agency) (2003). *Periodic Safety Review of Nuclear Power Plants, IAEA Safety Standards Series No. NS-G-2.10,* IAEA, ISBN 92-0-108503-6, Vienna, Austria.

INPO (Institute of Nuclear Power Operations) (2002). *Human Performance Fundamentals Course Reference*, National Academy for Nuclear Training, Revision 6, INPO, Atlanta.

Jung, Y. S., Lee, J. S., & Chi, M. G. (2006). Development and Application of Web Based K-HPES for Human Error Analysis, *Fall Conference of Korea Nuclear Society*, Kyungju, Korea, November, 2006.

KINS (Korea Institute of Nuclear Safety) (2006). *2007 Directions of Nuclear Safety Policy (proposal)*, 32th Nuclear Safety Commission.

Ko, H. O., Chang, Y. S., Choi, J. B., Kim, Y. J., & Choi, Y. H. (2006), Development of Web-based Aging Assessment System for Major Components of CANDU Type Nuclear Power Plants, *Proceedings of Spring Conference of the Korean Society of Mechanical Engineers*, pp. 1186-1191, Jeju, Korea, July, 2006 (in Korean).

Lee, Y. H. (2003). An Effective Counter Plan of Human Errors to Promote Preventing Activity Major Industrial Accidents, *Safety and health*, the October issue, 2003 (in Korean).

Lee Y. H., Lee, H. C., & Son. H. S. (2004a). *Development of a Methodology for Analysing Information Requirement & Workload for Periodic Safety Review*, KAERI/TR-2699/2004, KAERI, Daejon, Korea (in Korean).

Lee Y. H., Lee, J. W., Park. J. C., Hwang, I. K., & Lee, H. C. (2004b). *Evaluation of Procedures for Periodic Safety Review of Nuclear Power Plants*, KAERI/TR-2880/2004, KAERI, Daejon, Korea (in Korean).

Lee Y. H., Lee, J. W., Park. J. C., Hwang, I. K., & Lee, H. C. (2004c). *Evaluation of Man-Machine Interface for Periodic Safety Review of Nuclear Power Plants*, KAERI/TR-2881/2004, KAERI, Daejon, Korea (in Korean).

Lee, Y. H. (2006). A More Effective Approach to the Analysis and the Prevention of the Human Errors, *Proceedings of Spring Conference of Ergonomics Society of Korea*, Seoul, May, 2006 (in Korean).

Lee Y. H., Park, J. C., Lee, J. W., Lee, H. C., Han, J. B., & Hwang. I. K. (2006a). *Human Factors Evaluation of Procedures for Periodic Safety Review of Yonggwang Unit # 1, 2,* KAERI/TR-3123/2006, KAERI, Daejon, Korea (in Korean).

Lee Y. H., Park, J. C., Lee, J. W., Han, J. B., Lee, H. C., & Hwang. I. K. (2006b). *Human Factors Evaluation of Man-Machine Interface for Periodic Safety Review of Yonggwang Unit # 1,2,* KAERI/TR-3124/2006, KAERI, Daejon, Korea (in Korean).

Lee Y. H., Lee, J. W., Park, J, C., Hwang, I. G., Lee, H. C., Moon, B. S., Jang, T. I., Kim, D. H., Hwang, S. H., Park, J., Lee, J. G., Lee, S. K., Koo, J. Y., & Hong, J. M. (2007). *Human Error Cases in Nuclear Power Plants: 2002~2007*, KAERI, ISBN 978-89-88114-26-1, Daejeon, Korea (in Korean).

Lee, Y. H., Lee, Y., Kwon, S. I., Jung, Y. S., & Kang, J. C. (2009). A Development of Posters for Human Performance Improvement in Nuclear Power Plants, *Proceeding of Asia Pacific Symposium on Safety 2009*, pp. 296-299, Osaka, Japan, October, 2009.

Lee, Y. H., Jang, T. I., Lee, Y., Oh, Y. J., Kang, S. H., & Yun, J. H. (2011). Research Activities and Technique for the Prevention of Human Errors during the Operation of Nuclear Power Plants, *Journal of the Ergonomics Society of Korea*, Vol. 30, No. 1, pp. 75-86, ISSN 1229-1684 (in Korean).

O'Reilly, C. A., Chatman, J., & Caldwell, D. F. (1991). People and Organizational Culture: A Profile Comparison Approach to Assessing Person Organization Fit, Academy of Management Journal, Vol. 34, No. 3, pp. 487-516, ISSN 0001-4273.

Park, J., Lee, Y. H., Jang, T. I., Kim, D., & Hwang, S. H. (2008). Applicability of an ecological interface design approach for an information requirement and workload assessment in nuclear power plants, *2nd International Conference on Applied Human Factors and Ergonomics*, ISBN 978-1-60643-712-4, USA Publishing, Las Vegas, USA, July, 2008.

Rasmussen, J. (1982). Human Errors: A Taxonomy for Describing Human Malfunction in Industrial Installations, *Journal of Occupational Accidents*, Vol. 4, No 2-4, pp. 311-333, ISSN 0376-6349.

Rasmussen, J. (1988). Human Error Mechanisms in Complex Work Environments, *Reliability Engineering and System Safety*, Vol. 22, pp. 155-167, ISSN 0951-8320.

Reason, J. (1990). *Human Error*, Cambridge, England, Cambridge University Press, ISBN 978-0521314190.

Swain, A. D. (1982). Modeling of Response to Nuclear Power Plant Transients for Probabilistic Risk Assessment, *Proceedings of the 8th Congress of the International Ergonomics Association*, Tokyo, Japan, August, 1982.

Yoo, T. Y. (1999), Scale Development for Organizational Personal Type Indicator (OPTI), *Journal of Korean Psychological Association*, Vol. 12, No. 1, pp. 113-139. ISSN 1229-0696 (in Korean).

20

Virtual Control Desks for Nuclear Power Plants

Maurício Alves C. Aghina[1,2] et al.*
[1]*Comissão Nacional de Energia Nuclear, Instituto de Engenharia Nuclear,*
[2]*Universidade Federal do Rio de Janeiro,*
[3]*Instituto Nacional de Ciência e Tecnologia de Reatores Nucleares Inovadores, CNPq*
Brazil

1. Introduction

Nuclear power is a very important option that meets the global needs for power generation. But nuclear plants' operation involves high safety requirements, due to all the potential risks involved. Nuclear power plants (NPP) must be operated under safety conditions in all stages, since its start up, and during all the process. For this reason, control desks and rooms must be designed in such a way operators can take all the procedures safely, with a good overview of all variable indicators and easy access to actuator controls. Also, operators must see alarms indication in a way they can easily identify any abnormal conditions and bring the NPP back to normal operation. These matters have been taken into account through the years in NPP control desks and rooms design, through ergonomics or human factors evaluations, to help design safer NPP control systems (Hollnagel, 1985; Pikaar, 1990; ANSI ANS-3.5, 1993; Foley et al., 1998; Feher, 1999).

Operator training routines used to be carried out in full-scope simulators that resembled real control desks with high fidelity. Then, a new concept emerged, using control systems simulators based on synoptic windows interface, with NPP dynamics computer-based simulation system. These later usually include all the dynamics involved in a NPP operation. All variables are affected by operators' actions in the synoptic windows-based interface, with responses to their actions readily presented on screen. Although the high fidelity in the NPP dynamics simulations, synoptic windows-based interfaces does not resemble much real NPP control desks, since operators have to deal with graphical diagrams on computer screens.

Virtual reality technology help NPP operation simulation, since it enables virtual control desks (VCD) prototyping, thus adding to NPP dynamics computer simulation the design of control desks with high visual fidelity with real ones. Operators can now take advantage of both the online simulation capabilities of NPP dynamics computer-based simulation systems, with a more suitable interface such as VCDs, which resemble more closely the real

* Antônio Carlos A. Mól[1,3], Carlos Alexandre F. Jorge[1,2], André C. do Espírito Santo[1], Diogo V. Nomiya[2], Gerson G. Cunha[2], Luiz Landau[2], Victor Gonçalves G. Freitas[1,2] and Celso Marcelo F. Lapa[1]
1 Comissão Nacional de Energia Nuclear, Instituto de Engenharia Nuclear, Brazil,
2 Universidade Federal do Rio de Janeiro, Brazil,
3 Instituto Nacional de Ciência e Tecnologia de Reatores Nucleares Inovadores, CNPq, Brasil.

ones. This approach have been under development and upgrading at *Instituto de Engenharia Nuclear* (IEN, Nuclear Engineering Institute) an R&D center of *Comissão Nacional de Energia Nuclear* (CNEN, Brazilian Commission of Nuclear Energy). This R&D enabled dynamic simulations and online operator interaction.
An existing NPP simulator, originally designed with a synoptic windows-based interface, was then interconnected with the developed VCD through network, either local or through the Internet, by using TCP/IP protocol. First tests were reported earlier (Aghina et al., 2008), showing the remote operation of the computer-based NPP dynamics simulator through the VCD.
Our staff has lately made improvements to this VCD, to turn its design an easier task, through the use of modules that can be added, deleted or modified. Besides this, new interaction modes have been included for an easier interfacing. There is a 3×2 m projection screen in one of our Labs, which can be used for the VCD simulation, among other ones. With friendlier interaction modes, users can interact with the VCD in front of this projection screen without using computer keyboard and mouse.
One of these new modes makes use of speech recognition-based commands. Also, an alternative interaction mode makes use of head tracking, with or without visual markers.
This Chapter spans many topics related to this R&D, ranging from the VCD development and the interfacing with the existing computer-based NPP dynamics simulator, to the new interactions modes. Thorough the chapter, related topics will be commented, such as the importance of ergonomic evaluations for safe NPP operation.

2. Nuclear power plant operation

NPP operation involves high safety requirements, due to the nature of this power source itself. Nuclear (fissile) materials must be dealt with very safely to achieve the desired objective, – that is, to generate power –, through the use of highly efficiency control systems. The nuclear fission reactions must be taken in very controlled conditions.
In NPP, fission takes place by inserting or removing control rods from the NPP core, where is the fissile material. The more operators remove the control rods from the core, the higher the operating power level.
In the following, a simplified description of pressurized water reactor (PWR) NPP is given, since this is a very common type of NPP currently in operation. PWR NPP resemble much thermoelectric power plants, in that water is heated by a power source, to move turbines associated with electric generators. The main difference is that nuclear fissile material is used as power source, instead of coal or gas.
A PWR NPP consists basically of three main parts, named: (i) primary, (ii) secondary, and (iii) tertiary. But one can consider also the electrical part, to suppy power to transmission lines. Fig. 1 illustrates a PWR NPP with its three main parts and the electrical part.
Primary includes the NPP core, that contains the fissile material and the control rods, and also the associated primary circuit, – indicated as "reactor vessel", in Fig. 1 –, where water is heated through the nuclear fission reactions. There is a pressurizer, which keeps the pressure high enough to prevent water to vaporize above 100 degrees Celsius; this is the reason for the term "pressurized" in PWR. In the primary circuit, water – shown in red, in Fig. 1 –, may have nuclear contamination. Thus, the primary and secondary water circuits are completely isolated from each other, to prevent nuclear contamination among them.

Heat is transfered from the primary to the secondary by a heat exchanger, named steam generator, the interface between both isolated circuits. This water in the secondary circuit is not kept at high pressure, but vaporizes itself to move the turbines, that are also part of the secondary.

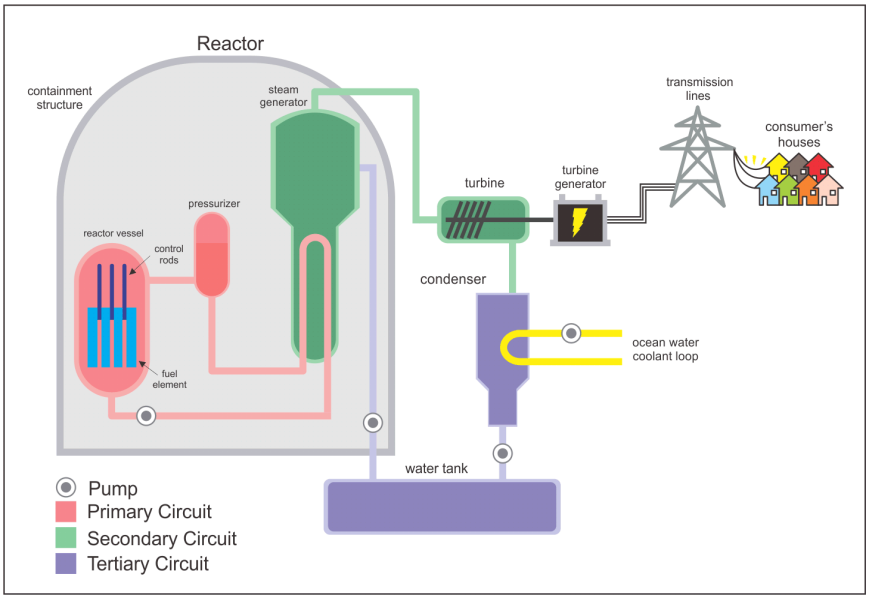


Fig. 1. PWR NPP simplified diagram[1].

As shown in Fig. 1, there is a containment structure for the nucler plant core, all the primary, and the steam generator. It consists, in the plants currently in use, of two containers, in fact, an inner container made of steel, resistant to high pressures and to corrosion, and an outer container made of concrete. This is an advancement in relation to older nuclear plants where only a concrete contaniment structure was used.

The vapour at the secondary, after moving the turbines, is condensed back to water at the condenser, a heat exchanger that is another isolated interface between the secondary and the tertiary. The water used for this purpose comes in some cases from a nearby sea or other natural water source, by pumps.

The electrical part after the NPP consists of the electrical generators coupled with the turbines axes, and all further devices and systems needed to supply power to the transmission lines, as transformers, power back-up and the frequency and phase synchronization controls.

All these parts, – primary, secondary, tertiary and the electrical part –, and related equipment, have their associated control subsystems, with all sensors, displays, actuator controls and alarm indicators. Operators are given specific tasks in NPP operation, and a supervisor must coordinate their actions. They all have to set operational conditions and monitor variables and any possible malfunctioning through the displays and alarms indicators. Once any abnormal conditions detected, they must identify fast and correctly

[1] This figure was made at IEN/CNEN with CAD software.

their cause, and mitigate their effects, bringing the NPP back to normal operational conditions. This is carried out through pre-defined procedures that must be followed during any incident or accident.
The following subsections give an overview on some topics related to this R&D, and on related R&D run by other groups.

2.1 Nuclear power plant simulators

Given the high safety requirements for NPP operation, operators must run very efficient training programs. These are usually carried out by using full-scope control desks and rooms, which resemble the real ones with high visual fidelity, with associated NPP computer-based simulation systems. The former requires the physical construction of control desks, similar to the real ones, what involves high costs and time. For this reason, only a few of such full-scope simulators are constructed, meaning operator trainees usually have to wait for some time to be trained, among other teams, besides to the need of travelling to the training site.
Fig. 2 shows, as an example, a reduced-scale full-scope simulator, in that the developed virtual control desk was based. The leftmost module is dedicated to neutron flux monitoring, control rods and general alarms. The central (and main) module is dedicated to the monitoring and control of nuclear plant's core, pressurizer, stem generator and turbines, among other tasks. The rightmost module is dedicated to the monitoring and control of the electrical power generation besides other general alarms.

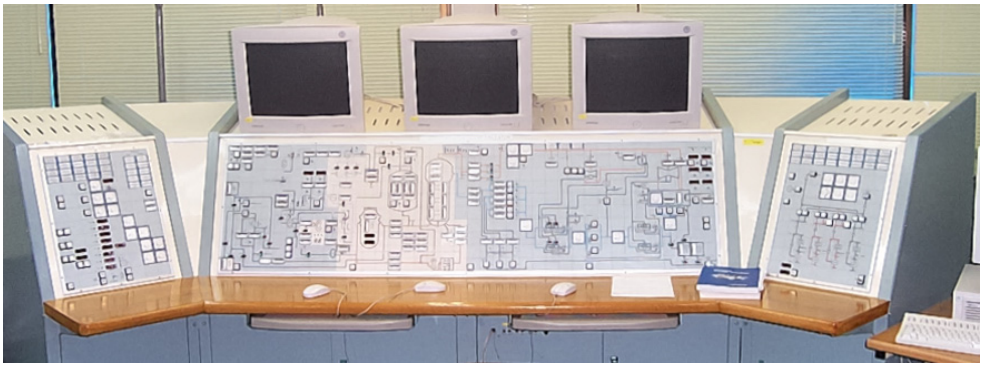


Fig. 2. Reduced-scale full-scope NPP simulator[2].

Computer-based simulators perform numerical simulation of a NPP dynamics, – in general, a simplified modelling of its dynamics, online (IAEA-TECDOC 995, 1998). This involves all the mathematical models related to neutronics, thermohydraulics, chemical, and other NPP dynamical subsystems. It deals also with all input and output variables, the former set by operators, and the later monitored by them.
Through the years, the use of physical full-scope simulators tended to be surpassed by the use of synoptic windows-based interfaces, where the NPP control desk is represented

[2] This photo was obtained by IEN/CNEN's staff during a visit to Korea Atomic Energy Research Institute.

through diagram views in computer screens. This approach is an upgrade relatively to the full-scope interfaces, but it lacks the visual fidelity with real control desks, since diagrams have quite different appearance, instead.

2.2 Virtual control desks and rooms

More recently, a new approach has come to use by some R&D groups, in which computer graphics and virtual reality technology are used to simulate control desks and rooms as visual interfaces (Drøivoldsmo and Louka, 2002; Nystad and Strand, 2006; Markidis and Rizwan-uddin, 2006; Hanes and Naser, 2006). This new approach thus combines both the NPP computer based simulator systems functionality with high visual fidelity with real control desks and rooms. These virtual interfaces become then virtual prototypes of real ones, for operator training or for ergonomics evaluation.

The fields of ergonomics and human factors have become very important topics in the design of control desks and rooms (Hollnagel, 1985; Pikaar, 1990; ANSI ANS-3.5, 1993; Foley et al., 1998; Feher, 1999). The relevance of these fields is independent of the end user application, either for nuclear plants, or for any other industrial plants, as chemical, petrochemical or industrial plants in general. The following explanation concentrates in the design of control desks, but it could be applied to control rooms too.

Ergonomics analyses enable the evaluation of the control desks' design, relatively to the location of displays, actuator controls and alarms indicators, for a safer operation. A design which does not consider these factors may turn operation a more difficult and unsafe task for personnel, due to possible misallocation of all the above mentioned devices in the control desks, so operators may not act properly in the case of incidents or accidents. A design that considers these factors takes into account the evaluation of operators' behaviour through many simulations, before a final decision about their location. Besides new control desks design, existing ones can be also evaluated and modified, to meet ergonomics requirements for safe NPP operation conditions.

3. IEN's nuclear power plant simulator

IEN's staff had been involved in a computer-based NPP simulator R&D, in a cooperation with the Korea Atomic Energy Research Institute (KAERI), and with the International Atomic Energy Agency (AIEA). This cooperation resulted in a new laboratory at IEN in 2003, named *Laboratório de Interfaces Homem-Sistema* (LABIHS, Human-Systems Interface Laboratory), (Carvalho and Obadia, 2002; Santos et al., 2008). This simulator comprises a computer-based PWR NPP dynamics simulation system, with synoptic windows-based interface, and is used mainly for operator training and ergonomics evaluation.

Fig. 3 shows a view of the LABIHS simulator room, while Fig. 4 shows its layout. There is a projection screen with a window showing an overview of the whole system. One can see two operators, – the primary (or reactor) operator and the secondary (or turbine) operator –, near these projection screens and a supervisor in the back part of the room. There is also an instructor, not shown in this figure, who stands at "Simulator Set-up Controls", shown in Fig. 4, that starts operation and may insert malfunctions into the system, which operators do not know in advance. One can also notice the use of multiple computer screens in front of operators, to minimise the need of switching among different views.

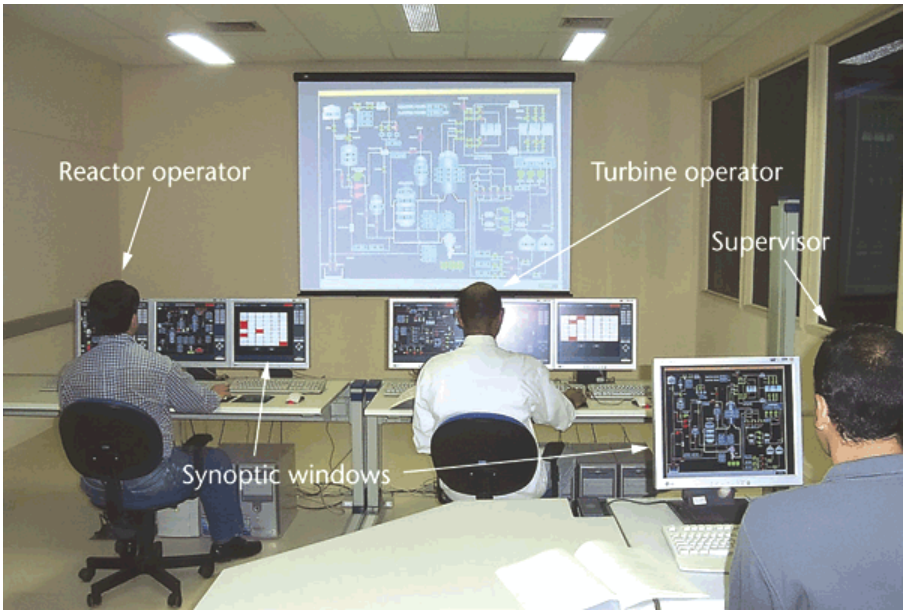


Fig. 3. LABIHS simulator room's view.

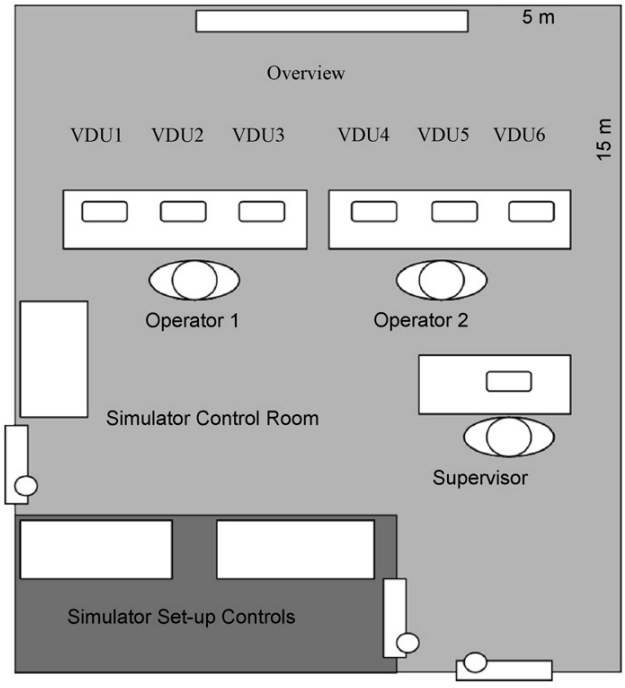


Fig. 4. LABIHS simulator room's layout.

3.1 Synoptic windows

Synoptic windows, as already briefly explained in section 2.1, mean the NPP control desk is represented through many diagram views on computer screens. This approach consists, from a point of view, in an upgrade relatively to the full-scope interfaces, which consisted of physical control desks construction. But, from another point of view, this approach lacks the visual fidelity with real control desks, since diagrams are used instead. Operator trainees have to switch among many windows, to see different parts of the NPP, to monitor variable indicators and also alarms. Multiple computer screens reduce this effort, – as can be noticed in Fig. 3 –, but even so this can be a confusing task, besides the poor appearance, far from that of a real control desk (Carvalho et al., 2008).

Some R&D have been carried out to improve these synoptic windows, following recommendations from ergonomic evaluations performed by IEN's staff, which are detailed in the references (Carvalho et al., 2008; Santos et al., 2008; Oliveira et al., 2007).

Fig. 5 shows a close view of an original synoptic window in a computer screen.

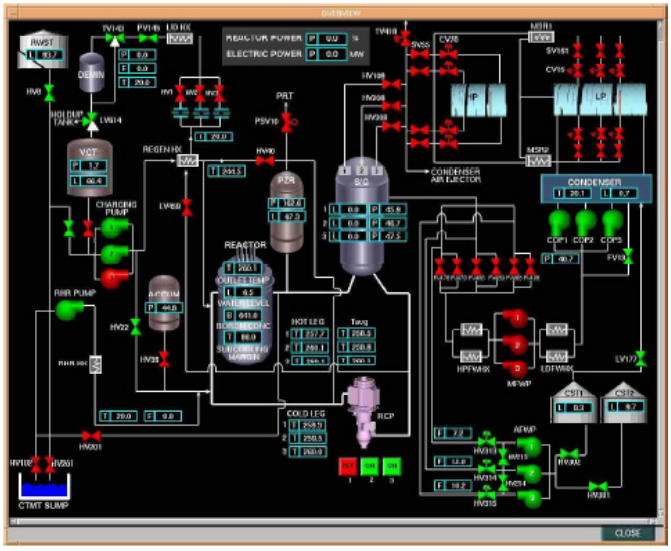


Fig. 5. Example synoptic window.

3.2 Networking

The LABIHS simulator networking can be represented in the following form (Carvalho et al., 2008). There is an interface between the simulator code and the synoptic windows, the shared memory. The later keeps updated values of all input and output variables that can be accessed by both sides, from the simulator side, or from the synoptic windows one.

Variables values, as temperature, pressure, flows, among others, are updated periodically as simulation runs, and feed the synoptic windows to inform personnel about operation conditions. Also, they can modify some other variables through actuator controls, such as "close valve A", "open valve B", "remove rods", "insert rods", and so on. These actions are readily input to the simulator code, that in turn updates simulation computation based on these new instructions.

All the tasks are performed in a local network, in which there is a central computer, operated by the instructor and where the simulator also runs, and other terminals operated by the trainees. Fig. 6 illustrates the networking scheme used, showing the shared memory rule.

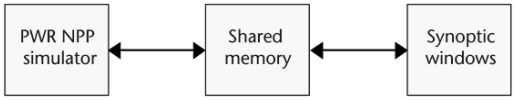


Fig. 6. The networking scheme used at LABIHS simulator.

4. IEN's virtual control desk

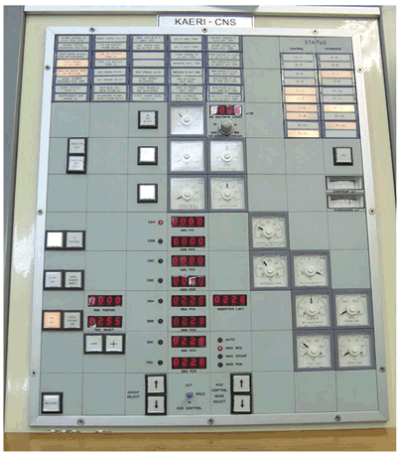
4.1 Visual interface

Due to the interactive nature of this R&D, the VCD makes use of programming, with the OpenGL graphics library, what enabled online operation training and dynamical ergonomics evaluation. The VCD developed was based on the reduced-scale full-scope physical control desk shown in Fig. 2 (Aghina, 2009).

OpenGL is a free, public domain, and high performance 3D graphics function library written for C/C++ languages. It is dealt with directly by graphics card, setting the computer processor free of tasks. It can be defined as interfacing software for graphics hardware. There is also an auxiliary library, OpenGL Utility Toolkit (GLUT), for user interaction tasks through interface devices such as computer keyboards, mouses, among others.

The VCD consists of more than five hundred graphic components, belonging to seven classes, as plane displays, cylindrical displays, buttons, switches, among others. The VCD panel is mapped in discrete Cartesian coordinates, for component location. As inserting text boxes in OpenGL is not so easy a task, textures are used instead, by pasting images from the real control desk, for the front panel's background and for the displays and buttons, for a more realistic appearance. Fig. 7 shows an overall view of the VCD; compare it with Fig. 2. Fig. 8a and 8b show a comparison of partial views of the real and the corresponding virtual versions. This partial view corresponds to the leftmost module shown in Fig. 7. Fig. 9 shows a close perspective view of the VCD, showing the realistic impression one can have by interacting with it.

Fig. 7. Overall view of the VCD.

a) 

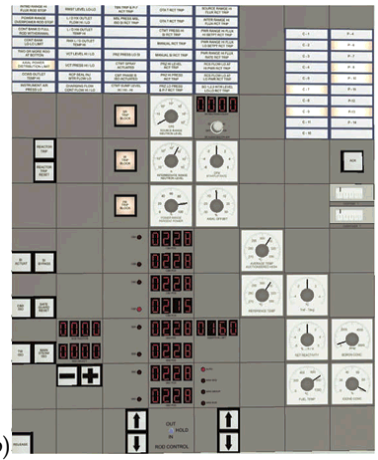
b) 

Fig. 8. a) Partial view of the real control desk; b) The corresponding virtual model.

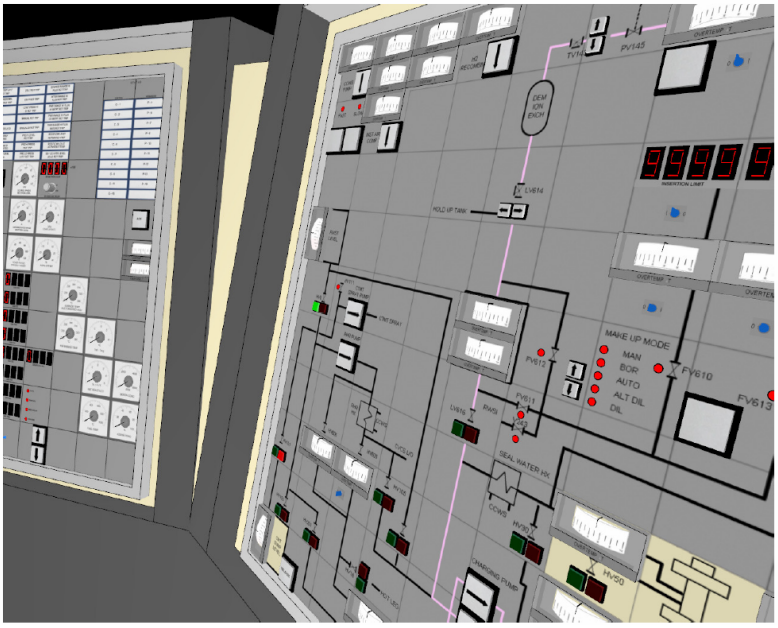


Fig. 9. A perspective view of the VCD.

4.2 Networking

The VCD now substitutes the former synoptic windows-based interface, and thus must communicate with the shared memory, for both reading variable values and feeding the simulator with input operator commands (Aghina, 2009; Aghina et al., 2008). This was done through networking, using TCP/IP protocol. Therefore, the VCD is able to communicate

with the simulator not only through local networks but also remotely, through the Internet. Fig. 10 illustrates the new networking scheme adopted. Comparing it with Fig. 6, one can see the VCD along with the TCP/IP socket now play the role formerly played by the synoptic windows-based interface.

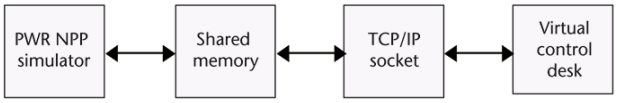


Fig. 10. The networking scheme used at LABIHS simulator.

TCP/IP protocol was chosen because it suited very well in the communication needs for this R&D. It enables friendly bi-directional data transmission between two computers, besides the versatility to be used locally or remotely through a global medium such as Internet. This is very attractive for training, as it avoids unnecessary trainees travelling to the training site where the simulator is installed, since the VCD is portable and can be used in the end-user site.

4.3 Interaction modes

In a first moment, interaction was carried out through usual interaction modes as computer keyboard and mouse. But soon our staff planned to develop other friendlier interfaces through other mode such as voice, to set user free from keyboard and mouse, and turn user interaction with the VCD a more natural task. In IEN's Labs there is a 3 × 2 m projection screen, where user can visualise different applications in an environment that enables immersion, in some sense. It would be very desirable if users could be stood up in front of this projection screen, interacting directly with the VCD, free from wired interfaces such as the ones cited before.

First, an automatic speech recognition (ASR) system was developed and implemented (Jorge et al., 2010). Lately, other interaction modes were implemented, based on head tracking, with and without visual markers. Also, a head-up display was also tested as an alternative form of visualisation, besides the projection screen or simply desktop computer screens.

All interaction modes in use are explained in the following sections.

4.3.1 Automatic speech recognition

An ASR system was developed based on well-known signal processing techniques, as cepstral analysis (Furui, 1981; Oppenheim and Schafer, 1989) of the spatio-temporal speech signals, and neural networks (NN), (Haykin, 1999; Cichocki and Unbehauen, 1993), for the pattern recognition stage. The good system's performance has been demonstrated by tests, and was published elsewhere (Jorge et al., 2010). It was first implemented offline, thus partially online by using OpenAL library and also OpenGL, for integration with the VCD. But then, it was implemented as a self-contained system, for full online operation. Other purpose was to perform direct interaction with the operating system (OS), independent of the application, so it could be used for any other application in our Labs, besides the VCD. This was achieved by using MS Windows Application Programming Interface (API). The computer thus receives the spoken command inputs as they were keyboards' or mouse's ones.

Tests were performed, showing also good performance of the system for changing views of the VCD through spoken commands, as moving it for left, right, up or down, or zooming it in or out, as examples of possible commands.

4.3.2 Head tracking

Head tracking was implemented through two approaches, with and without visual markers, as described next. But, independently of using or not markers, the main purpose of head tracking is to turn interaction a much more natural task, because an user can turn his or her head towards the specific scene location he or she needs to see with more detail.

Considering the current R&D, some examples might be: (i) one might need to look towards the leftmost VCD module (see Fig. 7) to see an alarm or any other indication; he or she needs simply to move head to the left, and the image on the projection or computer screen turns to that side. (ii) one might need to look in more details a variable indication in a specific display; he or she needs simply to approach head towards the screen (projection screen or computer screen), and the projected image zooms in. Other examples might easily be thought of, for moving to the left, up or down, or for zooming out.

These tasks can be also executed by spoken commands, as former tests that we performed. Of course spoken commands is a more natural interaction mode than using keyboard or mouse. But head movement seems a more natural interaction mode than that supplied by the developed ASR system, so our staff moved towards this type of interaction too. In fact, each interaction mode has its own advantages and disadvantages, as will be discussed later.

4.3.2.1 Head tracking with markers

After the acquisition of a head-mounted head-up display, it soon became an alternative for visualisation, besides the projection screen and desktop computer screen. Coupling of visual markers in this head-mounted display readily enabled the use of an available computer code for head pose estimation, with the use of an infrared sensor, – Natural Point's Trackir5 –, attached to a fixed location in front of user.

The tracking system with markers does not need to be used with the head-up display, as the markers can be fixed in user's head by other means. But the head-up display can not be used with the face tracking system without markers, as will be explained in the following section.

This tracking system makes use of three reflective markers filed at user's head (in the head-mounted display or by other means). The pose is estimated based on their positioning, through projective geometry computation, by a freeware library, supplied by the same company, – Natural Point. This library is OptiTrack, and performs tracking with six degrees of freedom.

This approach has a strong advantage related to accuracy in head pose estimation, due to the precise markers position detection by the fixed infrared sensor. The disadvantage, if one could mention it, is the need to use the markers at one's head.

4.3.2.2 Head tracking without markers

Another approach is the use of a code for head pose estimation based on tracking some points detected in user's face. This code is a proprietary library named FaceAPI, by Seeing Machines. The source code is not available, and the company does not supply any details about the tracking methodology used. Thus, it has been used as an executable called by our application. It also performs tracking with six degrees of freedom, and operates with either webcams or infrared cameras, as described in section 4.3.2.1.

After some tests, evaluation showed it did not offer good accuracy when user turns his or her head to the sides; at twenty degrees to both sides, the estimation of head angle, – its pose –, sometimes oscillates, making the VCD image on screen to shake for both sides. The

advantage is that there is no need to use any markers at one's head, turning interaction more natural.
This system can not be used with head mounted display, because it comes trained to detect faces, and thus any other device in one's head makes face undetectable by the code.

4.3.3 Combining the interaction modes

After many tests with the interaction modes explained above, our staff has noticed each one had its own advantages and disadvantages, and thus could be combined, each one used for those tasks where it performs better. It results, therefore, in a more complete and friendlier interfacing. More specifically, the combination is performed between the ASR and head tracking (with or without markers), not between both head tracking approaches.
The ASR system performs very well when one needs to switch quickly among different VCD views, as for example, switching from an overall view, as shown in Fig. 7, to a partial view, as shown in Fig. 8a, or to return to a previous view. One has to speak simple commands such as "left module" (or simply "left"), "up", "down", or "back", to return to a previous view.
Once switched to the desired view, through the ASR system, the head tracking system enables the user to see details in that view, by simply moving head a just bit for each side, up or down, to direct sight to a specific part of interest. User can also zoom the view in or out by approaching or apparting his or her head just a bit from the VCD image.
Using head tracking to switch among different views can be not so easy and natural a task, because one would have to move not a bit, but a lot, his or her head for both sides, up, down, approaching or apparting from screen.
From another point of view, using speech to move a bit for any position within a VCD module's view would also be not so easy a task, as user would have to speak many times for little movements.
Thus, combining ASR and head tracking results in a more natural interaction for users than using each of the modes alone. Table 1 summarises this combination.

		Task	
		Switching views	Seeing details
Interaction mode	ASR	Applies	–
	Head tracking	–	Applies

Table 1. Interaction modes choice according to the desired task.

4.3.4 The head-up display

Tests also indicated the advantages and disadvantages of using the head-up display. It is good to remind that using head markers does not mean necessarily using also head-up display. The disadvantage, if one could mention it, is the need to use such a device in his or her head.
But there is a strong advantage. When one uses head tracking (with or without visual markers), he or she must turn his or her head to the sides, or up, down, as examples, to direct sight towards a specific VCD module. But then, he or she must keep head in that position to keep looking at that module. This cause discomfort because he or she must direct only his or her eyes towards the desired detail, keeping the head turned to the sides, or up

or down. Eyes become directed towards the opposite head position, say, head turned to the left, and eyes turned to the right side.
By using the head-up display, when turning head to the desired position, that module's view is projected in the glasses in front of his or her eyes in a frontal angle, so anyone needs to direct eyes to the opposite head position. Thus, this is an important factor to be considered, if one has to keep interacting with the VCD for a long term.

5. Concluding remarks and perspectives

The developed VCD proved to be a very good alternative relatively to the synoptic windows-based approach, by aggregating both the high fidelity visual appearance with the corresponding real control desk, to the computer-based PWR NPP simulation system's functionality.
Relatively to the interaction modes, after developing and testing only the ASR system, in a first moment, and then the head tracking approaches, it suddenly turned out to be that a combination of ASR with head tracking would result in a more natural users' interaction with the VCD. Each interaction mode is used for the tasks it fits best, with ASR performing better for switching among different VCD views, or what one could call macro movements; head tracking performing better for small movements within a particular VCD module or region, or what one could call micro movements.
Another conclusion was related to users' better comfort when using the head tracking approach with head-up displays, in relation to projection screen or computer screens. But the head-up display can only be used with the visual marker approach.
The matters related to the markerless head tracking closed algorithm could be surpassed in the future, by using another methodology available in the literature.

6. Acknowledgements

This research is sponsored by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ and Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq

7. References

Aghina, M. A. C.; Mól, A. C. A.; Jorge, C. A. F.; Pereira, C. M. N. A.; Varela, T. F. B.; Cunha G. G.; Landau, L. (2008). Virtual control desks for nuclear power plant simulation: improving operator training, *IEEE Computer Graphics and Applications*, Vol. 28, No. 4, 6-9.

Aghina, M. A. C. (2009). Virtual Control Desk for Operators Training: A Case Study for a Nuclear Power Plant Simulator (in Portuguese), MSc Dissertation, Universidade Federal do Rio de Janeiro.

ANSI/ANS-3.5 (1993). Nuclear power plants simulators for use in operators training and examinations.

Carvalho, P. V. R.; Obadia, I. J. (2002). Projeto e implementação do laboratório de Interfaces Homem Sistema do Instituto de Engenharia Nuclear, *Revista Brasileira de Pesquisa e Desenvolvimento*, Vol. 4, No. 2, 226-231.

Carvalho, P. V. R.; Santos, I. J. A. L.; Gomes, J. O.; Borges, M. R. S.; Guerlain, S. (2008). Human factors approach for evaluation and redesign of human-system interfaces of a nuclear power plant simuilator, *Displays*, Vol. 29, No. 3, 273-284.

Cichocki, A., Unbehauen, R. (1993). *Neural Networks for Optimization and Signal Processing*, John Wiley & Sons, Baffins Lane.

Drøivoldsmo, A.; Louka, M. N. (2002). Virtual reality tools for testing control room concepts, In: Liptak, B. G.; *Instrument Engineer's Handbook: Process Software and and Digital Networks*, third ed., CRC Press.

Feher, M. P. (1999). Application of human factors to the Candu 6 Control Room upgrade, *IAEA Workshop Specialists' Meeting on Approaches for the 171 Integration of Human Factors into the Upgrading and Refurbishment of control rooms*, Halden, Norway, 23 to

Foley, J. D.; Wallace, V. L.; Chan, P. (1998). *Human Computer Interaction*, Prentice-Hall.

Furui, S. (1981). Cepstral analysis technique for automatic speaker verification, *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP 29, No. 2, 254-272.

Hanes, L. F.; Naser, J. (2006). Use of 2.5D and 3D technology to evaluate control room upgrades, *The American Nuclear Society Winter Meeting & Nuclear Technology Expo*, Albuquerque, NM, 12 to 16 November.

Haykin, S. (1999). *Neural Networks – A comprehensive foundation*, Prentice-Hall,. Upper Saddle River, NJ.

Hollnagel, E. (1985). A survey of man-machine system evaluation methods, Technical Report, Halden Reactor Project Organisation for Economic Cooperation and Development.

IAEA TECDOC 995 (1998). Selection, specification, design and use of various nuclear power plant training simulators.

Jorge, C. A. F.; Mól, A. C. A.; Pereira, C. M. N. A.; Aghina, M. A. C.; Nomyia, D. V. (2010). Human-system intrerface based on speech recognition: application to a virtual nuclear popwer plant control desk, *Progress in Nuclear Energy*, Vol. 52, No. 4, 379-386.

Markidis, S.; Rizwan-uddin (2006). A virtual control room with an ebedded, interactive nuclear reactor simulator, *The American Nuclear Society Winter Meeting & Nuclear Technology Expo*, Albuquerque, NM, 12 to 16 November.

Nystad, E.; Strand, S. (2006). Using virtual reality technology to include field operators in simulation and training, *27th Annual Cannadian Nuclear Society Conference and 30th CNS/CNA Student Conference*, Toronto, Canada, 11 to 14 June.

Oliveira, M. V.; Moreira, D. M.; Carvalho, P. V. R. (2007). Construção de interfaces para salas de controle avançadas de plantas industriais, *Ação Ergonômica*, Vol. 3, 8-13.

Oppenheim, A. V.; Schafer, R. W. (1989). *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ.

Pikaar, R. N. (1990). Ergonomics in control room design. *Ergonomics*, Vol. 33, No. 5, 589-600.

Santos, I. J. A. L.; Teixeira, D. V.; Ferraz, F. T.; Carvalho, P. V. R. (2008). The use of a simulator to include human factors issues in the interface design of a nuclear power plant control room, *Journal of Loss Prevention in the Process Industries*, Vol. 21, No. 3, 227-238.

21

Risk Assessment in Accident Prevention Considering Uncertainty and Human Factor Influence

Katarína Zánická Hollá
University of Žilina in Žilina, Faculty of Special Engineering, Department of Crisis Management, Slovakia

1. Introduction

Nowadays, every manufacturing enterprise, company providing services or transport wants to be successful. Its goal is to secure prosperity of business and to achieve it through manufacturing or services. The technological processes and equipment are closely linked with the industrial risks which have become the object of assessing their decrease on the acceptable level and last but not least permanent monitoring of risks from the side of enterprises as well as selected bodies of the public administration. Perception and awareness of the need to prevent rising the crisis phenomena is for the society as well as legal entities or natural persons very important. Overlooking and insufficient attention paid to the risks could have negative impacts for all people. An important part of the risk assessment process is also inspecting and monitoring the adopted measures in connection with the effort to decrease their number.

As it was mentioned the risks connected with the technical and technological processes can be a source of unplanned interruption of the manufacturing processes or can violate providing a service and can cause material losses, damage the environment, threaten the health and lives of people. They can endanger not only the participating employees but also inhabitants from the surroundings and in the case of a leakage of dangerous substances it can become the source of violating the nature, the environment and endangering the inhabitants for a long-time period, if not forever. In the time period of increasing demands on the security of the technological processes the prevention is becoming the dominant idea and this article deals just with this area. It characterises the selected areas of prevention of industrial accidents which are the basic assumption for securing the safety of the industrial processes.

One of the objectives of this article is to show a possible approach to assessing the risks of the industrial processes in the form of structured diagrams. The logical procedure coincides with the currently used procedures created on the EU level as well as the national structures. Our attention will be aimed at especially at the area of the industrial accidents and the process of the risk assessment of the industrial processes in the Slovak Republic.

The uncertainty and its influence is significant in the process of the risk assessment and it is valid especially when assessing the risk from the quantitative point of view. The variance of the results is caused especially by different approaches to assessing and presenting the risk.

However, an extensive investigation of uncertainty is missing here. In the next part of the article I will deal with the influence of the uncertainty in assessing the risks and I will show the main sources of its presence. Several investigations have been realised worldwide which are to find out the main sources of the uncertainty influence when assessing the risks. (Amendola et al., 2002)
The last part of this article deals with the main cause of rising the industrial accidents, namely with people – the human factor. The human factor operates in an interaction with the industrial accidents and plays an important role in their rise; however, it is affected also by the consequences of their demonstrations. Here we can mention its three positions. The first one is the often analysed position of people as the human error that is the main cause of developing the industrial accidents. James Reason (1990) analyses it thoroughly in his book "Human Error". In his next book "Human Contribution" he describes also another position of the human factor – the hero who is able to prevent the rise of the industrial accidents. It is just the position of the human factor whose causes and characteristic features are not investigated in depth and therefore there is the space for further research. Last but not least they are the people – the human factor who is affected by the rise of the industrial accidents and which puts them to a position of a victim. The victims can be the so called direct victims (dead, injured people…) and indirect ones (the families of the missing or dead people).

2. Risk assessment in industrial accident prevention

The effort of people to achieve higher and higher standard of living reflects in the dynamic development of technologies, which are, on the other hand, still more and more complicated and can lead to the industrial accidents. The industrial accidents as the explosion in Flixborough (1974) or in Seveso (1976) as well as the disaster of the firm Union Carbide in Bhopal (1984) or in Chernobyl (1986) as also a whole range of others have shown the failure of the technology, its attendance. Due to these failures a great number of people died, or the consequences of these accidents caused durable health damage, not to speak about the losses of the material values and environment which can be of a long-term character or even irreversible.
The aim of the first part of this chapter is to describe the environment of the industrial accidents in the EU and in the Slovak Republic and to show an algorithm for assessing the risks through structured diagrams which should be mainly used in Slovak Republic. The model created complies with the currently valid legal regulations in the area of prevention against industrial accidents and is in balance with the used procedures for practical risk assessment in Slovakia as well as the EU.

Industrial accidents in the context of the EU and the Slovak Republic

The company SWISS RE works out an annual overview and summary of natural disasters and anthropogeneous disasters (man-made disasters) according to selected criteria (see the note). For working out an overview I decided to use selected indicators from 1998 to 2008 and to create the table 1 which depicts an overview of technological disasters in a chosen time period and is divided in three groups. The main group is created by the anthropogeneous disasters where belong all technological disasters caused by people and are designated in the table and the graphs as MMD (man-made disasters). This group comprises explosions, fires, air disasters, naval disasters, road and railway disasters, collapse of buildings and bridges, mining disasters and terrorism. The anthropogeneous

disasters where an explosion or fire arose or their combination create the second group and are designated as FED (fire and explosion disasters). There are disasters occurring in the industrial plants and warehouses, manufacturing processes working with petrol and gas, hotels and other buildings as well as remaining fires and explosions. The last group is created by accidents where an explosion, fire or leakage of hazardous substances in the industrial environment arose (factories and warehouses) are designated as ID (industrial disasters). In the table which is transformed into graphs I depict all three groups from the point of view of the overall number of the technological disasters, the number of victims and financial losses (million USD) during the time period of 1998 – 2008.

	Number of events				Number of victims		Financial losses (mil. USD)		
YEARS	MMD	FED	ID	MMD	FED	ID	MMD	FED	ID
1998	219	34	19	9788	1445	94	3534	1454	835
1999	188	36	20	7238	723	189	4140,3	2551	2107,7
2000	230	34	20	9694	1368	349	3049	1334	773
2001	204	40	17	10247	921	371	24381	3748	2086
2002	214	27	14	13066	2111	1562	2130	935	915
2003	238	36	15	7914	1071	139	2320	1137	905
2004	216	44	15	7275	1330	47	2889	1713	887
2005	248	60	31	8935	692	162	5066	4095	2346
2006	213	42	21	8677	906	185	4043	2110	1722
2007	193	34	15	6923	611	163	4295	2145	1170
2008	174	45	24	5618	454	159	7812	5255	2146
Total	2337	432	211	95375	11632	3420	63659,3	26477	15892,7

(Source: Swiss Re, 1998 – 2008)
Note: The company SWISS RE understands a disaster as an event when at least 20 people lose their lives, or the total amount of damages represents the sum of 72 million USD or the damages on property exceed 36 million USD

Table 1. Overview of selected anthropogeneous disasters according to number of events, number of victims and financial losses

The table shows that the disasters in the industrial environment create a relatively great part of the anthropogeneous disasters especially from the point of view of their number and financial losses. Their impacts on the employees and inhabitants from this point of view are not negligible which is proved by the numbers of victims in the individual categories. The financial losses caused by the largest disasters in the industrial environment create a relatively great proportion of the anthropogeneous disasters in the individual years.

Seveso II directive

The growth of the number of industrial disasters is the reason why new methods rise or the old ones are modified, i.e. the so called systematic procedures are developed which attempt to increase the security in the industrial enterprises. An example is the implementation of the SEVESO II directive in the framework of the EU as the basic pillar of preventing serious industrial disasters in the member states. Forming the directive began after the consequences of the large industrial disasters in the 1970s and 1980s when the EU in 1982 adopted a directive on serious industrial disasters. The EU called this first document "SEVESO Directive" – it got its name after the Italian town Seveso where after an explosion in a chemical factory dioxin leaked and caused a mass intoxication of the inhabitants. The prevention of the serious industrial disasters was later adapted by the Council Directive 96/82/EC on the control of major-accident hazards involving dangerous substances also called "SEVESO II" which is aimed not only at the prevention of large disasters but also at reducing their consequences for people and the environment.
Due to serious industrial disasters (breaking the dam of the sludge bed in the Rumanian town Baia Mare which caused intoxicating the river Tisza, the explosion of the pyrotechnics factory in the Dutch town Enschede, the explosion in the factory for producing fertilisers in the French town Touluse) a requirement for updating this directive arose. In 2003 the Council Directive 2003/102/EC was adopted. It formulates the environmental objectives of the EU as well as the decisive procedures in adopting measures for achieving these goals. The objects of this legal adaptation are specific duties of the operators and corresponding bodies concerning the enterprises where the selected hazardous chemical substances can be found. These issues are solved from the view of supervising the risk management of the possible serious industrial disasters. This law concerns companies of heavy chemistry, firms dealing with pressurised gases, equipment working with a higher amount of ammonia (firms using refrigerating equipment), petrochemical operations, but also companies with a higher supply of oil substances, etc. It does not concern the military premises, transport of hazardous substance by pipelines, mining activities, garbage dumps, etc.

EU study in the area of serious industrial disaster prevention

In 2008 the EU – Vri (The European Virtual Institute for Integrated Risk Management) realised a questionnaire study whose aim was to acquire information about the transposition of the requirements concerning the SEVESO II Directive in the individual member states and its general procedure, practical experience with making use of the weaknesses and problems connected with its practical implementation, effectiveness of its implementation and the impacts of the directive on the competitiveness of the European industry and subsequently to respond to these comments (to improve the directive). The target industrial sectors for processing the questionnaire were as follows: production of metals, explosives, petrochemistry, pesticides, pharmaceutical industry, basic chemical production, plastics and rubber, production of energy and its distribution, food industry and beverages. The questionnaire assessment brought conclusions and lessons necessary for a partial updating of the directive and preparing new accompanying documents. The selected conclusions from the research realised are as follows:

- the respondents have recognised a possibility to work out next accompanying documents in some areas – the area with the highest priority is the analysing and assessing the risks (risk assessment),

- a problematic area is the non-universality of the approach of the risk assessment, insufficient criteria for quantifying the risk and methods, tools and data for implementing these procedures,
- a lot of enterprises work out more a qualitative rather than a quantitative analysis which can conceal a higher level of the result uncertainty,
- the procedure for the risk assessment according to the SEVESO II Directive should be harmonised with the legal standards for the given area in the given country. /SALVI, O. et al

Similarly the responsible bodies in the area of serious industrial disaster prevention recommend proposing and creating the European database for supporting the risk assessment and working out the other documents. There exist some "guaranteed practices" for working out the analysis and risk assessment, however, in general it is necessary to create a clear and understandable procedure for processing documents and most respondents are missing such a document.

If new accompanying documents are created, the following issued should not be forgotten:

- the criteria of risk acceptability (impacts and probability),
- the assessment of the security measure management,
- the assessment of emergency planning,
- the calculation of the dangerous events´ consequences (explosion, fire, spreading a toxic substance),
- the methodology taking into account prevention and protecting measures,
- the methodology for assessing the domino effects.

The final EU recommendations in the area of the questionnaire assessment head to two levels:

- creating an accompanying document which will deal with what is to be done step by step and will explain how the directive requirements are to be interpreted,
- creating manuals for individual industrial sectors which would specify the environment for risk analyses and procedures necessary for its processing (SMEs).

Existing procedures for risk assessment

The environment of preventing the industrial disasters in the EU member states is affected by obligations which result for them from the membership in the international organisations. The individual EU countries implement the directives in their legal guidelines and create new procedures for the risk assessment which should contribute to harmonising in the area of the industrial disaster prevention.

There are several procedures for assessing the risks of the industrial processes. Systematic procedures, methods and techniques are used. The systematic procedures are structured operations which utilise selected methods and techniques in the individual steps. In the Slovak Republic the risk assessment also fulfils the requirements of the laws introduced in the Figure 1.

The risk assessment is part of the risk management. Its activity as well as expending resources for preventing the rise of serious industrial disasters is often pushed to background both by the wide lay public and professionals during a time period when no crisis phenomenon arises. However, when any technological disaster occurs, e.g. the accident which happened on 27th October 1995 in VSŽ, a.s. Košice – the leakage of CO, on 2nd March 2007 in Nováky - the explosion of the delaboration hall – both of them in Slovakia, then the losses of lives as well as material prove that a lot of tasks in this area are fulfilled

only in a formal way, their complex securing from the organisational, personnel, technical as well as material point of view is not solved. However, fulfilling these tasks is to be mutually harmonised and it is necessary to ensure them on a corresponding level.

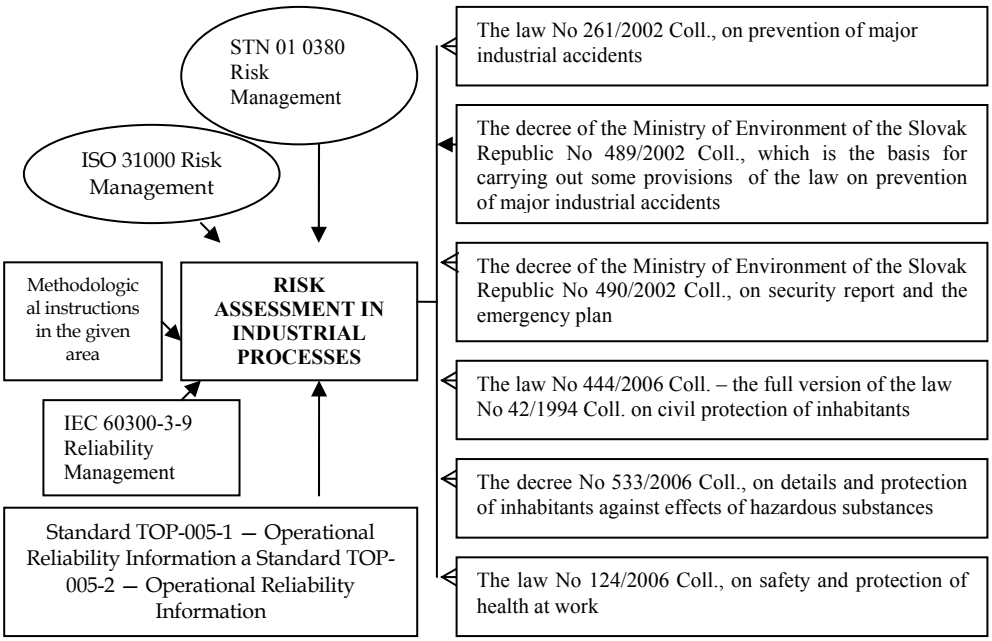


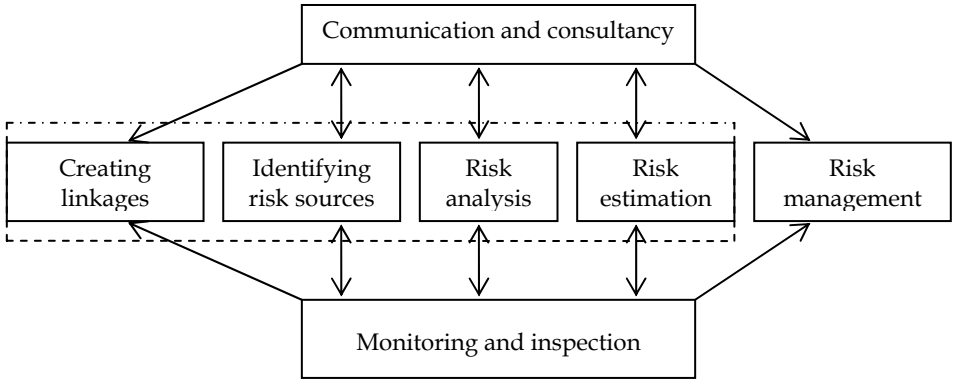
Fig. 1. Selected legal guidelines in the area of preventing the industrial disasters which require the risk assessment

Prevention in risk assessment

To avoid the industrial disasters, it is necessary to deal with prevention which is part of the crisis management model (prevention – preparedness – response – recovery). We utilise several procedures in the area of prevention whose main goal is to reduce the probability of the rise of the crisis phenomena or their negative impacts. One of these tools or more or less idea procedures or philosophy is the risk management, i.e. the process which is utilised not only on the microeconomic but also on the macroeconomic and global levels. Its procedures, methods and techniques contribute to reducing the probability of rising crisis phenomena and reducing their negative impacts which plays a positive role for the object assessed. It is implemented in different spheres of the social life and is applied in various forms in the practice. A consequent implementation of the risk management requires not only realising a thorough identification, analysis and risking assessment, their minimising by suitable procedures, but also a regular inspection of the measures realised.

In the Slovak Republic in the area of risk management the standard STN 01 0380 Risk Management is used, however, it has become outdated in several directions and the professional circles criticise it. If we wanted to identify the decisive phases of risk management we could realise it according to the standard ISO 31 000 Risk Management Guidance Standard. According

to it the process of risk management consists of the parts depicted in the figure 2. The risk assessment (outlined by an interrupted line in the figure 2) in this standard includes creating linkages, identifying the risk sources, risk analysis and evaluating the risk (risk estimation).



Source: ISO 31 000, 2009 – adapted.

Fig. 2. Risk management according to the standard ISO 31 000

The individual phases are in the accessible sources, legal norms and regulations, methodological manuals frequently introduced in different ways and this fact can cause misunderstandings in communication in the given area (a problem is often caused by a translation from a foreign language).

The risk assessment should be based on a systematic identification of the risk sources, on detecting what can be damaged, on creating scenarios in the form of trees of knowledge, trees of failures, and assessing the probabilities and their consequences. Expressing the risk should always comply with the mathematical formulation and represents a product of the probability and consequences. The consequences are determined in continuation to the rate of the threatened activities through calculations, and the probability either by a qualified estimation, or based on the historical experience. Quantitative risk analysis has its unique place in determining the level of adequacy of the security measures in the area of industrial process security. The quantitative criteria are, from the point of view of the level of subjectivity which enters the process, more credible than the qualitative ones.

Risk assessment is the core of risk management. After its realisation, the corrective measures for carrying out the stabilisation of the system and decreasing the risks can be stated. Both phases are burdened by subjective as well as objective factors which affect their overall result (uncertainty). The objective factors comprise defining the real quantities when assessing the risk quantitatively. In practice it is a problem to define the probability and consequences of an undesirable phenomenon because often the relevant data required for stating the risk is missing.

Existing procedures, methods and techniques for risk assessment

Assessing the risks in the industrial processes and their decreasing has a whole range of specifics whose recognising and accepting is very important for improving the level of the safety of the whole society and its continual progress. There are lots of models and methods for assessing the risks, however, most of them use a special terminology and specify the same facts in a different way.

In Slovak Republic there should be used these types of systematic approaches:

- PRA (Probabilistic risk analysis)
- ARAMIS (Accidental Risk Assessment Methodology for Industries)
- MOSAR and others

PRA is also called quantitative risk analysis (QRA) or probabilistic safety analysis (PSA) is widely applied to many sectors. In many of these areas PRA techniques have been adopted as a part of the regulatory framework by relevant authorities (so do in the Slovak Republic). In other areas the analysis PRA methodology is increasingly applied to validate claims for safety or to demonstrate the need for the further improvement. The trend in all areas is for PRA to support tools for management decision making, forming the new area of risk assessment. In the Slovak Republic the approach is worked out in the document "Methodological Procedure for Risk Assessment of Hazardous Operations and Study of Companies in the Slovak Republic" (Ministry of Environment of the Slovak Republic, Bratislava, 2000). The document shows the advantages of implementing the PRA (probabilistic risk analysis) compared to other methodologies as well as its broad implementation. The usage of induction and deduction methods described by it is emphasised. Next systematic approach is MOSAR which is a relatively new, systematic approach for analysing technical and technological risks developed in France. It can be used for analysing both a new and existing system. Two of its basic modules are known, namely Module A and Module B. The principle consists in realising a double analysis. In the first step the macroscopic view is searching for risks created by transmitting a danger (the so called risks of proximity) and this is solved by the Module A. In the second step the risks of individual sources are analysed, here we make use of the so called classical methods of the risk analyses (Module B). In the framework of the first step, i.e. the macroscopic view the so called black-boxes are used. The key when we use them is a simplified view at the considered system depicted as the black-box. The inputs are entered and concrete outputs are picked up. The way from the input to the system to the output from it is not determined in a greater detail.

The European approach ARAMIS is a less utilised method. It serves for the risk assessment in the industry and combines the strengths of determinism and acknowledged objective regularities. Its aim is to create a unified procedure for the risk assessment in all companies which belong to the group which has to fulfil the SEVESO II Directive with the possibility of the mutual comparison of the "companies´ danger rate" regardless to the fact to which industrial sector they belong. This methodology was optimised for the gas industry, specifically for the company NAFTA, a.s. The methodology's output is to determine the risk rate, suggesting suitable measures with a subsequent investment aim of the company in the area of increasing the operation security. The systematic procedure ARAMIS is recommended for implementation in the Slovak Republic. Currently only few companies in Slovakia use it for working out the risk assessment. A thorough depiction of the method is shown in the figure 3.

The following types of analyses affect the selection of the methods and procedures of the risk assessment in an industrial environment:

- *the a priori analysis* is based on the phenomenon which is the source of the risk and has occurred in the past at least once. The nature of the object assessed, the probable behaviour of the phenomenon is known and thus we can a priori forecast its behaviour and properties in the future;
- *the a posteriori analysis* is used when the analyst has to work with information, phenomena and events about which he/she thinks can develop, although they have not happened in the past. It means that the risk is estimated based on the assumed behaviour of the phenomena which develop after the analysis.

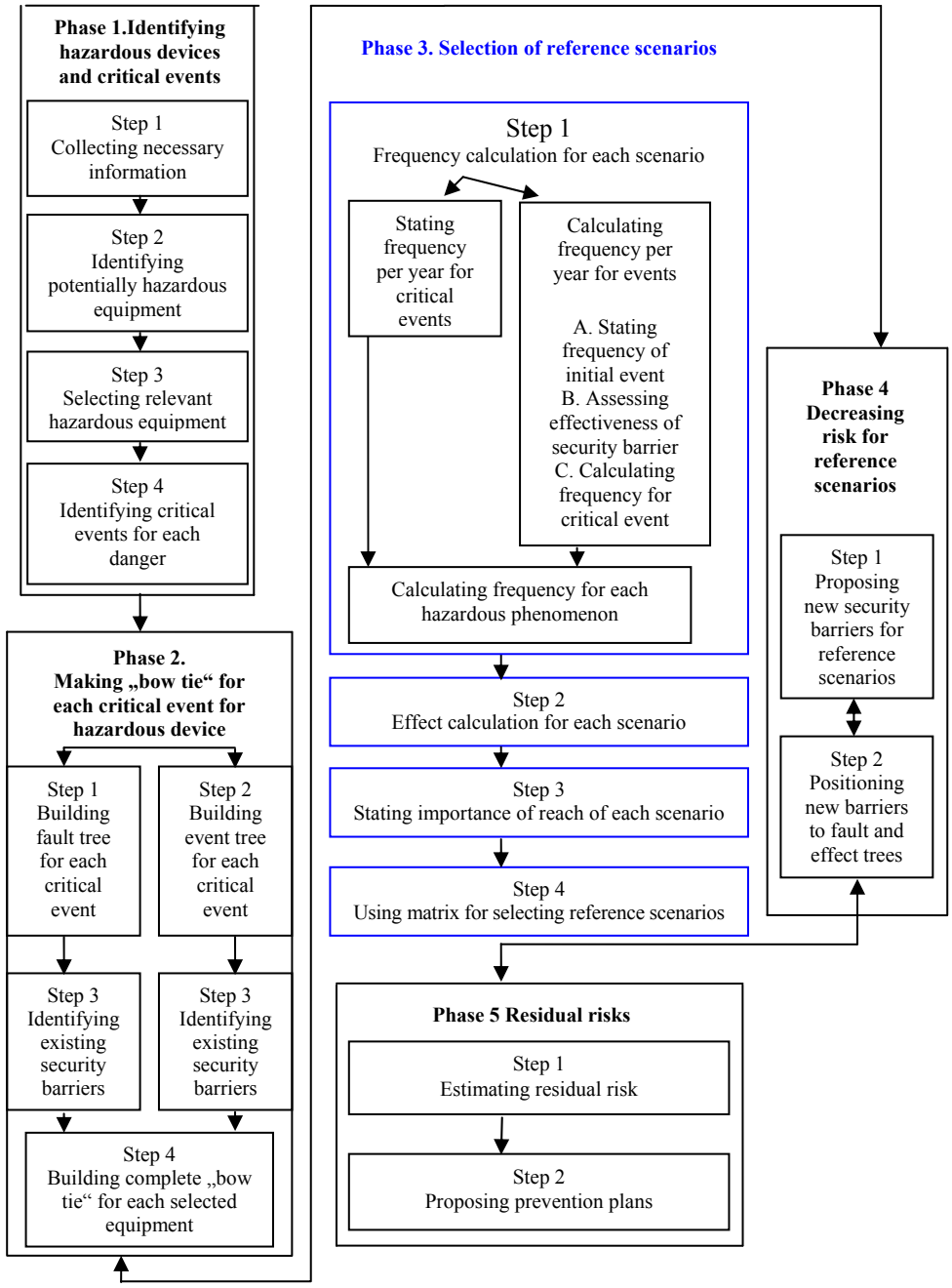


Fig. 3. Systematic approach ARAMIS (ARAMIS final user guide)

From the point of view of the inputs used and their character we distinguish:

- *the qualitative analysis* – is used for the qualitative estimation of the risk of a certain event, i.e. non-digital description consisting of identification and description of the risk sources, the relative verbal evaluation of the seriousness of the risk sources, identification, setting up and describing the accident scenarios;
- *the semi-quantitative analysis* – makes use of the semi-quantitative estimation of the risk of a certain event, i.e. the category of frequencies and effects and certain levels of seriousness are determined both verbally and quantitatively for the scenarios. The risk is stated similarly as in the qualitative risk analysis, however, the category of seriousness of the effects and scenario frequency are rendered more precisely;
- *the quantitative analysis* – a systematic procedure of numerical quantification of the expected number and effects of the potential accidents connected with the equipment or operation based on an engineering estimation, assessment and mathematical methods. (Paleček et al., 2000)

The decision about selecting the qualitative, semi-quantitative or quantitative analysis depends especially on the depths of the study and the purpose of the analysis realised.
The approach to the analysis from the point of view of stating the consequences and probabilities can be as follows:

- *the deterministic approach* – can be used if the problem formulated by one question or several questions can be answered clearly and understandably by one answer. The analysis itself is connected with a relatively simple determining of the causes, effects and impacts (by the relationships among them). We assume in the case of each problem it will have one result or one possible solution. It can happen that this approach does not result in any solution, i.e. there is no answer to the given question, or it cannot be answered. In this case only an approximate result is achieved. The uncertainty is not connected with a probabilistic result and is not easily detectable. When the effects which can develop are defined correctly we sometimes recognise the probability in the form of 100 % of the probabilistic occurrence or 0 % of the probabilistic occurrence (i.e. the phenomenon either develops or it does not);
- *the probabilistic approach* – is based on an assumption that several possible results of one assessed problem (situation) can develop. Probabilistic modelling aims at studying several results from the given data. The input data itself for the deterministic model cannot be used for a probabilistic study of the same problem. The probabilistic approach is currently preferred more. It is also recommended in the Slovak Republic for processing the analysis and risk assessment in the area of serious industrial accidents.

Model for assessing risks of industrial processes

Based on the previous information in the further text I characterise analyses affect the selection of the methods and procedures of the risk assessment. The subjects of investigating the model for the risk assessment are especially the technological processes in the industrial environment utilising hazardous substances. The systematic procedure created can form a supporting apparatus for analyses, especially in the SMEs. It is similarly usable for the analysis in the process of managing continuity in the operational company processes (*the business continuity management)* whose mission is to ensure the operation of all important processes inside the organisation if any unexpected events occur.

A systematic procedure serves the processors of the risk assessment of the technological processes with the presence of a hazardous substance for a better orientation in the given area as well as for approximating the fulfilment of the individual phases and will make the selection of methods and techniques for their application in the individual steps easier. The creation of a logical sequence of the phases and their steps according to which the analyst should proceed are emphasised. The phases of the risk assessment can be depicted by a simplified model which shows the involvement of the analysts, the responsible manager (decision-maker) and the working team to the overall process. The figure 4 shows the basic structure of the model of the risk assessment.

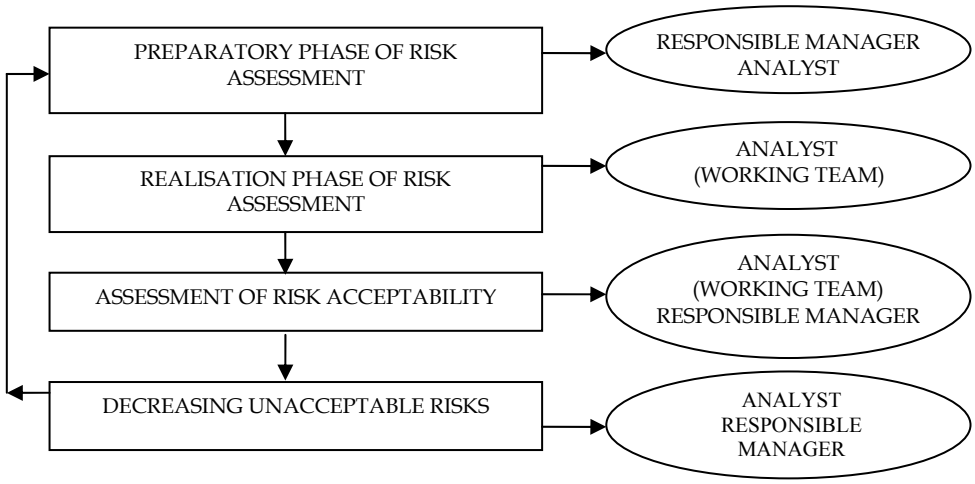


Fig. 4. Basic structure of the model of risk assessment

Further text explains the individual phases of the simplified model. As the first one, the preparatory phase of the risk assessment is characterised whose realisation is often underestimated or is not carried out correctly. The process of the risk assessment is implemented in the realisation phase and then the assessment of risk acceptability continues. Decreasing the risks is a decision which is realised on the basis of identifying unacceptable risks and subsequent work with them.

Preparatory phase of risk assessment

The preparatory phase of the risk assessment is followed by its implementation phase. In this part the risk analyst and the working group (if the decision is being made the presence of a responsible company manager is also necessary) are the most important players. The figure 5 depicts preparatory phase of risk assessment.
The figure 6 depicts the individual steps which create the realisation part of the risk assessment. Their interpretation as well as the content can differ in dependence on the

resources and type of the environment investigated as well as on the systematic approach used.

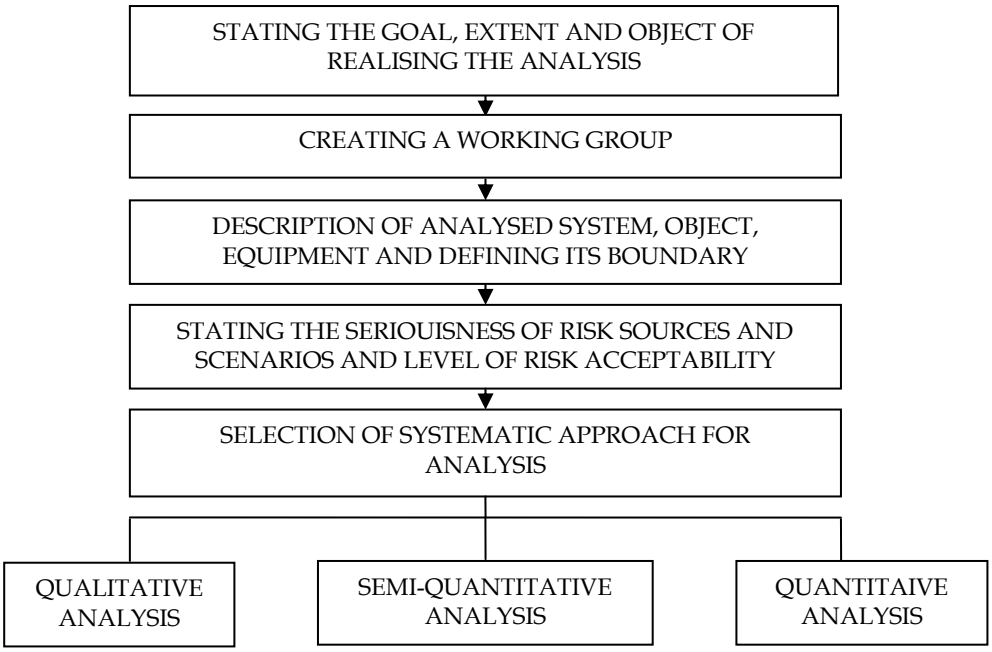


Fig. 5. Preparatory phase

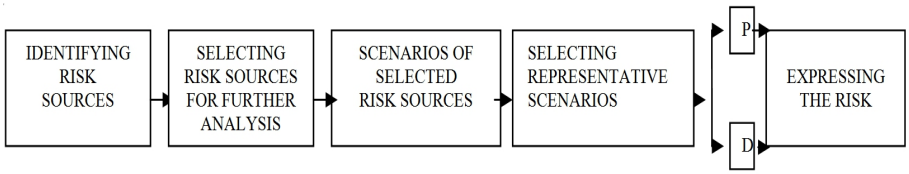


Fig. 6. Steps of implementation phase of risk assessment

Assessment of risk acceptability

The phase of stating the risk acceptability is important from the point of view of their further control. In most cases the criteria of acceptability are stated already in the preparatory phase of the risk assessment.

The decision about the acceptability, or unacceptability the risks is based on its two following levels:

- the negligible (acceptable) level of the risk – it represent a socially acceptable level of the risk in which the probability of occurrence of an adverse effect is small, the effects of its operation are moderate and the profit from the situation (the real or perceived one) is that large that the persons, groups or the whole society is willing to take the risk. It means that this level of risk does not require any regulation or other measures for its decrease neither from the point of view of people's health nor the protection of other live systems;
- the unacceptable level of the risk – requires inevitable taking of regulation measures or other specific measures for its decrease.

Every individual as well as every society has own values for the risk acceptability which are a compromise in many cases or sometimes a consensus reflecting its real "cultural", technological or operational maturity – in the technical practice often designated as the culture of operation. However, the term culture of operation comprises much more than the personal and technological security. It involves except for other things also the overall philosophy and approach of an individual or society to understanding the needs of the society.

3. Uncertainty in risk assessment

The second part of book chapter will talk about the uncertainty in risk assessment. It is known that results of any risk assessment are inevitably uncertain to some degree. Because of inevitable limitations of the risk assessment approach it must be acknowledged that the true risks could be higher or lower than estimated. In general, the word 'uncertainty' means that a number of different values can exist for a quantity, and 'risk' means the possibility of loss or gain as a result of uncertainties. The uncertainty should be divided into two categories: aleatory and epistemic. Aleatory stochastic uncertainty or due to randomness should result from bad knowledge of risk figures and their distribution, quantities such a failure rates, meteorological conditions at the time of release. Epistemic (reducible) is related to incomplete knowledge about phenomena of concern and inadequate matching available databases to the case under the assessment.

Besides, we know also the so called operational uncertainty. When comparing the physical models, the experience shows the importance of the human factor, e.g. using the same computer code by several specialists can lead to variations. The estimation variability of the commonly defined "representatives" of values expressing the risk and complexity of dangerous and main/temporary events which were identified by various experts from the teams, reflects the types of uncertainty, both operational and epistemic ones. If the values are defined as the "point assessment", in this case the variability is tied to an aleatory uncertainty. A different point assessment can be assumed for the main events or the parameters can be selected by an equal division.

Benchmark studies

EC´s Joint Research Centre in Ispra and RisØ National Laboratory were coordinators of projects that showed the acute presence of uncertainty when carrying out the risk assessments and emphasised the resources the uncertainty stems from and also the fact how it can decisively affect the final result of the analysis. In the first comparison study 7 teams carried out the risk analysis in a chemical factory at an undetermined place in Europe. Their results in spite of equal input data mutually differed which was caused especially by

utilising different methods and approaches. It was detected in the risk identification phase that the scenario assessment by probabilistic and deterministic approach can lead to fully different conclusions. The comparison study consisted of five main phases: the documentation phase, three working phases and the assessment (enlarging) phase. The working phases include the qualitative and quantitative phase – through study of the technological process mechanisms through case studies. The uncertainty is in this case bound to a lot of components, inspection mechanisms which are used in the technological process and interactions between them and the human factor. On the other hand we count on an uncertainty which is linked with meteorological and environmental conditions. The table 2 shows an example of a difference when stating uncertainties (6^{th} team chose deterministic approach).

S	Team 1	Team 2	Team 3	Team 4	Team 5	Team 7	Size Deviations
1	9.10^{-7}	1.10^{-6}	$1,4.10^{-3}$	9.10^{-7}	1.10^{-6}	$1,8.10^{-7}$	$1,8.10^{-7} - 1,4.10^{-5}$
2	1.10^{-5}	3.10^{-6}	$1,4.10^{-5}$	9.10^{-7}	$7,3.10^{-7}$	$4,6.10^{-7}$	$7,3.10^{-7} - 1,4.10^{-5}$
3	$4,8.10^{-4}$	$4,8.10^{-6}$	8.10^{-3}	5.10^{-7}	$5,4.10^{-7}$	$1,3.10^{-5}$	$4,8.10^{-6} - 8.10^{-3}$
4	1.10^{-6}	------	$4,6.10^{-6}$	9.10^{-7}	8.10^{-7}	$1,8.10^{-6}$	$8.10^{-7} - 4,6.10^{-6}$
5	$2,8.10^{-7}$	1.10^{-8}	$5,7.10^{-3}$	-----	$2,3.10^{-6}$	$4,9.10^{-6}$	$6,4.10^{-10} - 5,7.10^{-5}$
6	5.10^{-7}	1.10^{-8}	4.10^{-8}	-----	5.10^{-8}	5.10^{-7}	$1.10^{-8} - 5.10^{-7}$
7	6.10^{-7}	1.10^{-6}	5.10^{-6}	9.10^{-7}	4.10^{-7}	4.10^{-7}	$4.10^{-7} - 6.10^{-6}$
8	1.10^{-6}	5.10^{-7}	1.10^{-6}	$4,5.10^{-7}$	$1,3.10^{-5}$	4.10^{-7}	$4,5.10^{-7} - 1,3.10^{-5}$
9	3.10^{-6}	$3,4.10^{-7}$	$1,5.10^{-5}$	9.10^{-7}	$2,2.10^{-6}$	8.10^{-7}	$3,4.10^{-7} - 1,5.10^{-5}$
10	$2,4.10^{-6}$	$1,5.10^{-7}$	$2,1.10^{-3}$	$2,7.10^{-6}$	6.10^{-6}	5.10^{-7}	$1,5.10^{-7} - 2,1.10^{-3}$
11	$5,5.10^{-9}$	$1,5.10^{-9}$	$1,2.10^{-7}$	$1,2.10^{-7}$	$4,7.10^{-6}$	$1,4.10^{-7}$	$1,5.10^{-9} - 4,7.10^{-6}$

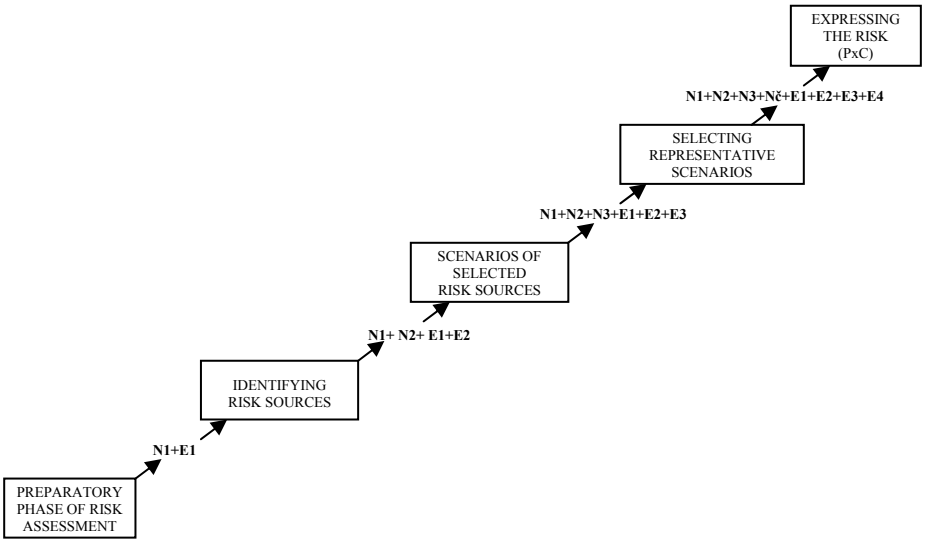
Table 2. Probability of „top events" of the individual 7 teams´ scenarios (Amendola, 2002)

The whole afore-mentioned procedure of assessing the consequences is full of uncertainties. In general we can thus say that there are two types of uncertainties: the uncertainty due to an incidental nature of the phenomena and uncertainties due to imperfect knowledge. The first type takes into account some phenomena and variables which incidentally change with time. The meteorological conditions can be such an example; it is impossible to determine with a 100 % certainty to forecast the direction and speed of the wind at a certain place of the space and at a certain time in the future, even if we knew exactly the conditions at present and in the past as well. The second type considers the lack of information which is presents at almost each step of the analysis. Our knowledge of phenomena following an unexpected leakage is not perfect and usually is based on empirical rules and observations of a limited number of accidents. The input parameters are also uncertain because exact conditions of accidents cannot be defined in advance. For an analyst to be able to cope with these uncertainties and insufficient knowledge, he/she usually has to state broad assumptions and to implement subjective judgement, i.e. an additional source of uncertainty

into the whole procedure. The result is then characterised as an output of assessing the consequences with the occurrence of a whole range of uncertainties. The analysts and the decision-making segment should be aware of these uncertainties connected with the results of the risk assessment and to take them into account in the case of the risk-oriented decisions. Some uncertainty sources can serve as an example:

- the meteorological conditions,
- the conditions in the closed equipment (e.g. pressure, the state of the substance, the quantity of the substance in the vessel at the time of damage),
- the size and dimension of the opening,
- the proportion of the removed liquid,
- the drops in the material that leaked,
- the presence of an initiation resource and the exact initiation time,
- the behaviour of the flying ruins,
- the vulnerability of the persons and buildings, etc. (Paleček, 2000)

One of the methods how the uncertainty can be reduced is the repeating of the calculations for all possible combinations of uncertain input values and all possible changes of the used models and to assign them the individual uncertainties. However, this results in rising of an unbelievable large number of scenarios. In this case we can orient on a few important variables, or to choose some representative categories, we pay attention to calculating a significant expected frequency, as well as a great number of scenarios can be analysed and assessed or in the end the Monte Carlo simulation can be implemented. In the framework of the uncertain variables in assessing the risks the main attention should be paid especially to the correlation among them.



N- uncertainties, E- error

Fig. 7. Cumulating of uncertainties in phases of risk assessment

On the figure 7 there is shown gradualness of risk analysis where in each phase there are partial uncertainty and partial error increasing to the final N and final E. Each phase is characterized by its own uncertainty and errors and input uncertainty and errors from previous phase. Finally we need to count not just with results of risk analysis but also estimate an uncertainty related to final figure.
As a part of it, risk assessment is inevitably uncertain to some degree. And there is a question how issues of uncertainty are dealt with in existing safety regulations and in existing standards for risk analysis and management. I want to point on fact that there is a big need to deal with uncertainty and to count with it in risk assessment. Benchmark studies could serve as a guide to areas where caution must be taken when performing risk analysis.

4. Human factor influence on accident occurrence and demonstration

The last part will point out the problem which is very important to talk about. This is also the crucial part of crisis events occurrence and arising – the human factor. The aim of this part will be to show the human factor and his contribution to crisis events occurrence. The human factor will be assessed from two points of view as a hazard component which cause industrial accidents occurrence by errors and human as a hero element whose adaptations and compensations have brought troubled systems back from the bring of disaster.
In the past models of accidents dealing with the causes and relationships of accident rise were created. They insubstantially emphasized the human factor, it was only introduced as an immediate cause of events leading to an accident. Currently there is an effort to understand why and when the human factor affects the rise and development of serious accidents (it is the cause or part of accidents). What makes it possible to forecast, to prevent accidents as well as to decrease the share of the human factor on the rise and development of serious disasters? (Feyer, 2010)
The analysis of events which occurred and were caused by the human factor is one of the methods for creating the preventive measures. According to this method it is possible to foresee partially the human behavior in the crisis situations.
Over the past 50 years has been a dramatic widening of the scope of accidents investigation across many different hazardous domains:

- system and cultural issues (1960s Metal fatigue, Aberfan Inbrox)
- unsafe acts (errors and violations) (1970s Flixbourgh, Seveso, Tenerife TMI MT Erebus)
- equipment failures (hardware – software) (1980s Chernobyl Zeebrugge Bhopal PiperAlpha Dryden, 1990s Paddington Long Island Alabama Eschede, 2000s Linate Uberlingen Columbia). (Holla &Moricova, 2010)

Chemical incident statistics are very sketchy with respect to root causes and many reported incidents do not furnish much detail about the cause. Chemical safety and hazard investigation board published in 600K Report that:

- Among cases where the cause was known, 49% were as a result of mechanical factors, 39% from human factors and just 2% to weather-related phenomena, 10% causes not found,
- Among cases involving mechanical factors, an overwhelming 97% were attributed to general equipment failure; 63% of human factors cases were attributed to human error. (Garcia, 2002)

The high rate of general equipment failure among reported incidents suggests that mechanical integrity/maintenance issues are significant and from the human error that training and proper procedures should also be examined.

There should be introduced instances of accidents which were caused by failing the human factor or saving lives by human factor. The first of them is the *Chernobyl disaster.* An industrial accident of exceptional size had a lot of victims that cannot be counted exactly (the epidemiological analysis is not available). Various scientific studies assume from 9,000 to 475, 000 victims. The most frequent conclusions and maybe the most probable values are in several tens of thousands (30,000 to 60,000). The 1986 Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident (INSAG-1) of the International Atomic Energy Agency's (IAEA's) International Nuclear Safety Advisory Group accepted the view of the Soviet experts that "the accident was caused by a remarkable range of human errors and violations of operating rules in combination with specific reactor features which compounded and amplified the effects of the errors and led to the reactivity excursion." In particular, according to the INSAG-1 report: "The operators deliberately and in violation of rules withdrew most control and safety rods from the core and switched off some important safety systems."

Another example of the human factor failure in the environment of the nuclear power stations is the disaster *Three miles island* which happened at 4 am on 28th March 1979 and where the second nuclear reactor was partially melted. The operational building was contaminated and an extensive leakage of radioactivity to the environment also occurred. The investigation commission later designated for the reason of the accident a breakdown of the safety valve. The proportion of the human factor was that operators were unable to diagnose or respond properly to the unplanned automatic shutdown of the reactor. Deficient control room instrumentation and inadequate emergency response training proved to be root causes of the accident.

Last example is connected to another type of accident - nearly accident. As an example we can introduce the *Apollo 13* programme. Its objective was the third landing of the human crew on the Moon surface, this time in the area of Fra Mauro. The typical sentence: "Houston, we've had a problem," says how very close the crew was to a disaster. During the flight one of the oxygen tanks exploded and seriously damaged the service module. The consequences of this explosion were serious. Not only this situation caused the crew did not fulfill the task of this flight but it threatened the lives of the crew members. The Manned Spacecraft Centre (today Lyndon B. Johnson Centre) had to develop with an extreme effort emergency scenarios thanks to which they succeeded in transporting the crew alive back to the Earth. Hundred of people were involved in the rescue: off – duty controllers, astronauts, simulation technicians, contractors' personnel and many more. But this case is only to show how the team effort, and a magnificent display or sheer unadulterated professionalism, both in he spacecraft and on the ground brought the crew to the Earth alive. (Reason, 2010)

There is a stark contrast between unsafe acts and these intrepid recoveries. Errors and violations are commonplace, banal ever, they are as much as a part of human condition as another ordinary human activities. Successful recoveries, on the other hand, are singular and remarkable events.

The human factor in relation to the rise and demonstrations of the industrial accidents can play several roles. These roles are as follows:

- the human factor as the cause of the rise of the industrial accidents (hazard - human error),
- the human factor as the recipient of the negative consequences of the industrial accidents (victim - negative impact),
- the human factor as a hero or anticrisis factor (hero - heroic recoveries).

Human factor as the cause of the rise of industrial accidents

When the human factor fails, there is a whole chain of small errors which if occurred individually they would not have fatal consequences. However, from a certain point on the tragedy is unavoidable.

There are several definitions of the human error. One of them says that the error is an action or a decision which was not determined (planned) and which leads to undesirable result. Furthermore, the human error defines a certain fact, statement or decision which deviates from the standard and the result is an actual or potential unfavourable event. However, this event can but also need not lead to an unfavourable result.

There are several possible definitions and there are also many ways in which errors can be classified. When we are talking about deviations concerning the human error we should mention such deviations that could be from upright (trip or stumble), from the current intention (slip or lapse), from an appropriate route towards some goal (mistake), or in some circles, it could even involve straying from the path of righteousness (sin).Human error classification should be done based on possible generic classification based on action: omission, intrusions, repetitions, wrong objects, disordering, mistiming, blends etc.

In the industrial processes there are the following possible causes of errors and failure of the human factor: bad reflection of risks of the attendants; errors in communications; insufficient or incorrect knowledgeability of the employees, insufficient qualification, insufficient experience (lack of training) – practice, personality and health assumptions of the employees; failing to keep the working procedure; unsuitable working conditions and working environment; inattentiveness (momentary) of the employees and many others. (Malý, 2002)

Human factor as hero (intrepid recoveries)

Another perspective according to human factor, one that has been relatively little studied in its own right is human factor as a hero. This presents a human factor as an element whose adaptation and compensation have brought trouble systems back from the brink of disaster on a significant number of occasions. We have already presented an example Apollo 13 where human factor saved several lives of astronaut. Other examples to be mentioned concerned to intrepid recoveries are connected to aeroplane crashes for example British airways flight 09 from London Heathrow to Aucland then BAC 1 – 11 flight to Malaga and many others.

Reason (2010) presents:" I find the heroic recoveries of much greater interest and in the long run, potentially more beneficial to the pursuit of improved safety in dangerous situations (operations)."

Human factor as recipient of negative consequences of industrial disasters

As already mentioned people are in many cases the reason for rising industrial accidents and they also significantly affect their development. However, on the other hand people are also affected by them, tangibly by their negative consequences. The accidents affect the people – their lives, health, property but also the environment in dependence on the concrete form of the accident. The impacts on people can be divided into two groups,

namely the impacts on the employees working in the company and impacts on the non-employees (the general public). The impacts of the industrial disasters on the employees according to their levels can be: death of the employee; serious damage of health with permanent consequences; serious industrial accident; light industrial accident; dangerous event (almost an accident); stress resulting from the situation arisen. (Zanicka Holla et.al, 2010)

Several scientific disciplines participate in solving the area of the human factor. They are especially disciplines as psychology, ergonomics, physiology, cybernetics, anthropology, hygiene, medicine, sociology and others. The human being as part of the working system is the most flexible, adaptable and valuable element, however, the most predisposed to making errors. An important role of the scientific disciplines which deal with the area of the human factor is to solve practical tasks in the real life, to increase the security, effectiveness and work comfort.

There is nothing in the people's history that would have prepared the human being for mastering the environment of the most modern technique, although we have adapted this technique to our capabilities and limitations. However, the technique is not sufficiently adapted to our psychical properties. In the field of the crisis management the area of the human factor is a cross-sectional area and therefore it is necessary to pay it increased attention.

5. Conclusion

The object of this article was the area of preventing the industrial accidents with an emphasis on the process of the industrial processes risk assessment, the influence of uncertainty on the results of the realised analysis and last but not least the position of the human factor in the process of the rise and operation of the industrial disasters´ effects. In Europe for the time being there are discussions concerning the utilisation of the same procedures, methods and techniques in the area of preventing the industrial disasters by the member states. This unification can bring positives but also negatives. One of the positives is the possibility to compare the results among individual companies and in this way to assess the level of their danger in the European context; however, this would be only possible in the area of serious industrial disasters, i.e. for the companies controlled by the SEVESO II Directive. The systematic procedure ARAMIS has been created and it is to serve these purposes, however, only a few countries are making use of it. The EU requirement also heads to utilising especially the quantitative approaches in regard to reducing the uncertainty rate in the analyses.

A problem could be also the variance of approaches used by individual countries, selecting the probabilistic or deterministic approach of stating the risk, the a priori or a posteriori approach, the qualitative, semi-qualitative approach. The selection of the procedure depends especially on the size of the company assessed, the pre-disposition of the employees (the educational and personality one) who carry out these analyses, the financial possibilities of the company or institutions and many others.

In my opinion the common approach which will work on the quantitative calculations can be selected only for the so called SEVESO companies which are monitored by the EU and have to work out these analyses based on the legal requirements. For other companies (but for the SEVESO firms as well) it is possible to state at least a structured approach. The structured approach should state how it is to proceed when assessing the risks phase by

phase and subsequently step by step in the framework of the individual phases. The auditor would choose the individual methods based on the criteria for the risk assessment. The utilised methods should be, in my opinion, at least semi-quantitative and of course, the quantitative methods should be preferred.
Another challenge for solving this area is to create a risk matrix which would be able to compare the quantitative expression of the risk components of several objects (loss of life, damaging health, damaging property, and environment). In such a case we would come to the issue of calculating the price of the human life by financial means which is today considered as non-ethical and impossible by many experts. Another problem is the presence of uncertainty in the risk analysis which causes deviations in the analysis results. It is necessary to identify the critical places in the analysis for the influence of uncertainty to be reduced as much as possible. In the Slovak Republic we are missing the investigation of uncertainty and due to this fact research and searching for critical places of uncertainty specific for Slovakia due to several differences compared with other countries in this region could be realised.
However, we must not forget that the human factor is the weakest segment in this process. According to several investigations and analyses the human factor is the most frequent cause of the rise of the industrial disasters. The analysis of the human reliability should create an integral part of the risk assessment. It would be suitable to create a methodological instruction for processing the analysis of the human factor reliability which is missing in Slovakia for the time being. Creating some space for a further investigation in the area of the human factor I see especially in researching the specifics of surviving and behaving the human factor (personality) in three positions identified.

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0043-10.

6. References

Amendola, A. et al. (2002) M: *Assessment of Uncertainties in Risk Analysis of Chemical Establishment*, The Assurance report, Final summary report, Denmark

Aramis final user guide.[on line]. [cit.2011-11-6]. Available on http://mahb.jrc.it/index.php?id=447

Aven, T (2002). *Foundations of Risk analysis*. Norway: John Wiley & Sons, ISBN 0- 471 – 49548 – 4.

Bedford, T. & Cooke, R (2001).: *Probabilistic Risk Analysis (Foundations and .methods).* Cambridge university press. ISBN 978 – 0 – 521 – 77320- 1, Cambridge, Great Britain

Bell, J. & Holroyd, J. (2009). *Review of human reliability assessment methods.* [on line]. Norwich: First published. [cit. 2010-11-2]. Available on: http://www.hse.gov.uk/research/rrpdf/rr679.pdf

Berry, L. M. (2009). *Psychology at work*. Ikar, ISBN 978-80-551-1842-0, Bratislava

Dzvoník, O., Kríž, J. & Blaško, P. (2001). Human *factor in flying. Human effectiveness and its limits.* : EDIS – ŽU, ISBN 80-7100-811-7, Žilina

Feyer, A. M., & Williamson, A. M. (2010). *Human Factors in Accident Modelling*. [on line]. [cit. 2010-11-13]. Available on: http://www.ilo.org/safework_bookshelf/english?content&nd=857170643

Fotr, J. & Švecová, L.(2006). Risk and uncertainty in strategic decisions.. In: Political Economy. Economic University. .ISSN 0032-3233. Prague, Czech Republic

Garcia, D., (2002). *The Debate About Chemical Accidents: Where Do We Stand?* [on line]. [cit. 2010-11-12]. Available on: http://www.acusafe.com/Newsletter/Stories/0800News-ChemicalAccidentsDebate.htm

Kopecký, Z. (2005). *The use of system access in safeguard of Business Continuity Management*. In: Economics, finance, company management, Economic University, Bratislava

Loveček, T., Kampová, K. (2009) *Application of quantitative methods in protection of strategic subjects* , In: Varstvoslovje: Journal of criminal justice and security, Vol. 11, no. 4, ISSN 1580-0253.

Malý, S. (2002). *Human factor in safety documentation according to act no. 353/1999 Sb. about major accidents prevention.* [on line]. [cit. 2010-11-10]. Available on: http://www.bozpinfo.cz/utf/knihovna bozp/citarna/clanky/lidsky_cinitel/lc020308.html, VÚBP Prague, Czech Republic

Merna,T. AL & Thani, F.F. (2007). *Risk management*. Computer Press, ISBN 978-80-251-1547-3, Brno, Czech Republic.

Míka, V., Šimák, L., Hudáková, M. & Horáček, J. (2009). *Management and crisis management*. EDIS, ISBN 978-80-554-0079-2, Žilina, Slovak Republic

Moricová, V. 2011. Stressful situations and its influence on crisis manager. In: *Zborník z konferencie „Riešenie krízových situácií v špecifickom prostredí", 1. – 2. 6. 2011 v Žiline.* Žilina: EDIS – vydavateľstvo ŽU, 2011. ISBN 978-80-554-0365-6, s. 475 – 482.

Hollá, K., & Moricová, V. (2010). Risk assessment of human factor in industrial processes. In Zborník z konferencie „Riešenie krízových situácií v špecifickom prostredí" on 2. – 3. 6.2010 v Žiline. Žilina : EDIS – vydavateľstvo ŽU, 2010. ISBN 978-80-554-0202-4, s. 221-227, Slovak Republic

Paleček, M. a kol (2000). Procedures and Methodologies Of Analyses and Risk Assessments for Purpose of Law No 353/1999 Coll., on Prevention of Major Accidents. Praha: VÚBP, Czech Republic.

Paleček, M. a kol. (2006) *Risk Prevention*. VŠE, ISBN 80-245-1117-7, Prague, Czech Republic

Reason, J. (2010). *The human contribution*, MPG Books Ltd, ISBN 978-0-7546-7402-3, Burlington, United states.

Reason, J. (1990). *Human error*. Cambridge University Press, ISBN 978-0-521-31419-0, Cambridge, Great Britain

Salvi, O. et al(2008) : *F – SEVESO, Study of the effectiveness of the Seveso II directive*, Brussels: EU–Vri

Simak, Ladislav and Ristvej, Jozef (2009) "The Present Status of Creating the Security System of the Slovak Republic after Entering the European Union," Journal of Homeland Security and Emergency Management: Vol. 6 : Iss. 1, Article 20, ISSN: 1547-7355. DOI: 10.2202/1547-7355.1443. Available at: http://www.bepress.com/jhsem/vol6/iss1/20

Sluka, V. – Bumba, J: VÚBP, *Major Industrial accidents prevention* , Personal consultation, (11.11.2008)

Tichý, M.: Risk controlling: analysis and management, C.H.Beck, 2006. ISBN 978 – 80 – 7179 – 415 – 5, Prague, Czech Republic

Vose, D.: *Risk Analysis – A Quantitative Guide*, third edition (2008), John Wiley & Sons Inc., 2008, ISBN 978-0-470-51284-5.

Zánická Hollá, K., Ristvej, J., Šimák, L. (2010). *Risk assessment in industrial processes*. Iura Edition, spol. s. r. o., ISBN 978-80-8078-344-0, Bratislava, Slovak Republic