Early detection and identification of anomalies in chemical regime based on computational intelligence techniques

FIGEDY Stefan[1], and SMIESKO Ivan[2]

1.Department of Diagnostic Systems, VUJE,a.s., Okruzna 5, Trnava 91864, Slovak Rep. (figedy@vuje.sk)

2.Chemistry Unit, Slovenské elektrárne, a. s., subsidiary of ENEL, Bohunice Nuclear Power Plant, 919 31 Jaslovské Bohunice Slovak Republic (ivan.smiesko@enel.com)

Abstract: This article provides brief information about the fundamental features of a newly-developed diagnostic system for early detection and identification of anomalies being generated in water chemistry regime of the primary and secondary circuit of the VVER-440 reactor. This system, which is called SACHER (System of Analysis of CHEmical Regime), was installed within the major modernization project at the NPP-V2 Bohunice in the Slovak Republic. The SACHER system has been fully developed on MATLAB environment. It is based on computational intelligence techniques and inserts various elements of intelligent data processing modules for clustering, diagnosing, future prediction, signal validation, *etc*, into the overall chemical information system. The application of SACHER would essentially assist chemists to identify the current situation regarding anomalies being generated in the primary and secondary circuit water chemistry. This system is to be used for diagnostics and data handling, however it is not intended to fully replace the presence of experienced chemists to decide upon corrective actions.

Keyword: computational intelligence; early fault detection; chemical regime diagnostics

1 Introduction

In nuclear power plants, water is used in primary circuits, secondary circuits as well as auxiliary systems. Water becomes an aggressive medium especially at high temperature when in contact with structural materials. Hence the reliability of many nuclear power plant systems is dependent on water chemistry during normal operations, startups, shutdowns and abnormal operation cases. In water-cooled power reactor, some undesirable effects may occur even under normal operating conditions, such as corrosion, erosion or deposition of corrosion products and other insoluble substances on heat transfer surfaces. Chemistry control in nuclear reactors is important at least from four different perspectives, namely integrity of barriers, plant radiation levels, deposit buildup and safety [1] .

Chemistry and radiochemistry data measured with conventional or advanced methods and sensors, as well as plant data (such as temperatures, mass flow rates, and other thermal-hydraulic and operational data) are normally collected with data acquisition and diagnostic system. For this measurement, on-line monitoring is preferred. Nevertheless, the off-line sampled data serve as supplemental data, which may positively or negatively influence the credibility of diagnostics in substantial manner.

2 Chemical information system

Originally there were four units operated at Bohunice Nuclear Power Plant in the Slovak Republic. However, currently only the 3rd and 4th units are operating, because the first two units have been closed as the requirements for Slovakia's accession to European Union. The Bohunice Nuclear Power Plant V2 (NPP-V2), which is the newer type of Soviet designed VVER-440 pressurized water reactor, has a water chemistry monitoring system called SYMOCHER (SYstem for MOnitoring of CHEmical Regimes) together with CHEMIS (CHEMical Information System) as a part of LIMS (Laboratory Information Management System). It is a distributed complex information system for performing, monitoring and control of chemical aspects of NPP operation processes. The system is based on combination of Linux and WinXP ORACLE database-platform. The main features of this system include:

- collecting and archiving raw or pre-processed information on continuous operational measurements within the technology,

Received date: January 24, 2012
(Revised date: March 19, 2012)

- automatic control of chemical injection into secondary circuit,
- support for sample collection and chemical analyses according to good laboratory practices,
- generating schedules of laboratory analysis, monitoring the degree of analyses completion,
- elementary statistical processing of measured data,
- checking and alerting for the occurrence of events such as exceeded limits of monitored parameters, including in the context of user-adjustable relations of these parameters,
- preparing gathered or measured data for verification and analyses,
- presenting collected data in graphs or tables and further processing the collected data to show the plant status by easier way on the interactive displays,
- calculation of pre-defined parameters and investigation of mutual parameter interdependencies,
- quality assurance modules (*e.g.* calibrations and quality control charts), and
- generation of protocols to export data to various formats.

The system is a very powerful tool for the management of plant chemistry and associated activities. Nevertheless, the knowledge of experienced plant chemists is still necessary for effective utilization of the system. In order to maintain plants' ability to effectively respond to the existing or possible new anomalies, introduction of support function by higher level information and technology (IT) than at present is necessary for the future challenges along with plant operation optimization and the related staff reduction with the reflection of knowledge management aspects in mind. Thus, a computational intelligence-based application package is a promising way to achieve this future direction.

3 Intelligent extension

Extensive measures have been implemented within the modernization project, such as the improvement of power plant's nuclear safety, seismic resistance, and fire safety, as well as the increase of operation reliability and availability of the Bohunice V2 units. A new safety control system TELEPERM XS was commissioned in both the 3rd and 4th units. The replacement of section and auxiliary switchboards had been completed. The replacement and extension of plant information systems had also been completed.

The following goals had been achieved through the completion of the Bohunice V2 upgrade program: increased level of nuclear safety in accordance with international standards for nuclear power plants of the same type; increased operation reliability of units; establishment of conditions for up-rating of units to 107% of the original rated power and prolongation of the power plant lifetime.

The availability of prompt information about the chemical conditions of the primary and secondary circuit is very important to prevent the undue corrosion and fouling build-up. For this reason, VUJE Inc. has been developing a new diagnostic system for early detection and identification of anomalies being generated in water chemistry regime. This system is called SACHER (System of Analysis of CHEmical Regime) and was installed within the major modernization project at the NPP-V2 Bohunice, units 3 and 4, as supplemental extension to the existing chemical information system CHEMIS. Both SACHER and CHEMIS systems are common for units 3 and 4 and are located in the water chemistry laboratories, which are also common for both units.

The typical chemical information systems that exist in operation at the NPPs provide users with values of the measured parameters together with their time trends and other derived values. Further to that, the roles of experienced users are required in order to identify the situation of the monitored process, to make the subsequent decisions, and to take upon the appropriate measures. The SACHER system, based on the computational intelligence techniques, inserts various elements of intelligent data processing modules for clustering, diagnosing, future prediction, signal validation, *etc*, into the overall chemical information system. The relation of CHEMIS and SACHER is shown in Fig. 1.

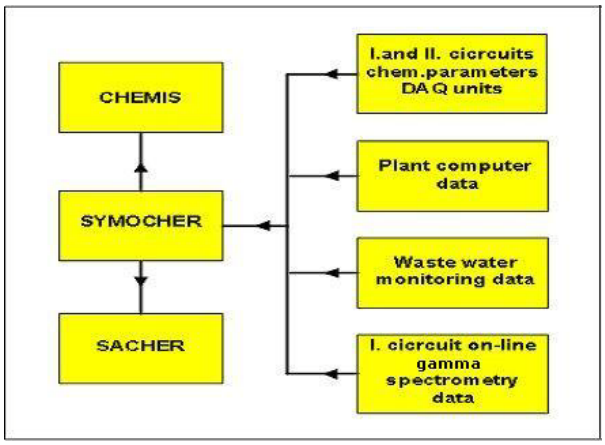

Fig.1 The structural diagram of water chemistry monitoring and diagnostic system at NPP-V2 Bohunice.

Although both systems receive the data from the plant computer through the chemistry monitoring system SYMOCHER, the SACHER system functions independently from CHEMIS. Such a configuration allows easier implementation of SACHER in the other Slovakian NPP Mochovce in the future.

SACHER has been fully developed on the MATLAB environment and has the modular structure described in the following sections (see also Fig.2).

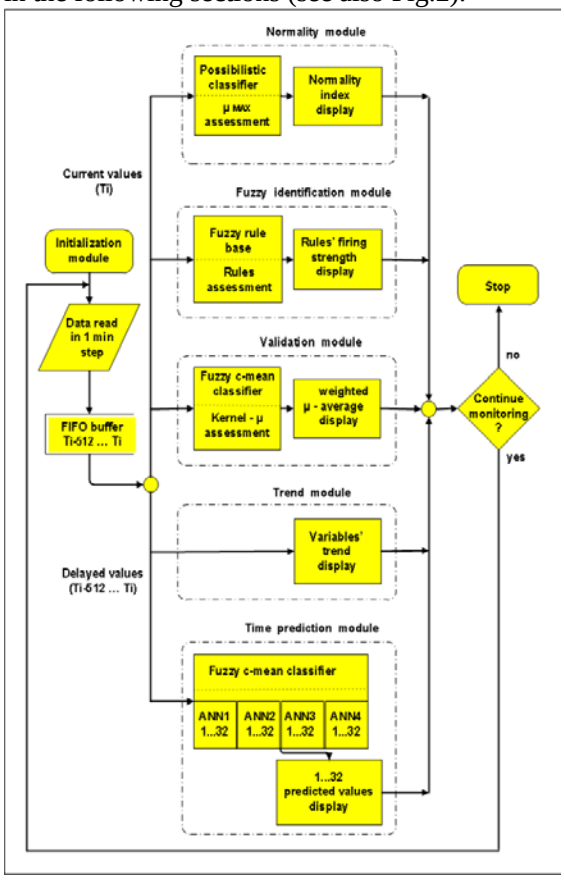

Fig.2 The functional block diagram of SACHER water chemistry diagnostic system.

3.1 Initialization module

The initialization module serves to launch the continuous surveillance of the chemical regime of both the 3rd and 4th units concurrently. Prior to that, the user of the system manually initializes the values of a few chemical parameters and the connection status of some plant systems which are not available on-line. The module also enables the adjustment of these parameters on the run whenever such requirement arises.

There are approximately 130 process and chemistry parameters that are inputted into SACHER with the sampling frequency of 1 minute. They are placed in a FIFO buffer with 512 values for each parameter. The normality, fuzzy identification, and validation modules process only the current data values, while the time prediction and trend modules make use of the delayed data snapshots from the buffer. The size of a buffer, which is currently set as 8.5 hours, is going to be increased substantially in the next upgrade of the system. The low sampling rate of 1 min is taken due to the slow changes of chemical processes. Another reason for the selection of the low sampling frequency was the expected high computing requirements which the standard personal computer may not be able to meet within shorter sampling interval. This limitation naturally becomes insignificant when the hardware and software upgrade will be further made in future.

3.2 Normality module

The objective of normality module is to recognize whether or not the situation of the process starts to deviate from the normal one. This module is based on the possibilistic fuzzy clustering algorithm[2],[3] implemented into the neuro-fuzzy process signal validation system PEANO[4] developed by the OECD Halden Reactor Project. It is a joint undertaking of national organizations in 18 countries including the Slovak Republic.

The clustering algorithm must be able to generate the representative clusters to which the patterns of process parameter values belong. As the pattern vectors may have the characteristics of several classes, the classification must assign any single pattern to the representative clusters through the degree of

membership. Otherwise, the pattern must be discarded if it is not represented by any cluster. Another requirement on the clustering algorithm is the smooth transition between the clusters as the situation represented by process signals evolves due to power maneuvers or transients.

Fuzzy clustering reflects the probabilistic requirement that the total probability for an input data set pattern that belongs to any cluster is 1. It implies the patterns that are not reflecting any of the identified cluster prototypes are classified and assigned to the clusters, merely because of the implicit certainty that all of the patterns belong to the established partition. There can be uncertainty (or fuzziness) on where the incoming pattern could be assigned; otherwise no uncertainty exists if the incoming pattern can be assigned somewhere. A number of issues may arise, such as:
- lack of robustness against noisy data.
- The inability to provide the statement: "I do not know", although it might be the best answer for a given situation. An incoming pattern might be given a high grade of membership in a cluster, even if it is far away from all the centroids, only because it is relatively closer to one specific cluster.

Relaxation of the requirement imposed on the established cluster partition leads to a possibilistic approach. For this reason, the possibilistic fuzzy clustering approach has been used in this module. A possibilistic classifier initially learns a dataset X of pattern samples only from normal situations. In other words it calculates the cluster prototypes and the corresponding membership grades of pattern samples.

The step-by-step procedure used to develop the fuzzy and possibilistic classifiers can be summarized as follows:
- Given a set X of only normal situations samples, compute an initial set of cluster centroids using the crisp ISODATA[5] algorithm that has been chosen because it automatically optimizes the number of required clusters.
- Initialize the elements of the partition matrix U with crisp values (0 or 1), using ISODATA. Subsequently, run the Gustafson-Kessel[2] algorithm, which produces the fuzzy classifier.
- Use the updated fuzzy partition from the previous step to start the iterative process of the Krishnapuram-Keller[3] algorithm to arrive at a possibilistic partition.

During this process, the model increases its robustness to noisy data and many patterns in X could be discarded as not representative of any developing cluster. As new patterns are examined during the monitoring phase, the possibilistic model evaluates in which cluster or clusters the incoming pattern could possibly be assigned, if any. It means that patterns that do not reflect any of the identified cluster prototypes, *i.e.* do not belong to the clusters of normal situations, are discarded as unknown if they do not fit into any cluster. The low degree of membership to all normal situations clusters, which represents a normality index, is the indication of an incoming anomaly and serves as early warning to staff in order to take upon the appropriate measures.

There are 26 signals of chemical quantities selected by an experienced chemist. All acquired snapshots of these signals were then filtered and only those satisfying the specific criteria set up in advance were chosen to build up the possibilistic fuzzy clusters of normal situations and for normality index calculations.

An example of a normality index window is depicted in Fig.3.

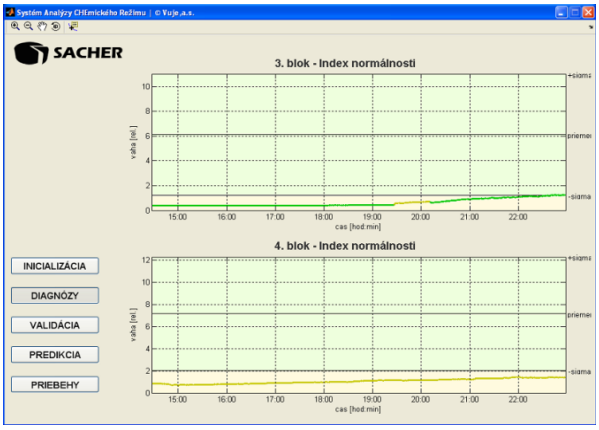

Fig.3 An example of a normality index window.

The normality index values for both the 3rd and 4th units over the last 512 minutes (8.5 hours) are shown in the sliding windows. In the case presented, both indices are slightly below the normal band, which is a strip of light green background color. The normality

index line is color-coded with green, amber, brown, or red for all signals available, one, two, or three and more signals respectively missing. The situation that the less inputs available, the worse evaluation of the normality index can be understood by the change of line color.

The bigger normality index value means the situation is closer to normal. The results of the normality module, which is based on data- driven models, are in consensus with the identification of a normal situation by the fuzzy identification module. It is conversely based on the rules set up by an experienced chemist. Nevertheless, the normality index usually remains in the range of the low values. It is the result of the first version of the normality module having been built up on data acquired in the periods some time ago when the water chemistry regime was controlled under slightly different operational conditions than the current standard. This means that the established clusters are not the best representatives of the current operational situations. The new cluster partitioning is being prepared and will be issued after the completion of thorough evaluation of the collected data against possible failures.

3.3 Fuzzy identification module

The objective of fuzzy identification module is to identify the anomaly in the chemical regime. Despite the previous normality module in which only the deviation from the normal regime is expressed by the normality index as the degree of membership to clusters of normal situations, in the fuzzy identification module, the situation, whether normal or anomalous, is recognized on the basis of a set of fuzzy rules. The fuzzy if-then rule is made up of a number of antecedent and consequent linguistic statements, suitably related by fuzzy connections. They were proposed by an experienced chemist and apply his knowledge on the chemical process, which is acquired over a long period of time.

The fuzzy rule base consists of a set of R rules (currently 80), each assigned to an anomaly to be identified, *e.g.*:

R_j: if (p_1 is P_{1j}) or (t_1 is T_{1j}) and…and (p_n is P_{nj}) or (t_n is T_{nj}) then (a is A_j)

The measured antecedent variables p are represented by the fuzzy sets P, *e.g.* HIGH, NORMAL, or LOW, into which range of each variable is partitioned. The estimated trend variables t_i are represented by fuzzy sets T, *e.g.* INCREASING, STABLE, or DECREASING. Although the fuzzy rules were originally proposed with the trend variables for diagnosing the anomalous situations, yet the current version does not include them. The reasons are justified in the later section of this article.

The membership functions to the fuzzy sets are of typical trapezoidal shape, which again were proposed by an expert. Similarly to the zero-order Takagi-Sugeno-Kang fuzzy model, the consequents a are represented by singleton spikes A, which refer to the anomaly to be identified.

The first step is to fuzzify the crisp numerical input values, *i.e.* to determine by what degree they belong to each of the appropriate fuzzy sets via membership functions. In the second step, the connective operators *and* and *or* are interpreted as minimum and maximum, respectively. Implication method implemented for each rule is the product operator. The firing strength of each rule then reflects by what degree the rule is activated by the incoming inputs, *i.e.* the respective anomaly degree in which the current operational situation is at that moment, as shown in Fig.4.

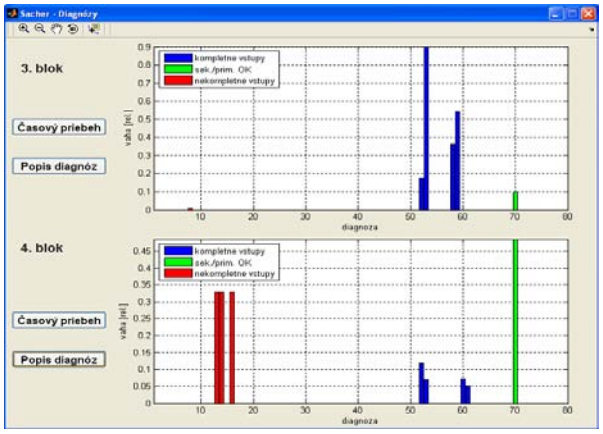

Fig.4 An example of a fuzzy identification window.

According to Fig.4, usually more diagnoses, *i.e.* more fuzzy rules, can fire concurrently. The reason is that the antecedent conditions of more rules, matching the anomalous situations near to each other, are fulfilled to

a certain degree. Normally in the diagnostic systems, the rules in the fuzzy rule base are connected by *else* connective operator interpreted by maximum, which would end up with the final output. In the current system, this step has been omitted even if the user can see more rules firing at the same time. This approach was adopted on purpose as in reality, the current first version of the fuzzy rule base does not include the trend variables. Thus the rules are not the optimal representatives of the anomalies. It has been found that the recognition of the trends in the inputs was not sufficiently reliable to avoid degraded evaluation of the rule. It is therefore better from the perspective of human-factor issues to leave the final judgment on experienced user. Despite these limitations, the users may experience an overall satisfaction from the application of the fuzzy identification module. Nevertheless it will be the further tuning task for the next upgrade of the system to minimize this effect by refining the rules and membership functions.

In Fig. 4 the columns are color-coded. Blue represents the situation when all chemical quantities required for evaluation of the fuzzy rule are available. Green is reserved for normal situation. This allows users to easily identify the degree of its fulfillment. If one or more quantities entering the antecedent part of the rule are missing (this corresponds to the incomplete input information), the color of the column changes to red. The more inputs missing, the more the firing strength of the rule might be negatively affected. This would be expressed through the change of line color.

The users may also obtain the time trends of all diagnostics in the form of a 3D waterfall diagram and also their description by clicking a proper button.

3.4 Time-prediction module

The objective of time-prediction module is to predict the behavior/trend of the selected measured chemical quantities 8 hours ahead in 15 minutes step from the moment of request. The model uses properly-trained artificial neural networks, each giving the prediction to the specific time step. Each network has 3 layers, input layer of 48 equidistant time-delayed inputs, 1 hidden layer and one output layer with one node for prediction. It means that for one quantity to be predicted 8 hours ahead, there are 32 neural networks to be engaged. In reality the neural networks would even be more, as the entire operating range of each quantity has been partitioned into 3 or 4 fuzzy clusters by fuzzy c-mean algorithm. This partitioning aims to avoid training of the neural networks to capture all situations at once. Each cluster is assigned to its own set of neural networks. It has been proven that the neural networks perform better if they are trained to recognize only a specific types of situations.
Figure 5 illustrates the time prediction of a selected chemical quantity.

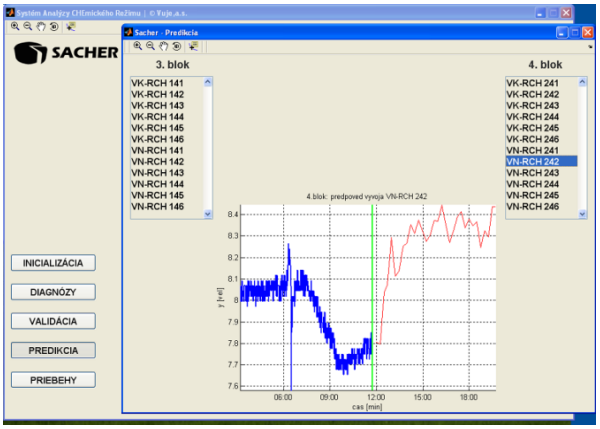

Fig.5 An example of time prediction window.

During the prediction task, the cluster that reflects the most of the given situation is found in the beginning. Afterwards the prediction is conducted by neural networks that are trained to situations of that respective cluster.

The training and testing data were obtained from the real process. The mean absolute percentage prediction error (MAPE) in most cases is approximately 1% for short-time prediction and has a growing trend to 10% for long-time prediction 8 hours in advance, which satisfies the original project criteria. Yet there are situations when the signals change their stationary behavior and the prediction accuracy decreases. As mentioned above, the neural networks' inputs are 48 delayed values representing 8 hours of past time. For successful prediction of the time series, one needs a much longer sliding time delay window in the input than the prediction time horizon applied for the neural networks output. The current 8 hour window was originally selected as the minimal trade-off to comply with the computer speed and memory capabilities. It limited the choice by the authors of this paper to the

present length of the data buffer as 512. The 2^k length was chosen for discrete wavelet based denoising of inputs to the neural networks.

In this first version, only a limited effort on the selection of delays was performed and they were left equidistant. The other limitation of the current system is that only the past values of the same time series are used for the univariate time series prediction. Forecasting better prediction performance may be attained if variations in the other related variables are also taken into account. However by the authors' trial case, it resulted in more complicated neural networks without significant improvement to the forecasting outcome. The authors of this paper conjecture that this might be ascribed to the short input time-delay window which the authors adopted in this paper.

In the further step of the authors' study, the time-prediction module will be rearranged after switching to 64bit architecture of the new computer with the aim of improving the prediction accuracy.

3.5 Validation module

The objective of validation module is to validate the measured quantities. The originally intended neuro-fuzzy approach to validation of readings of chemical quantities did not produce satisfactory results. This is perhaps due to the relations between the chemical quantities, which are many times vague and with substantial delays.

As a consequence, the auto-associative kernel regression method based on fuzzy c-mean classification and adaptive distance measure[6],[7] was chosen for validation. The true expected value of the measured quantity is calculated as the weighted average of the values obtained from the nearest clusters in a certain vicinity to the most representative one. The centers of these clusters have been found in the training phase in advance. The mismatch between the measured values and their true expected calculated counterparts can be estimated whether or not exceeding the properly chosen error band.

Figure 6 shows the validation of a selected chemical quantity from the 4th unit.

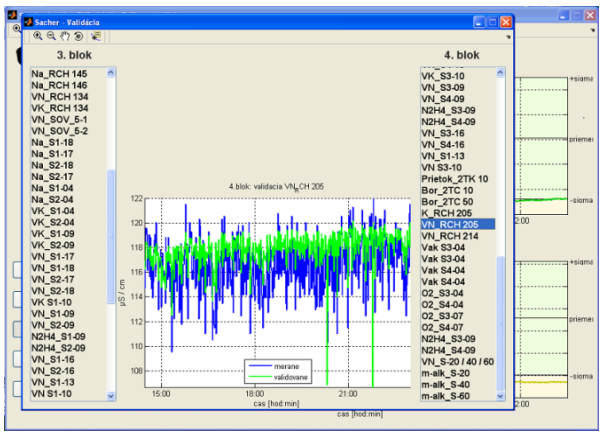

Fig.6 An example of validation window.

3.6 Trend module

This module serves for showing the trends of the acquired quantities as seen in Fig.7.

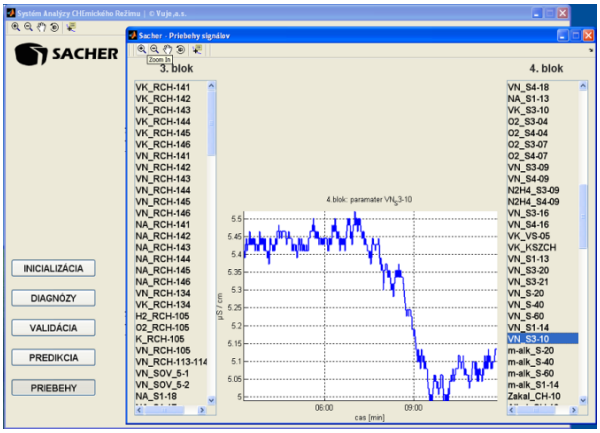

Fig.7 An example of a trend window .

4 Conclusions

This article provides a short description of the chemical information system SYMOCHER/ CHEMIS and its intelligent extension SACHER implemented at the Slovakian Bohunice NPP.

SACHER aims to support chemists in the early detection and identification of anomalies in the primary and secondary circuit water chemistry. This would help chemists in identifying the latest and developing situations at the NPP.

SACHER, which makes use of computational intelligence techniques, has been newly-developed and is being further tuned on the basis of acquired experience. There are further plans for improvements, mainly to taking into account the new operational

situations in the water chemistry regime. The shortcoming of the current version is that the fuzzy rules do not consider the trends of ascending, stable, or descending. The reason is that unreliable recognition would cause inaccurate evaluation of the strength of anomalies. The next step to make the fuzzy identification module more accurate is to develop a reliable algorithm of chemical quantity trend recognition.

This expert-like intelligent extension to the existing chemical information system, particularly for its ability of early notification of abnormal situations in the chemical regime will assist young chemists as well as experienced chemists with overloaded responsibilities. In the future it is also expected that it will assist control room staff even when the absence of chemists in the shift crew. Although this system is effective, it is not intended to fully replace the presence of experienced chemists to decide upon corrective actions. It is expected that SACHER will be used for diagnostics and as an additional tool for data handling.

Acknowledgement

This article presents the outcomes of implementation of the project: “Increase of Power Safety of the Slovak Republic” (ITMS: 26220220077) supported by the Research & Development Operational Program funded by the European Regional Development Fund.

Nomenclature

CHEMIS	CHEMical Information System
LIMS	Laboratory Information Management System
NPP	Nuclear Power Plant
MAPE	Mean Absolute Percentage Error
SACHER	System of Analysis of CHEmical Regime
SYMOCHER	SYstem for MOnitoring of CHEmical Regimes

References

[1] IAEA-TECDOC-1505: Data processing technologies and diagnostics for water chemistry and corrosion control in nuclear power plants (DAWAC), Report of a coordinated research project 2001-2005, June 2006

[2] GUSTAFSON, D. E., KESSEL, W. C.: Fuzzy clustering with a fuzzy covariance matrix, In: BEZDEK, J. C., Pal,S.K.: Fuzzy models for pattern recognition, IEEE PRESS,1991: 117-122.

[3] KRISHNAPURAM, R., KELLER, J.: A possibilistic approach to clustering, IEEE Trans. on Fuzzy Systems, 1993, 1(2):98-110.

[4] FANTONI, P. F., FIGEDY, S., PAPIN, B.: A neuro-fuzzy model applied to full range signal validation of PWR nuclear power plant data. In: Second OECD Specialist Meeting on Operator Aids for Severe Accident Management (SAMOA-2), Lion, 1997

[5] TOU, J. T. , GONZALES, R. C.: Pattern recognition principles, Addison-Wesley Publishing Comp., Reading, Massachusetts, 1974: 97-104

[6] HINES, J. W., GARVEY, D. R.: Traditional and Robust Vector Selection Methods for use with Similarity Based Models, NPIC&HMIT, Albuquerque, New Mexico, November 2006

[7] GARVEY, D. R., HINES, J. W.: Adaptive Distance Measure for Use with Nonparametric Models, NPIC&HMIT, Albuquerque, New Mexico,November 2006