Computing and Informatics, Vol. 33, 2014, 387–409

AI-BASED DIAGNOSTICS FOR FAULT DETECTION AND ISOLATION IN PROCESS EQUIPMENT SERVICE

Svetla Vassileva

Institute of System Engineering and Robotics
Bulgarian Academy of Sciences
Acad. G. Bonchev str., bl. 2, 1113 Sofia, Bulgaria
e-mail: si27v06@gmail.com

Lyubka Doukovska, Vassil Sgurev

Institute of Information and Communication Technologies
Bulgarian Academy of Sciences
Acad. G. Bonchev str., bl. 2, 1113 Sofia, Bulgaria
e-mail: doukovska@iit.bas.bg, sgurev@bas.bg

Abstract. Recent industry requires efficient fault discovering and isolation solutions in process equipment service. This problem is a real-world problem of typically ill-defined systems, hard to model, with large-scale solution spaces. Design of precise models is impractical, too expensive, or often non-existent. Support service of equipment requires generating models that can analyze the equipment data, interpreting the past behavior and predicting the future one. These problems pose a challenge to traditional modeling techniques and represent a great opportunity for the application of AI-based methodologies, which enable us to deal with imprecise, uncertain data and incomplete domain knowledge typically encountered in real-world applications. In this paper the state of the art, theoretical background of conventional and AI-based techniques in support of service tasks and illustration of some applications to process equipment service on bio-ethanol production process are shortly described.

Keywords: Process equipment service, fault detection and isolation, residuals, artificial intelligence, bio-ethanol production

Mathematics Subject Classification 2010: 68T01

1 INTRODUCTION

Today, product liability, variety and production flexibility require measures in the supervision of processes, which rule out errors during production. Modern industry asks for efficient fault discovering and isolation methods in process equipment service. This problem is a real-world problem of typically ill-defined systems, hard to model, with large-scale solution spaces. In such cases the design of precise models is impractical, too expensive, or often non-existent. A possible solution can be generated by leveraging resources as well as:

- problem domain knowledge of the process equipment faults that is necessary to fix and
- field-data that characterize the behavior of process equipment faults.

The above-mentioned two kinds of resources determine two main approaches found in artificial intelligence methods: knowledge-driven reasoning systems (probabilistic and multi-valued systems) [6, 7, 8, 9, 15, 16, 20, 22, 32, 33, 37, 44] and data-driven search and optimization approaches (neuro, neuro-fuzzy and evolutionary computing) [2, 10, 11, 12, 13, 23, 29, 36, 40, 41, 43]. The relevant available domain knowledge is typically a combination of first principles and experiential or empirical knowledge, which is usually incomplete, uncertain and erroneous [1, 2, 3, 5, 7, 14, 16, 21, 27]. In connection with the data-driven search, the available data are typically a collection of input-output measurements, representing instances of the system's behavior, and are usually incomplete and noisy [18, 19].

Process Equipment Service can be optimized to prevent failures and maximize uptime while avoiding superfluous maintenance. Some of these objectives can be accomplished by using tools that measure the system state and indicate arising failures. Such tools ask for high level of sophistication and incorporate monitoring, fault detection, decision making, possible preventive or corrective actions and execution monitoring [49]. Support service of equipment requires generating models that can analyze the equipment data, interpreting their past behavior and predicting the future one. These problems pose a challenge to traditional modeling techniques and represent a great opportunity for the application of AI-based methodologies.

Because of the complexity of these tasks, AI-methods have been forced in the implementation of fault detection and isolation tools. Some application of AI-based techniques in support of service tasks, such as anomaly detection and identification, diagnostics, prognostics, estimation and control, have been reported in [10, 11, 12].

In this paper the state of the art, diagnostics and prognostics tasks, theoretical background of conventional and AI-based techniques in support of service tasks and illustration of some most successful applications to process equipment service on bioethanol production are shortly described. The detection procedure is implemented on the laboratory plant in the Institute of Cryobiology and Food Technologies.

2 DIAGNOSTICS AND PROGNOSTICS TASKS

The task of diagnosis is to find an explanation for a set of observations and – in the case of prognosis – to forecast the course of events. Diagnosis can be broken down into anomaly detection and failure identification, depending on the desired granularity of information required [3, 4, 6, 7, 9, 14, 19, 20, 22, 30, 31]. Prognosis is concerned with incipient failure detection, margin prediction, or overall performance prediction. The latter can be prediction of efficiency, current system status, etc. The outcome of diagnosis and prognosis processes drives planning and execution.

Possible planning includes planning of corrective action that can be either reactive or proactive [25, 26]. Another possible plan is maintenance planning which has to take into consideration not only the current system status, but also the cost of maintenance (out of sequence vs. routine), disruption of service, and the cost of further damage. All these steps can be interim fixes or tactical decisions. Three concepts are discussed in [25]: clearly identified sources of data which identify problems that will be investigated; root cause analysis (RCA) to identify the cause of a discrepancy or deviation and suggest corrective actions to potentially prevent recurrence of a similar problem, or preventive action to ensure that discrepancies do not occur and finally – remedial corrections of a problem which is identified.

For example, improving the Department of Energy's project the Root Cause Analysis (RCA) concept identifies the key elements, necessary to make the meaningful changes required to consistently deliver projects within cost and schedule performance parameters; disciplined upfront planning; realistic estimates of cost and schedule; and straight forward communication between the project director and senior management [26]. Food-quality prevention is presented in [34], where the knowledge-based method (decision tree) is implemented in understanding the food safety aspects related to the brewing process and its technological equipment control in the Critical Control Points (CCPs). CCP is any point or procedure in a food process where loss of control may result in an unacceptable health risk. Hazard Analysis Critical Control Point (HACCP) forms a key component of many certified compliance standards and is recognized as a main element of international trade in food products. In operations research, specifically in decision analysis, a decision tree (or tree diagram) is a decision support tool that uses a graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility [31, 32]. A decision tree is used to identify the strategy most likely to reach a goal. Another use of trees is as a descriptive means for calculating conditional probabilities.

Some of the challenges that diagnosis and prognosis systems face (besides imprecise data and incomplete understanding of the problem domain) are the ability to a changing environment which must allow the distinction between "normal" wear or desired system changes and the reportable system deviation. The transients are often very similar and a proper distinction becomes important. Environments do not only change with time but also in space.

For example, the clustered damage precursors of the technological processes are usually correlated with underlying damage for anomaly detection and fault-mode isolation for prognostics health monitoring of electronics subjected to drop and shock in [30]. Feature extraction techniques in the joint-time frequency domain are developed along with pattern classifiers for fault diagnosis of electronics at product-level [1, 31, 42, 48, 50].

Both corrective and preventive actions include investigation, action, review, and further action if so required. Corrective action includes reconfiguration of the current system or sub-system, de-rating the set-point, or changing the goal. Examples of corrective actions include error proofing, visible or audible alarms, process redesign, product redesign, enhancement or modification of existing training programs, improvements to maintenance schedules, improvements to material handling or storage. A combination of such actions may be necessary to fully correct the problem.

Mobile diagnosis or remote monitoring and diagnosis systems are one possible answer to the recent industrial plants control. However, accessibility and ability to transmit are limited by bandwidth and cost. Therefore, the system may be equipped with remote repair capabilities [35], which is a step towards an autarkic system. This implies a more sophisticated decision maker that can reason about the information gathered and come to an optimal judgment within the constraints of the system.

3 THEORETICAL BACKGROUND

3.1 Principles of Model-based Fault Detection and Isolation

Fault isolation task can only be realized if the fault to be isolated has been previously taken into account in the model [28, 39]. There are different approaches for the design of diagnostic observers: the geometric methods [31], algebraic methods [7, 24, 49], spectral theory-based methods [48] and frequency domain solutions [17].

The main goal of this part is the presentation of a Fault Detection and Isolation (FDI) observer-based method applied to a non-linear process. The method is designed with a dynamic model and the observer is determined using the eigenstructure assignment approach. The principle of observer-based FDI approach is to compute residuals by comparing estimated states with the measured outputs of the actual plant. The residuals are ideally zero when a fault is not detected and they become non zero if the actual system differs from the model. The isolation of faults follows the detection stage and requires a set of residuals. Generation of residuals is realized in two different ways: the structure residuals method and the fault detection filter method.

Theoretical background of this method is as follows. An FDI system generally includes two stages (Figure 1). The first one uses input-output observations to elaborate a set of residual relations which enable fault detection and isolation. The second one, which is the decision-making stage, involves residual (r) evaluation. Independently of the method used for residual generation, the following conditions have to be considered:

- the mathematical model of the system is uncertain;
- measured signals are affected by noise with unknown characteristics;
- the occurrence of the fault in the course of time is unknown (unanticipated fault).

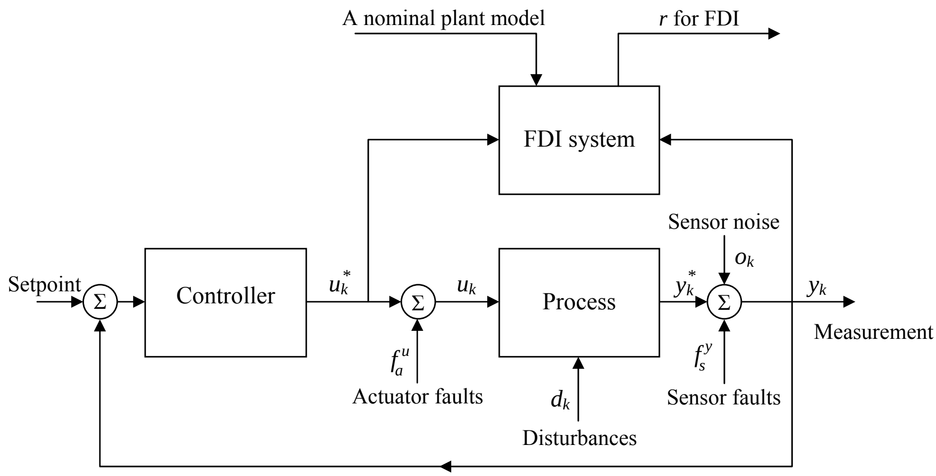


Figure 1. Fault Detection and Isolation (FDI) – general scheme

The principle of observer based FDI approach is to compute residuals by comparing estimated states with the measured outputs of the actual plant [34]. The residuals are ideally zero (i.e., free of faults) and they become non zero if the actual system differs from the model. Different reasons may cause non-zero residuals as faults, disturbances or plant-model mismatches (model errors). The discrete state equations describing the system can be written as:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) = F_1 f_a(k) + E_1 d(k) \\ y(k) &= Cx(k) + F_2 f_s(k) + E_2 d(k) \end{aligned} \tag{1}$$

where x is the system state vector, y is the system output vector, u is the system input vector, f_a is a vector containing actuator faults, f_s is vector containing the sensor faults and d is a vector containing unknown input while $A, B, C, F_1, F_2, E_1, E_2$ are known matrices of appropriate dimensions.

The corresponding observer equations are:

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + H\left[y(k) - \hat{y}(k)\right] \\ \hat{y}(k) &= C\hat{x}(k) \\ e(k+1) &= x(k+1) - \hat{x}(k+1) \\ e_0(k) &= y(k) - \hat{y}(k) \end{aligned} \tag{2}$$

where $e(k)$ is the dynamic estimating error, $e_0(k)$ is the observation error, H denotes the feedback gain matrix and is designed using eigen-structure assignment so that the observer is stable.

The dynamic equation of the estimating error is:

$$e(k+1) = [A - HC]\, e(k) + F_1 f_a(k) - HF_2 f_s(k) + (E_1 - HE_2)d(k), \tag{3a}$$

which can be rewritten as:

$$e(k) = (zI - A + HC)^{-1} \left[(E_1 - HE_2)\, d(k) + F_1 f_a(k) - HF_2 f_s(k).\right] \tag{3b}$$

The fault isolation is the step following the fault detection stage which requires a set of residuals. Each residual that belongs to this set has to be individually sensitive to the faults that may affect the process or the components. Usually, the generation of residual set is realized in two different ways: the structured residuals method and the fault detection filter method. When the process consists of interconnected subsystems, structured residuals can be generated individually for each of them. However, the decomposition in subsystems requires a good knowledge of the global plant. Structured residuals can then be generated with a bank of reduced order observers [19]. Each observer is dedicated to one measured output; and, as a consequence, one residual is generated for each observer and a fault occurring on a measurement will affect the corresponding estimation. On the other hand, the fault detection filter is a state observer designed in such a way that a residual deviation due to a particular fault is confined to a one-dimensional subspace of the output space. Considering the dynamic estimation error $e(k)$ and the observation error $e_0(k)$ in Equation (2), we obtain:

$$e_0(k) = Ce(k) + F_2 f_s(k) - E_2 d(k) \tag{4}$$

and the residual generator takes the following form:

$$r(k) = Ne_0(k) \tag{5}$$

where N is the filter matrix determined so that properties of isolability and detectability are verified. Matrices H and N are determined by using eigen-structure assignment procedure. This approach is built upon the fundamental eigen-pair equations:

$$\begin{aligned} (\lambda_j - G)v_j &= 0 \\ w_j^T(\lambda_j - G) &= 0 \\ \det(\lambda_j - G) &= 0, \end{aligned} \tag{6}$$

where λ_j is an eigen-value, v_j is a right side eigen-vector of G and $w_j T$ is a left side eigen-vector of G. The technique used for the detection filter is based upon the right assignment approach. Using expression of $e_0(k)$ given in Equation (4) and replacing

$e(k)$ by its expression in Equation (3b), residual in Equation (5) becomes:

$$\begin{aligned} r(k) &= NC(z - IA + HC) - 1((E_1 - HE_2)d(k) + F_1 f_a(k) - HF_2 f_s(k))) \\ &\quad + N(E_2 d(k) + F_2 f_s(k)) \end{aligned} \tag{7}$$

Using Equation (7), H is designed such that the columns of matrices F_1 and HF_2 are the side eigen-vectors of $A + HC$ belonging to $\lambda_i = 0$ eigen-values. If the unknown input $d(k)$ is neglected, then the residual can be written as:

$$r(k) = N(Cz^{-1}(F_1 f_a(k) - HF_2 f_s(k)) + F_2 f_s(k)). \tag{8}$$

The matrix N is designed in such a way that the effects of a fault are decoupled from the effects of the other faults. If NCF_1, $NCHF_2$ and NF_2 are zero in each row, except in the i^{th} row, the residual is affected by the i^{th} fault.

3.2 Decoupling the Actuator Fault from the Sensor Fault

In the following, the method is applied to decouple the effect of an actuator fault from a sensor fault [34]. The matrix N is decomposed into two matrices N_1 and N_2, designed such that:

$$r_a(k) = N_1 z^{-1}(CF_1 f_a(k) + CHF_2 f_s(k)) + N_1 F_2 f_s(k) \tag{9a}$$

and

$$r_s(k) = N_2 z^{-1}(CF_1 f_a(k) + CHF_2 f_s(k)) + N_2 F_2 f_s(k). \tag{9b}$$

The conditions for $r_a(k)$ to be sensible only to $f_a(k)$ are given by:

$$N_1 CF_1 \neq 0 \tag{10a}$$

$$N_1 CHF_2 = 0 \tag{10b}$$

$$N_1 F_2 = 0 \tag{10c}$$

The conditions for $r_s(k)$ to be sensible only to $f_s(k)$ are given by:

$$N_2 CF_1 = 0 \tag{11a}$$

$$N_2 CHF_2 \neq 0 \tag{11b}$$

$$N_2 F_2 \neq 0 \tag{11c}$$

3.3 Principle of AI-Based Fault Detection and Isolation

The AI-based fault detection and isolation is a data-driven technique where available data are typically a collection of input-output measurements, representing instances of the system's behavior, and are usually incomplete and noisy. Models, derived from the data for FDI, principally aim at the best possible estimation of those output measurements, which are influenced by the faults of interest. The data-based models are usually non-linear in contrast to the functional analytical models which are often linear and hence less complex.

The data-based models, usually black-box models, lie in the core of a modular diagnosis system concept (Figure 2) which has been chosen as separate fault detection system. Each of these systems is handling only partial information on the process. This is similar to different persons analyzing the same situation with different methods and/or different sources of information.

Low level fault detection systems for control of one technical unit are to be further combined within an overall system called a "state manager" [3]. This state manager is detecting faults when the set of information provided by the sensors is either incoherent (i.e., sensor fault due to fouling) or shows an abnormal working mode, the supervision module also can detect the presence of a steady state (of normal or abnormal state) based on principal components analysis (PCA) [9, 45], data space reduction and multivariate statistical analysis. The overall fault detection and diagnosis system must fit most of the plants and automatically adapts to any evolution (new sensor, change of an actuator, etc.).

Detection of faults is carried out by the developed knowledge-based system its rule-base consists of IF-THEN rules, defined by experts or by extracting knowledge from the data records. Detection and diagnosis of hardware sensors is expressed as "Normal status", "Damaged sensor", "No data", etc. Detection of software sensor is related with the process state variables status as well as "High", "Low", "Middle" or "Normal", "Toxic", etc. for concentrations, specific rates, product quality factors and ecological parameters.

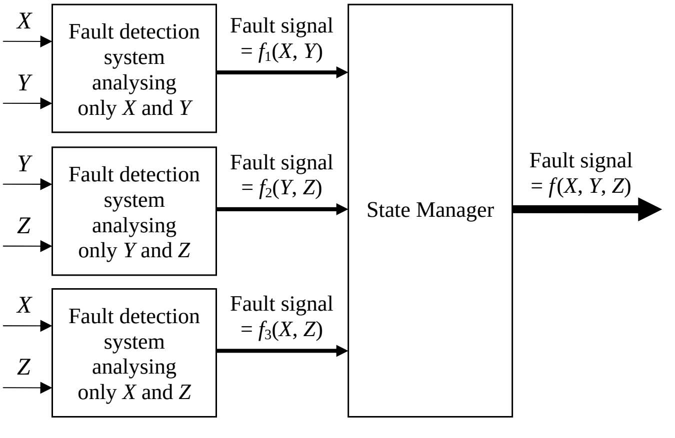


Figure 2. Concept of a modular diagnosis system

Keeping in mind that for FDI only those outputs need to be estimated, which are affected by faults, one can conclude that data-based models for FDI may be of reduced order and hence of less complexity due to an adequate input space in comparison to functional models used for control.

The given set of input-output data of the possibly non-linear system can be used to train a properly pre-structured non-linear model. The learning can be achieved by adaptation due to a given performance index.

3.3.1 Fuzzy-Logic-Based FDI

Fuzzy models as a nonlinear black box-structure represent the relationships between past observations $[u(t-1), y(t-1)]$ and the future outputs $y(t)$ of a general discrete time dynamic system:

$$y(t) = g(u(t-1), y(t-1)) + v(t). \tag{12}$$

The additive term $v(t)$ accounts for the fact that the next output $y(t)$ will not be an exact function of past data. The goal of adequate modeling is to minimize $v(t)$ in order to achieve a good prediction. The function $g(.)$ can be found from its parametrized form with a finite dimensional vector θ by the approximation. For convenience of calculations, a concatenation of two mappings is introduced: one that takes the increasing number of past observations $[u(t), y(t)]$ and maps them into a finite dimensional vector $\varphi(t)$, which is a regression vector and its components are regressors, which give necessary freedom in the linear black-box case, and it is natural to implement them in the nonlinear case. Finally the following model structure is used:

$$\hat{y}\left(t\theta\right) = g(\varphi\left(t\right), \theta). \tag{13}$$

Fuzzy models, implemented according the concept of a modular diagnosis system, are presented in the form of rule-bases with several parameters, which contribute to the vague statements such as “large”, “small”, “middle”, to be precise of terms of membership functions. A fuzzy rule basis is a collection of rules:

$$\begin{gathered}\text{If } (\varphi_1 \text{ is } A_{1,1}) \text{ and } (\varphi_d \text{ is } A_{1,d}), \text{ then } (y \text{ is } B_1),\\ \cdots \\ \text{If } (\varphi_1 \text{ is } A_{p,1}) \text{ and } (\varphi_d \text{ is } A_{p,d}), \text{ then } (y \text{ is } B_p),\end{gathered} \tag{14}$$

where the fuzzy sets $A_{i,j}$ are double-indexed, i is the input coordinate and j is the index of rule. The membership functions are denoted $\mu_{A_{j,i}}(\varphi_i)$ and $\mu_{B_j}(y)$.

The fuzzy rule basis exhibits the model structure with some features related to the elementary functions in the decomposition:

$$y = \sum_{j=1}^{p} y_j \left(\prod_{i=1}^{d} \mu_{A_{j,i}}\left(\varphi_i\right) \right) \triangleq \sum_{j=1}^{p} y_j \omega_j\left(\varphi\right) = g\left(\varphi\right), \tag{15}$$

where $\varphi = (\varphi_1, \ldots, \varphi_d)\,, y_j$ is the point at which μ_{B_j} reaches its maximal value, and the definition of the weight functions $\omega_j(\varphi)$ is obvious.

If $\sum_{j=1}^{p}\prod_{i=1}^{d}\mu_{A_{i,j}(\varphi_i)} \neq 1$, then the defuzzification formula is modified according to Wang (1992, [47]) as follows:

$$y = g\left(\varphi\right) = \frac{\sum_{j=1}^{p} y_j \omega_j\left(\varphi\right)}{\sum_{j=1}^{p} \omega_j\left(\varphi\right)}. \tag{16}$$

In such a case, rule basis may be directly built by crisp conclusions, because no defuzzification is needed.

3.3.2 ANN-Based FDI

Concept of modeling the discrete time dynamic system, when ANN-based FDI is implemented, lies on Equation (12), function parametrization and regression vector $\varphi(t)$, as shown above. A related NN-based model was suggested in [36]:

$$\hat{y}(t) = f(\theta_1, \varphi_1(t)) + g(\theta_2, \varphi_2(t)), \tag{17}$$

where $\varphi_1(t)$ consists of delayed outputs and $\varphi_2(t)$ consists of delayed inputs. The parametrized functions f and g can be chosen to be linear or non-linear by a NN. A further motivation for this model is that it becomes easier to develop controllers.

Building linear model for the system is suggested by Qin and McAvoy in 1992 [38], because the residuals from this model will then contain all un-modeled nonlinear effects. The NN-model could then be applied to residuals (treating inputs and residuals as input and output) to pick up nonlinearities. This approach is attractive since the first step to obtain a linear model is robust and often leads to reasonable solution. The second NN-step assures to obtain at least as good a model as the linear one.

FDI realization is often based on multi-layer networks; its mapping is convolved with each other. Let the outputs of the basis functions be denoted by

$$\varphi_k^{(2)}(t) = g_k(\varphi(t)) = k(\varphi\left(t\right), \beta_k, \gamma_k) \tag{18}$$

and collect them into a vector

$$\varphi^{(2)} = \left[\varphi_1^{(2)}(t), \ldots, \varphi_n^{(2)}(t)\right]. \tag{19}$$

Instead of taking a linear combination of these $\varphi_k^{(2)}$ as the output of the model we could treat them as new regressors and insert then into another layer of the basis function forming second function expansion:

$$g(\varphi, \theta) = \sum_{l} \alpha_l^{(2)}, k\left(\varphi^{(2)}, \beta_l^{(2)}, \gamma_l^{(2)}\right), \tag{20}$$

where θ denotes the whole collection of involved parameters $\alpha_k, \beta_k, \lambda_k, \alpha_l^{(2)}, \beta_l^{(2)}, \gamma_l^{(2)}$. Within the NN-terminology, Equation (20) is called two-hidden layer network. The basis functions $k(\varphi(t), \beta_k, \gamma_k)$ then constitute the first hidden layer, while $k\Big(\varphi^{(2)}, \beta_l^{(2)}, \gamma_l^{(2)}\Big)$ give the second layer. This procedure can be repeated in an arbitrary number of times to produce multi-layer NN.

The question how many layers to use is not easy. In principle, with many basis functions, one hidden layer is sufficient for modeling most practically reasonable systems.

Another very important concept for applications to dynamical systems is that of recurrent NN. This refers to the situation that some of the regressors used at time t are outputs from the model structure at previous time instants:

$$\varphi_k(t) = g(\varphi(t-k), \theta). \tag{21}$$

It can also be the case that some components $\varphi_j(t)$ of the regressor at time t are obtained as a value from some interior node at a previous time instant. Such model dependent regressors make the structure more complex and flexible.

Neural networks (NN) originated in an attempt to build mathematical models of elementary processing units in the brain and the flow of signals between these processing units. After a period of stagnation, these formal models have become increasingly popular, with the discovery of efficient algorithms capable of fitting them to data sets. Since then, neural nets have been applied to build computerized architectures that can approximate nonlinear functions of several variables, and classify objects, which is the task of FDI-systems. A neural net is nothing more than a sophisticated black box nonlinear model that can be trained on data.

4 EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experiments and the Process Equipment

The studied process is fermentation process for bio-ethanol production in batch mode [46]. Bio-ethanol is ethanol (C_2H_5OH) produced by biological fermentation of carbohydrates derived from plant-material and from wastes of food and beverages industry.

The experimental work was carried out in a 3l laboratory bioreactor ABR01, equipped with automated system for measuring and control of the physical variables as well as temperature (by heating and cooling), partial pressure of the dissolved oxygen (pO_2) by the aerating volume (Q_{O_2}), agitation speed (n, rpm) and feeding rate (Figure 3).

The observed biochemical data include the biomass concentration X [million cells/ml], substrate content S [mg/l] and ethanol content [mg/l]. The technological conditions were maintained according to the full experiment plan.

In this case, the experimental plan design is oriented to discovering the influence of key technological parameters and process variables on the final concentration of ethanol, to eliminate measurement errors and to obtain measurements not biased by some unavoidable factor.

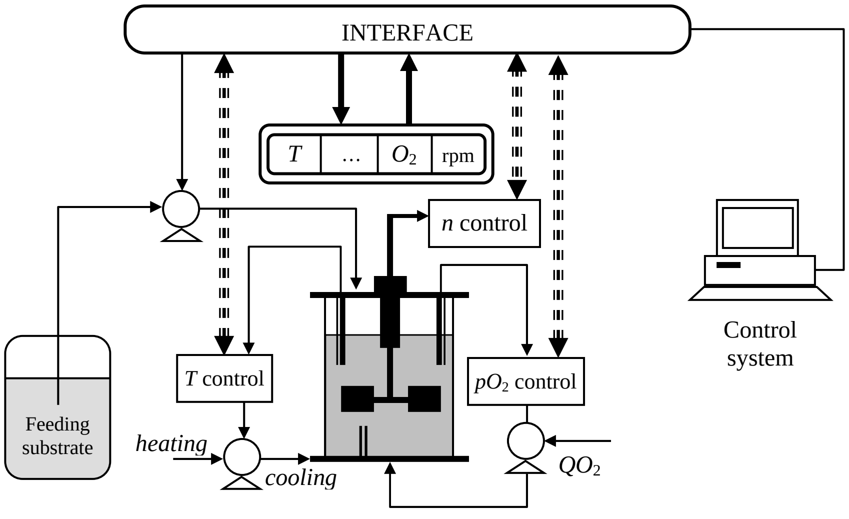


Figure 3. Instrumentation of the bio-ethanol production system

Monitoring bioprocesses by on-line (in-situ) techniques is highly desirable since it has the potential to produce significant improvements in process control. Whilst direct on-line measurement of the states may not be possible, the influence of their variation can be observed in available on-line measurements. It is therefore possible in certain instances to obtain an on-line inference of the process states – such an approach is termed a software-sensor [45]. In other words, a software-sensor is a software algorithm giving an on-line estimation for process state variables, whose analyses are normally time consuming, labor intensive and costly. A software-sensor calculates this prediction from the available on-line measurements using a model and proper mathematical inference.

The main idea of the advanced control strategy is that the efficiency of hardware sensors is complemented by software sensors which combine the information from the sensor network with a process model (or a model set from the Knowledge Base) in order to predict some key-process variables (e.g., biomass content, fermentation activity of microbial population, respiratory quotient (RQ), CO_2 concentration, chemical oxygen demand (COD), Higher Heating Values (HHV) evaluation of biomass fuels, CO_2-recovery and many others, which are generally not available on-line.

Software sensor systems are applicable to linear and non-linear systems, when uncertainty or incomplete information is available, for mono-phase and multiphase processes.

4.2 Observer-Based FDI – Solution and Results

Suppose that for the studied process the known matrices A, B, C, F_1 and F_2 are given as it was shown in Subsection 3.1:

$$A = \begin{bmatrix} 1.81 & 1 & 0 & 0 & 0 \\ -0.81 & 0 & 0 & 1 & 0 \\ 0 & 0 & -0.37 & 0 & 0 \\ 0 & 0 & 0 & 1.23 & 1 \\ 0 & 0 & 0 & -0.23 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.54 \\ -0.52 \\ 0 \\ -0.005 \\ 0.0048 \end{bmatrix}, \tag{22}$$

$$C = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$F_1 = \begin{bmatrix} 0.54 \\ -0.52 \\ 0 \\ -0.005 \\ 0.005 \end{bmatrix},$$

$$F_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{23}$$

Under conditions $(A+HC)F_1 = 0$ and $(A+HC)HF_2 = 0$, faults are decoupled from estimation error dynamics and one stabilizing solution is given by matrix H:

$$H = \begin{bmatrix} 0.99 & 17.50 \\ -0.81 & 0 \\ 0 & 0 \\ 0 & 0.27 \\ 0 & -0.23 \end{bmatrix} \tag{24}$$

Setting $N = \begin{bmatrix} 0 & -50 \\ 0.1 & 10.2 \end{bmatrix}$ allows decoupling of r given by Equations (10) and (11), to be satisfied with $NCF_1 = \begin{bmatrix} 0.25 \\ 2.10^{-3} \end{bmatrix}$, $NCHF_2 = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$ and $NF_2 = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$.

The computation of residuals is then realized as follows:

$$r(k) = \begin{bmatrix} r_a(k) \\ r_s(k) \end{bmatrix} = N(z^{-1}(CF_1 f_a(k) - CHF_2 f_s(k)) + F_2 f_s(k)). \tag{25}$$

4.3 Neuro-Fuzzy Prediction of Bio-Ethanol Production with Fault Detection

The work reported here is concerned with the design and application of model-based fault detection method to a biotechnological process for bio-ethanol production.

The main problems related to the alcoholic fermentation process include the lack of robustness of the fermentation in the presence of fluctuations in the quality of the raw material and modifications in microbial metabolism. These lead to changes in the kinetic behavior with impact on yield, productivity and conversion of the process. The lack of robustness can be corrected by adjustments in the operational and control parameters of the process when fluctuations occur.

In order to accomplish this, it is important that a mathematical model be available to aid in the decision making, mainly when the difficulties of monitoring the key process variables (concentrations of biomass, substrate and ethanol, cultivation conditions) are taken into account. However, the operational changes described make the prediction of the dynamic behavior of the process with a single model difficult, as they lead to changes in microorganism kinetics. Thus, it would be of great advantage to have a mathematical model that could be easily adapted to changes in operational conditions. A way to deal with this problem is to use fuzzy, neuro-fuzzy or neural models. Recent investigations show these models perform better than first principle models [45].

For building fuzzy model from data, generated by poorly understood dynamic systems well as multi-factorial bio-systems, the input-output representation in the form of NARX model is often applied. The NARX model can represent observable and controllable models of a large class of discrete-time MISO non-linear systems. In fuzzy modeling, the function F is represented as a collection of IF-THEN rules R_i:

R_i: IF $y(k)$ is $A_{i,1}$ and … and $y(k-ny+1)$ is $A_{i,n}$
and if $x(k)$ is $B_{i,2}$ and … and $x(k-nu+1)$ is $B_{i,n}$, THEN:

$$y(k+1) = \sum_{j=1}^{n_y} a_{i,j} y(k-j+1) + \sum_{j=1}^{n_u} b_{i,j} x(k-j+1) + c_i, \tag{26}$$

where $A_{i,l}, B_{i,1}$ are fuzzy sets and $a_{i,j}, b_{i,j}$ and c_i are crisp consequent parameters. The weighted means output $y(k+1)$ of the model is:

$$y(k+1) = \frac{\sum_{i=1}^{k} \lambda_i \left(y(k), \ldots, x(k-n_u+1)\right) y_i\left(k+1\right)}{\sum_{i=1}^{k} \lambda_i \left(y(k), \ldots, x(k-n_u+1)\right)} \tag{27}$$

where the normalized form of the fulfillment degree is presented by:

$$\lambda_i \left(y(k), \ldots, x(k-n_u+1)\right) = \frac{\lambda_i \left(y(k), \ldots, x(k-n_u+1)\right)}{\sum_{i=1}^{k} \lambda_i \left(y(k), \ldots, x(k-n_u+1)\right)}. \tag{28}$$

The model output $y(k+1)$, affine in $x(k)$ according to the following nonlinear form, is as follows:

$$\begin{aligned} y(k+1) &= \lambda_i\left(y(k),\ldots,x(k-n_u+1)\right) \\ &\quad\times\left[\textstyle\sum_{j=1}^{n_y} a_{i,j}y(k-j+1)+\sum_{j=1}^{n_u} b_{i,j}x(k-j+1)+c_i\right] \\ &\quad+\textstyle\sum_{i+1}^{k}\lambda_i\left(y(k),\ldots,x(k-n_u+1)\right)b_{i,1}x(k). \end{aligned} \tag{29}$$

In order to fine-tune the parameters which are related to the output in a nonlinear way, training algorithm known from the area of neural network can be employed. These techniques exploit the fact that a fuzzy model can be seen as a layered structure (network) similar to artificial neural network. Hence, this approach is usually referred to as neuro-fuzzy modeling.

The main step of NN-training algorithm employment in adaptive parameter identification of primary obtained fuzzy model is their transformation into a connectionist layer structured NN model. The structure of NN model has to correspond to the fuzzy model structure. The nodes in the first layer compute the membership degree of the inputs in the antecedent fuzzy set. The product nodes P in the second layer represent the antecedent conjunction operator. The normalization node N and the summation node S realize the fuzzy mean operator for Takagi-Sugeno fuzzy models:

$$y=\frac{\sum_{i=1}^{k}\lambda_i\left(x\right)\left(a_i^T x+b_i\right)}{\sum_{i=1}^{k}\lambda_i(x)}. \tag{30}$$

Using smooth antecedent membership functions MSF, such as Gaussian:

$$\mu\left(x;c,\sigma\right)=\exp\left[-\left(\left(x-c\right)/2\sigma^2\right)\right]. \tag{31}$$

The parameters c and σ can be adjusted by gradient descent learning algorithm, such as back-propagation.

Our previous investigations show that microbial growth of yeasts, which conduct the fermentation of sugars into ethanol, is influenced by the rheological conditions in the bioreactor. The rheological conditions are usually determined by the cultivation temperature T [°C] and the partial pressure of the dissolved oxygen pO_2 [%], which influences biotechnological variables – specific growth rate, connected directly with the biomass content (X) and the limiting substrate (S) utilization. In the practice, the both microbiological variables are controlled directly by sensors or inferentially by implementing Kalman-filter procedure followed by the FDI-system.

In our case the latter approach was implemented and fault diagnosis is presented in Figure 4. The isolation of faults from predicted X allows realization of bio-ethanol prediction.

Detailed study of bio-ethanol production shows that limiting substrate (S) or biomass content (X) could be implemented as best predictors of bio-ethanol content (y_2) and its best regressors – $y(k-1)$, $y(k-3)$ and $y(k-4)$. The fuzzy inference system has 8 fuzzy rules in which the input variables are presented by 2-Bellmann

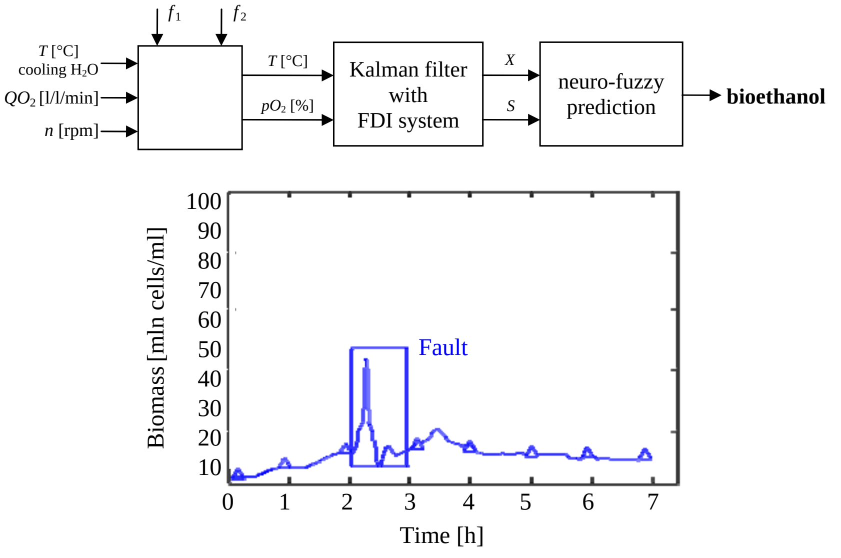


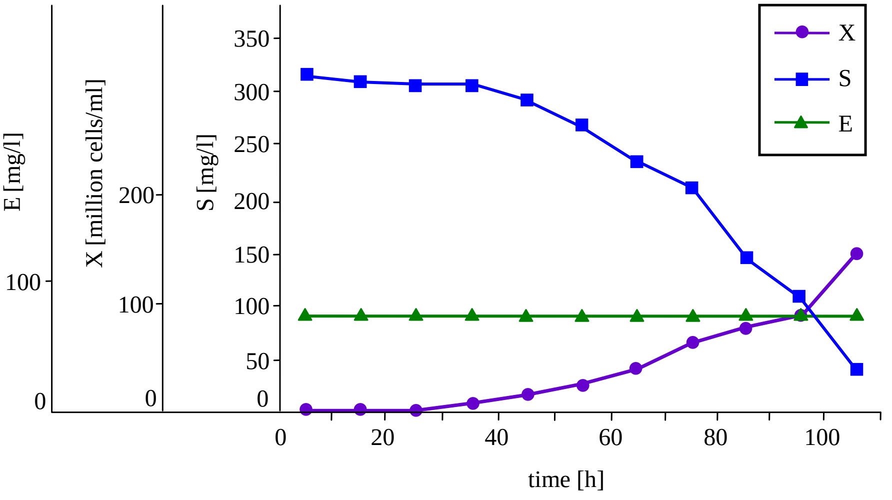
Figure 4. Bio-ethanol prediction: process scheme with the biomass fault diagnosis and isolation

shaped membership functions and the output variable is presented in the form of linear function. The prediction procedure was realized by implementing Fuzzy Toolbox in Matlab, version 6.5. with Simulink (Figure 5).

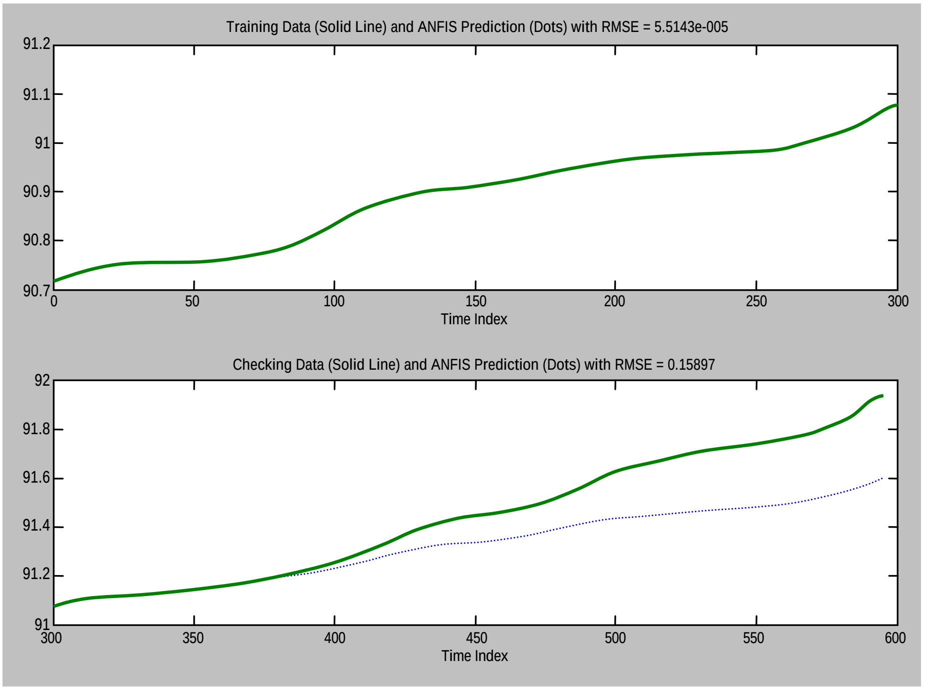
5 CONCLUSIONS

This paper presented the development of a Fault Detection and Isolation (FDI) systems, based on two different principles, which have been applied to a non-linear bio-ethanol production process. The first observer-based FDI method requires state space model in the form of discrete state equations, describing the system, and based on this description observer equations. The residual generator was determined so that properties of isolability and detectability are verified and its matrices are determined by using eigen-structure assignment procedure. As demonstrated, the observer-based method allows detecting and isolating unexpected faults. The principle of observer based FDI approach is to compute residuals by comparing estimated states with the measured outputs of the actual plant.

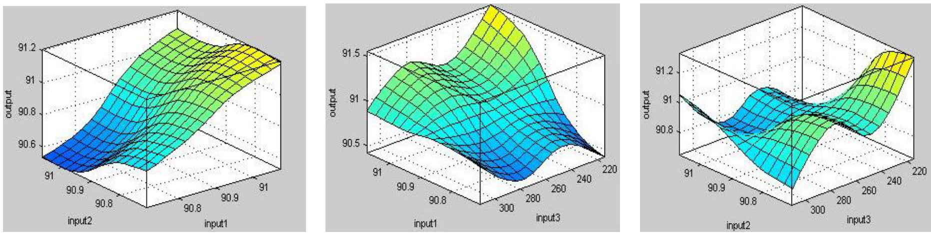
The second AI-based method is implemented for the bio-ethanol content prediction by using information about biomass or limiting substrate concentration, which is acquired from Kalman filter procedure, combined with FDI-system. After that one of the variables X or S is used for numerical prediction of bio-ethanol by Adaptive Neuro-Fuzzy Inference System (ANFIS). More details of prediction procedure are shown in [46].


E [mg/l]
X [million cells/ml]
S [mg/l]
time [h]
X
S
E

a)


Training Data (Solid Line) and ANFIS Prediction (Dots) with RMSE = 5.5143e-005
Time Index
Checking Data (Solid Line) and ANFIS Prediction (Dots) with RMSE = 0.15897
Time Index

b)



c)

Figure 5. a) Experimental data for a fermentation process, carried out at the cultivation temperature $T = 13[°\mathrm{C}]$, agitation speed $n = 0$ [1/min] and aeration volume $QO_2 = 1$ [l/l.min]; b) Prediction of ethanol by the exhaustive search (experiment No. 12) on the basis of limiting substrate (S). The best predictors according the exhaustive search from 36 fuzzy models are $y(k-1)$, $y(k-3)$ and $y(k-4)$. The FIS-model accuracy is 0.1589; c) Optimal 3D-surfaces of the FIS-model for bio-ethanol (output) prediction by using the best predictors – regressors input1 ($y(k-1)$, input2 ($y(k-3)$ and input3 ($y(k-4)$).

The AI-based FDI leverage the tolerance for imprecision, uncertainty, and incompleteness, which is intrinsic to the problems to be solved, and generate tractable, low-cost, robust solutions to such problems. The synergy derived from hybrid systems stems from the relative ease with which we can translate problem domain knowledge into initial model structures whose parameters are further tuned by local or global search methods. The payoff for a conjunctive use of techniques is a more accurate and robust solution than a solution derived from the use of any single technique alone. This synergy comes at comparatively little expense because typically the methods do not try to solve the same problem in parallel but they do it in a mutually complementary fashion. In other words, no single technique should be expected to be the best for finding every model structure and tuning all system parameters.

However, knowledge driven systems, which also involve fuzzy systems, have limitations, as their underlying knowledge is usually incomplete. Sometimes, these systems require the use of simplifying assumptions to keep the problem tractable (e.g., linearization, hierarchy of local models, use of default values). Theoretically derived knowledge may even be inconsistent with the real systems behavior. Experiential knowledge, on the other hand, could be static, represented by a collection of instances of relationships among the system variables (sometimes pointing to causality, more often just highlighting correlation). The result is the creation of precise but simplified models that do not properly reflect reality or the creation of approximate models that tend to become stale with time and are difficult to maintain.

Data-driven methods also have their drawbacks, since data tend to be high dimensional, noisy, incomplete (e.g., DBs with empty fields in their records), wrong (e.g., outliers due to malfunctioning/failing sensors, transmission problems, erroneous manual data entries), etc. Some techniques, such as feature extraction, filtering and validation gates, imputation models, and virtual sensors (which model the recorded data as a function of others variables) have been developed to address these problems.

Acknowledgements

The research work reported in the paper is supported by the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).

REFERENCES

[1] Alcorta-García, E.—Frank, P. M.: Analysis of a class of dedicated observer schemes to sensor fault isolation. Proceedings of the UKACC International Conference on Control, Exeter, UK 1996, pp. 59–64.

[2] Amann, P.—Frank, P. M.: On Fuzzy Model-Building in Observers for Fault Diagnosis. Proceedings of the 15^{th} World Congress on Scientific Computations, Modelling and Applied Mathematics (IMACS 1997), Berlin, August 1997.

[3] Basseville, M.: Model-Based Statistical Signal Processing for Change and Damage Detection. Encyclopedia of Structural Health Monitoring, C. Boller, F.-K. Chang, Y. Fujino (Eds.), John Wiley & Sons, Chichester, GB, January 2009, pp. 677–696.

[4] Basseville, M.: On Fault Detectability and Isolability. European Journal of Control, Vol. 7, 2001, No. 6, pp. 625–638.

[5] Basseville, M.: Information Criteria for Residual Generation and Fault Detection and Isolation. Automatica, Vol. 33, 1997, No. 5, pp. 783–803.

[6] Basseville, M.—Nikiforov, I. V.: Detection of Abrupt Changes – Theory and Applications. Prentice Hall Series in Information and System Sciences, Englewood Cliffs 1993.

[7] Basseville, M.—Nikiforov, I.: Fault Isolation for Diagnosis: Nuisance Rejection and Multiple Hypotheses Testing. Annual Reviews in Control, Vol. 26, 2002, No. 2, 2002, pp. 189–202.

[8] Beard, R. V.: Failure Accommodation in Linear Systems Through Self-Reorganization. Ph. D. Dissertation, MIT Press, Cambridge, MA 1971.

[9] Blanke, M.: Consistent Design of Dependable Control Systems. Control Eng. Practice, Vol. 4, 1996, No. 9, pp. 1305–1312.

[10] Bonissone, P.—Badami, V.—Chiang, K. H.—Khedkar, P. S.—Marcelle, K.—Schutten, M. J.: Industrial Applications of Fuzzy Logic at General Electric. Proceedings of the IEEE, Vol. 83, 1995, No.3, pp. 450–465.

[11] Bonissone, P.—Chen, Y.-T.—Goebel, K.—Khedkar, P.: Hybrid Soft Computing Systems: Industrial and Commercial Applications. Proceedings of the IEEE, Vol. 87, 1999, No. 9, pp. 1641–1667.

[12] Chen, S.—Billings, S. A.—Grant, P. M.: Recursive Hybrid Algorithm for Non-Linear System Identification Using Radial Basis Function Networks. International Journal of Control, Vol. 55, 1992, No. 51, pp. 1051–1070.

[13] Chen, Y. T.—Bonissone, P.: Industrial Applications of Neural Networks at General Electric. Technical Information Series, 98CRD79, General Electric CRD, October 1998.

[14] Chow, E. Y.—Willsky, A. S.: Analytical Redundancy and the Design of Robust Failure Detection Systems. IEEE Trans. Autom. Control AC-29, 1984, No. 7, pp. 603–614.

[15] Diestel, R.: Graph Theory. 3rd ed. Springer-Verlag 2005, ISBN 978-3-540-26183-4, http://www.math.uni-hamburg.de/home/diestel/books/graph.theory/.

[16] Frank, P. M.: Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-Based Redundancy Survey. Automatica, Vol. 26, 1990, pp. 459–474.

[17] Frank, P. M.—Ding, X.: Frequency Domain Approach to Optimally Robust Residual Generation and Evaluation for Model Based Fault Diagnosis. Automatica, Vol. 30, 1994, No. 5, pp. 789–804.

[18] Frank, P. M.—Alcorta García, E.—Köppen-Seliger, B.: Modeling for Fault Detection and Isolation Versus Modelling for Control. Mathematics and Computers in Simulation, Elsevier 2000, pp. 259–271.

[19] Gertler, J.: Structured Residuals for Fault Isolation, Disturbance Decoupling and Modelling Error Robustness. Proceedings of the IFAC-On line Fault Detection and Supervision in the Chemical Process Industries 1992, pp. 111–119.

[20] Gomm, J. B.—Williams, D.—Harris, P.: Detection of Incipient Faults Using Approximate Parametric Models. Proceedings of the IFAC-On line Fault Detection and Supervision in the Chemical Process Industries 1992, pp. 138–143.

[21] Hamscher, W.—Console, L.—de Kleer,J.: Readings in Model-Based Diagnosis. San Francisco, CA, USA, Morgan Kaufmann Publishers 1992, ISBN 1558602496.

[22] Han, Z.—Frank, P. M.: Fault Detection via Characteristic Parameter Estimation. Proceedings of the 3rd European Control Conference, Italy 1995, pp. 378–382.

[23] Himmelblau, D. M.: Use of Artificial Neural Networks to Monitor Faults and for Troubleshooting in the Process Industries. Proceedings of the IFAC Symposium on On-line Fault Detection and Supervision in the Chemical Process Industry, Newark, Delaware, USA 1992.

[24] Hou, M.—Müller, P. M.: Fault Detection and Isolation Observers. International Journal of Control, Vol. 60, 1994, pp. 827–846.

[25] http://en.wikipedia.org/wiki/Corrective_and_preventive_action.

[26] http://energy.gov/management/office-management/operational-management/project-management/root-cause-analysis-rca.

[27] Isermann, R.: Process Fault Detection Based on Modeling and Estimation Methods – A Survey. Automatica, Vol. 20, 1984, pp. 387–404.

[28] Keller, E.: Qualitative Model Reference Adaptive Control. Ph. D. Thesis, Heriot-Watt University, Department of Computing and Electrical Engineering 1999.

[29] Köppen-Seliger, B.—Frank, P. M.: Fault Detection and Isolation in Technical Processes with Neural Networks. Proceedings of the 34th Conference on Decision and

Control (CDC), 1995.

[30] Lall, P.—Gupta, P.—Panchagade, D.—Angral, A.: Fault-Mode Classification for Health Monitoring of Electronics Subjected to Drop and Shock. Proceedings of the Electronic Components and Technology Conference 2009, DOI 10.1109/ECTC.2009.5074086, 2009, pp. 668–681.

[31] Massoumnia, M.: A Geometric Approach to the Synthesis of Failure Detection Fiters. IEEE Trans. Autom. Control AC-31, No. 9, 1986, pp. 839–846.

[32] Mileva, S.—Vassileva, S.: Knowledge-Based Analysis of CCP for Food and Beverages Safety and Quality. Journal Ecology & Safety, ISSN 1313-2563, Vol. 3, 2009, Part 1, pp. 441–447.

[33] Mileva, S.—Vassileva, S.: Fault Diagnosis and Uncertainty Management (Illustration to Biotechnological Processes). Proceedings of the International Bulgarian-Turkish Conference – Computer Science 2006, Vol. II, pp. 36–41.

[34] Mileva, S.—Vassileva, S.: Knowledge-Based Analysis of CCP for Food and Beverages Safety and Quality. Journal Ecology & Safety (ISSN 1313-2563), `www.sciencejournals.eu`, Vol. 3, Part 1, pp. 441–447.

[35] Mironovskii, L. A.: Functional Diagnosis of Dynamic Systems. Automation and Remote Control 1980, pp. 96–121.

[36] Narendra, K.—Parthasarathy, K.: Identification and Control of Dynamical Systems Using Neural Networks. IEEE Trans. Neural Networks, 1990, No. 1, pp. 4–27.

[37] Querelle, R.—Mary, N.—Kiupel, N.—Frank, P. M.: Use of Qualitative Modeling and Fuzzy Clustering for Fault Diagnosis. WAC 1996, Montpellier 1996, pp. 527–532.

[38] Qin, J.—McAvoy, T. J.: Nonlinear PLS Modeling Using Neural Networks. Computers and Chemical Engineering, Vol. 16, 1992, pp. 379–391.

[39] Sorsa, T.—Suontausta, H.—Koivo, N.: Dynamic Fault Diagnosis Using Radial Basis Function Networks. TOOLDIAG 1993, Toulouse 1993.

[40] Sjöberg, J.—Hjalmarsson, H.—Ljung, L.: Neural Networks System Identification. In Preprint 10th IFAC Symposium on Systems Identification 1994, Vol. 2, pp. 49–72.

[41] Sugeno, M.—Yasukawa, T.: A Fuzzy Logic Based Approach to Qualitative Modeling. IEEE Trans. on Fuzzy Systems 1993, No. 1, pp. 7–31.

[42] The NACMCF Critical Control Point Decision Tree 1992. `http://meat.tamu.edu/CCP_tree.html#top`.

[43] Tzafestas, S. G.—Dalianis, P. J.: Fault Diagnosis in Complex Systems Using Artificial Neural Networks. Proceedings of the 3rd IEEE Conference on Control Applications 1994, pp. 877–882.

[44] Vassileva, S.—Mileva, S.: Observer-Based Approach for Fault Detection and Isolation. Proceedings of the International Bulgarian-Turkish Conference – Computer Science'06, Vol. II, pp. 21–26.

[45] Vassileva, S.—Wang, X. Z.: Neural Network Systems and Their Applications in Software Sensor Systems for Chemical and Biotechnological Processes. In Intelligence Systems: Technology and Applications, C. T. Leondes (Ed.), CRC Press 2002, pp. 291–335.

[46] Vassileva, S.—Doukovska, L.—Mileva, S.: AI-Based Prediction and Diagnostic on Bioethanol Production. Proceedings of the 6th IEEE International Conference on Intelligent Systems 2012, ISBN 978-1-4673-2782-4, 2012, pp. 270–274.

[47] Wang, L.: Fuzzy Systems are Universal Approximations. Proceedings of International Conference on Fuzzy Systems 1992.

[48] White, J. E.—Speyer, J. L.: Detection Fiter Design: Spectral Theory and Algorithms. IEEE Trans. Automat. Control, Vol. 32, 1987, No. 7, pp. 593–603.

[49] Wünnenberg, J.: Observer-Based Fault Detection in Dynamic Systems. VDI-Fortschrittsber, VDI-Verlag, Reihe 8, Nr. 222, Düsseldorf 1990.

[50] Yu, D.—Shields, D. N.: A Fault Isolation Method Based on Parity Equations with Application to a Lathe-Spindle System. Proceedings of the UKACC International Conference on Control 1996, pp. 317–322.

Svetla Vassileva is working as Associate Professor in the Integrated Systems Department of the Institute of Systems Engineering and Robotics (ISER), Bulgarian Academy of Sciences (BAS). She obtained her M. Sc. in electrical engineering from the Czech Polytechnic in Prague in 1976, and defended her Ph. D. in the Bulgarian Academy of Sciences in 1990. She received her habilitation rom Bulgarian Academy of Sciences in 1997. She is a member of IFAC TC 8.3., reviewer of the Signal Processing journal and of WSEAS journals and conferences, guest-editor of WSEAS Transactions on Systems, editor of Bulgarian scientific journals, etc. She is interested in intelligent and knowledge-based systems, inferential measurements, intelligent sensor design, intelligent control and fault detection and isolation in industry and ecology.

Lyubka Doukovska is Associate Professor at the Bulgarian Academy of Sciences (BAS). She is the Head of the Intelligent Systems Department at the Institute of Information and Communication Technologies (IICT). She received her MSc in automation of industry and process control from the University of Chemical Technology and Metallurgy, Sofia, Bulgaria. She defended her Ph. D. in 2006 at the Bulgarian Academy of Sciences in signal processing in radar target detection theory. Since 2009, she has been Associate Professor at the Bulgarian Academy of Sciences. She is the author and co-authour of more than 90 scientfic papers. Her current research interests include signal processing, radar target detection, parallel algorithms design, intelligent systems, process control and optimization.

Vassil Sgurev obtained his M. Sc. from the Electromechanical Department of the Sankt-Peterburg State Polytechnic University in 1960. Later, he defended his Ph. D. at the same university. Later on he held the positions of Associate Professor; Professor, Corresponding Member and Academician of the Bulgarian Academy of Sciences. His main research interests are in operations research and optimization network flows, Markov decision processes, artificial intelligence, formal logic, etc. He has important contribution to organizational measures in science and, notably, to the implementation of personal computers in practice, especially in the former Soviet Eastern European block. He is a member and a senior member of IEEE, AAAI, IFIP, IFAC, IBRO, CEPIS and other international professional organizations.