

On-line fault detection and isolation using analytical redundancy

Drago Valh, Božidar Bratina, Boris Tovornik

*University of Maribor, Faculty for Electrical Engineering and Computer Science,
Smetanova ul. 17, 2000 Maribor, Slovenija
E-mail: bozidar.bratina@uni-mb.si*

Abstract. In this paper fault detection and isolation schemes using an extended Luenberger observer for non-linear systems and linear fault sensitive filters are presented. The idea is to implement both approaches on the same plant, achieve on-line fault detection and show some practical issues related to on-line fault detection implementation on a real laboratory plant, where for evaluation of residuals and fault isolation an adaptive threshold together with Boolean decision logic is used. FDI methods tested in this paper were performed on a most popular case study, namely the three-tank system, which has in our case an unusual structure as the inflow to the tanks is mounted at the bottom of the tanks and contributes to additional non-linear behaviour. On-line data acquisition was realized by local controller using Ethernet communication and OPC interface, and FDI schemes were performed in the Matlab/Simulink environment. The implemented fault detection schemes proved themselves well even when small abrupt changes/faults were generated in the real process.

Keywords: observers, fault detection and isolation, hydraulic systems.

Sprotno odkrivanje in izolacija napak z uporabo analitične redundance

Povzetek. V prispevku predstavljamo shemo odkrivanja in izolacije napak z linearnim in nelinearnim pristopom. Na laboratorijskem procesu smo realizirali metodo z razširjenim Luenbergerjevimi opazovalniki in filtri, občutljivimi na napake. Za sklepanje o napakah smo preizkusili fiksni in adaptivni prag v kombinaciji z Boolovo logiko, sheme pa izvedli na laboratorijskem testnem sistemu (modelu treh posod) s specifično strukturo, saj je izveden dotok v posode na dnu, kar vnese v proces dodatne nelinearnosti. Z uporabo lokalnega krmilnika, vmesnika OPC in komunikacije Ethernet smo sproti zajemali podatke v okolje Matlab/ Simulink, kjer je bil izveden algoritem odkrivanja in izolacije napak. Izvedene sheme so se izkazale za ustrezne tako pri večjih kot pri manjših napakah, nastalih v procesu.

Ključne besede: opazovalniki, odkrivanje in izolacija napak, hidravlični modeli.

1 Introduction

The area of fault detection and isolation (FDI) has undergone considerable development in the last few years [8], leading to a wide variety of model-based approaches. Still, the concept of analytical redundancy remains unchanged: deviation between values of measured system inputs and outputs and its model-based computed inputs and outputs generates fault

indicators called residuals. The goal of the fault-detection algorithm is to produce such residuals that are sensitive only to a single fault, thus ensuring fault isolation in the system. This can be achieved by using many developed techniques [1], [2], [3], [5], however an adequate model of the system must be obtained first.

Deriving an adequate model for a part of the modern production plant requires a lot of effort, since the processes are usually of a large scale and complex, running with a large number of measured signals. Such derived models are non-linear, implicit with respect to the set of output signals and reflect the static and dynamic properties of the system in the whole range of operation. Furthermore, their non-linear nature and usually a closed-loop control can cause problems with detection and isolation of certain faults in the system. Nevertheless, many non-linear FDI approaches based on non-linear models have been developed, however only few of them are really applicable and can not be simply put into practice. For that reason, the use of linear models has been preferred in practice and in many cases they still offer satisfying results. The idea of this paper is to implement both approaches on the same plant, achieve on-line fault detection and show some practical issues related to on-line fault detection implemented on a real laboratory plant.

The paper is structured as follows: theory on deriving model-based fault detection and isolation using an extended Luenberger observer for nonlinear systems and linear fault sensitive filters is presented in the

second chapter. In the third, an adaptive threshold for evaluation of residuals and fault isolation together with Boolean decision logic is described. Both on-line FDI schemes were successfully implemented to a laboratory three-tank plant which is described in chapter four. The obtained experimental results are presented and evaluated in the end.

2 Fault detection using analytical redundancy

Faults in technical processes are unavoidable. Different approaches in the time or frequency domain have been proposed to realize the fault detection, isolation and diagnostic task. In the case of a known model structure, the problem reduces mainly to an appropriate selection of experimental conditions and detection algorithm with respect to the faults to be identified [7]. Unfortunately, the industrial cases are usually limited and may not be allowed at all to manipulate a system in the production mode. Also the use of linear approaches is limited if the system is strongly non-linear [8].

In order to evaluate which of the FDI approaches are more suitable to implement to our laboratory plant from the engineering point of view, a non-linear and linear approach were tested. The tested FDI schemes were an extended Luenberger observer and fault-sensitive filters.

2.1 Extended Luenberger observer for non-linear systems

One should consider that many industrial processes are of a non-linear nature and thus can be described in the form below (1):

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), \theta_f, \theta_d), \quad x(0) = x_0 \\ y(t) &= h(x(t), u(t), \theta_{fs}) \end{aligned} \tag{1}$$

where $x(t)$ is the state, $u(t)$ is the input of the system, $y(t)$ is the output of the system, θ_f represents the actual system parameters of the failure-free case, θ_{fs} represents parameters in the output equation and θ_d is the unmodeled dynamics of the plant. When there are no faults present in the process, the above parameters are denoted as: θ_{f0} and θ_{fs0} respectively.

To overcome the problem of model uncertainties, a design of the nonlinear unknown input observers can be performed, still the task is in many cases non-trivial. One way to overcome this problem is to employ linearization-based approaches, but on the other hand it is known, that this solution works well only when there is no large mismatch between the linearized model and the non-linear behaviour of the system.

An alternative approach was a synthesis of an extended Luenberger observer for non-linear systems, where:

\hat{x}, \hat{y} are estimated values, x_m, y_m, u_m are measured values

$$\begin{aligned} \dot{\hat{x}} &= K \cdot (y_m - \hat{y}) + f(x_m, u_m, \theta_{f0}, 0) \\ r &= y_m - h(x_m, u_m, \theta_{fs0}) = y_m - \hat{y} \end{aligned} \tag{2}$$

The residuals are, as a matter of fact, differentiating filters with respect to the fault signal, where K represents the vector of parameters, experimentally obtained to achieve proper behaviour of the observer, and has to fulfill two tasks. Firstly, it determines the cut-off frequency of the filter. The noise is better suppressed with the lower cut-off frequency but the fault effect becomes slower visible; the detection algorithm should not become too stiff. Secondly, the parameters set the steady-state gain of the time derivative of the fault signal. That means the effect of a fault on the residual is bigger with smaller values of the components of vector K . Finally, a trade-off between robustness to unmodelled dynamics, parameter uncertainties, sensitivity to slow arising faults and proper detection of abrupt changes must be achieved.

2.2 Fault-sensitive filters

The theory of fault-sensitive filters assumes linear, time-invariant process model written in the state-space form. One can simply construct the fault-sensitive filters if the system is fully measurable and the state variables x can be uniquely expressed as a linear combination of the system outputs y . From the observability point of view, all states are directly measurable and the output C matrix is actually a unit matrix I . To obtain the fault sensitive filters, one has to derive an observer with such parameters, to be able to detect and localize parameter changes.

According to Figure 1, the only possibility to define the feedback matrix H , is that the residuals r offer a unique inference of possible faults. Therefore, if a

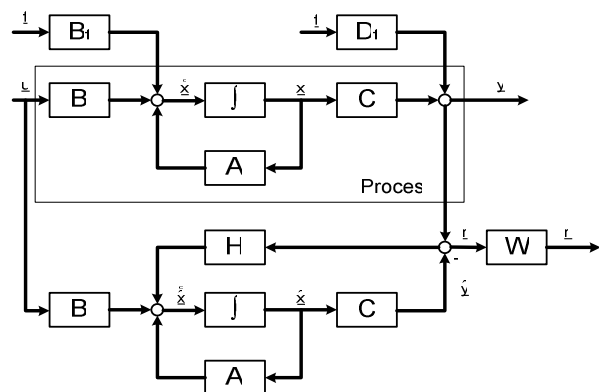


Figure 1. Full-order observer.
Slika 1: Shema opazovalnika polnega reda

change of parameters and output faults occurs, the faults should be properly described. While the fault vector f is not known, its influence can be modelled in the state-space form of the model.

$$\begin{aligned} \dot{x}(t) &= A \cdot x(t) + B \cdot u(t) + B_f \cdot f(t) \\ y(t) &= C \cdot x(t) + D_f \cdot f(t) \end{aligned} \quad (3)$$

If matrix C is considered to be regular, and $\text{Rank}(C)=n$, then n linear independent fault vectors can be formed. An output error is denoted as the residual:

$$\underline{r} = \underline{y} - \hat{\underline{y}} \quad (4)$$

Then the residual gets the following form:

$$\begin{aligned} \dot{r}(t) &= (C \cdot A \cdot C^{-1} - C \cdot H) \cdot r(t) + \\ &+ (C \cdot B_f - C \cdot A \cdot C^{-1} \cdot D_f) \cdot f(t) + D_f \cdot \dot{f}(t) \end{aligned} \quad (5)$$

There are n linear independent vectors available. If one needs more fault vectors (linear combinations), the system matrix of the residuals $C \cdot A \cdot C^{-1} - C \cdot H$ has to be diagonal and stable, fast and equal eigenvalues λ have to be chosen in order to obtain uniquely distinguishable residuals:

$$C \cdot A \cdot C^{-1} - C \cdot H = \lambda \cdot I \quad (6)$$

In that way the residuals become independent one from another and depend only on input and output signals.

$$\begin{aligned} r(s) &= \frac{1}{s-\lambda} \cdot [(s \cdot I - C \cdot A \cdot C^{-1}) \cdot y(s) - C \cdot B \cdot u(s)] = \\ &= \frac{1}{s-\lambda} \cdot [C \cdot B_f - (C \cdot A \cdot C^{-1} - I \cdot s) \cdot D_f] \cdot f(s) \end{aligned} \quad (7)$$

A faulty signal is filtered with $1/(s-\lambda)$ and affects the residuals in accordance with the output matrix C and fault matrix B_f and D_f . In addition, each element of the residual has the same dynamics and declination, but not the value of the fault. Therefore, a weight matrix W can be derived to achieve insensibility of a residual to certain fault. However, there are certain cases, when rows of B_f and D_f are linearly independent and one can isolate multiple faults. The eigenvalues need to be chosen negative and smaller than the smallest process eigenvalue. In that way the high-frequency noise can be suppressed. Nevertheless, λ affects the amplification of the residuals and needs to fulfill two tasks. Firstly, it defines the boundary frequency of the filter; the lower it is the better the noise is suppressed, however residuals become stiff. Secondly, λ defines the gain of the signal derivations, meaning that the influence of a fault on the residuals increases with smaller value λ .

3 Adaptive threshold function

Although unknown, the faults are assumed to be deterministic. Assuming that the noise has a zero mean,

the residuals have a time-varying mean contributed entirely by the faults and due to the actual difference between the model and real process; therefore they never have the zero mean value. This can be overcome by setting the threshold levels correspondingly higher or by using an adaptive threshold function.

Using the adaptive threshold as an alternative to the fixed one, or to another fault-isolation method [6] can sometimes significantly improve operation of the FDI scheme as it takes noise and model uncertainty into account. Incomplete knowledge about the process behaviour, disturbances and deviations, present unknown inputs to the system and can cause significant FDI problems. If the values of the threshold depend on the input process variable variations, problems caused

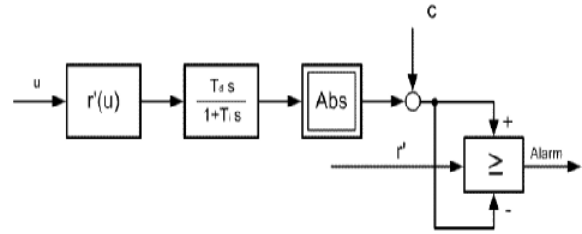


Figure 2. Adaptive threshold function.

Slika 2: Shema prilagodljivega praga

by the unknown inputs can be overcome. Figure 2 presents an adaptive threshold scheme, where inputs are driven by the system inputs. Function $r'(u)$ transforms process inputs ($Q_1(t)$, $Q_2(t)$) into residuals, where input signals affect the residuals and only then the adaptive threshold really depends on disturbances and unmodelled dynamics in the same way as the residuals do. Using a filter in addition, presents elimination of the mean value and sets the dynamics dependency of the input signals. The signal is then directed. This enables also a negative change of the input signal to influence the threshold. The c value is added to overcome the noise.

There are three parameters to be defined: T_d , T_i and c . T_i was chosen in accordance with the limit frequency of the process. The ratio $T_d/T_i=10$ was chosen since it offered promising results [10], [11]. The constant c was chosen by measurement of the maximum value of the noisy residual ($|r|_{max}$) when no faults were present. When Gaussian distribution of the stochastic noise is assumed, the standard deviation can be written as:

$$\sigma_r \approx \frac{1}{3} \cdot |r|_{max} \quad (8)$$

The threshold $c=k \cdot \sigma_r$ is defined by using the Chebyshev inequality:

$$P(-k \cdot \sigma_r \leq r(t) \leq k \cdot \sigma_r) \geq 1 - \frac{1}{k^2} \quad (9)$$

By choosing $k=4$ [4], the 93.75 % probability that no false alarm will be triggered in a fault-free case is achieved.

4 Implementation of FDI schemes to a three-tank laboratory plant

The process flowsheet of the three-tank laboratory plant is depicted in Figure 3. The upright tanks T_1 and T_2 are mounted above the tank T_3 , hence, the inflow of the upper tanks also depends on the level (hydrostatic pressure) in the tanks T_1 and T_2 , respectively (the pumps P_1 and P_2 are not an ideal generator to the system). The outlet pipes are mounted at the bottom of the tank T_3 , hence the amount of water in tank T_3 affects the outlet and the inlet flow of the tanks T_1 and T_2 . The nonlinear model was derived from the mass balance equations considering the Torricelli's rule. The model can be conveniently represented as:

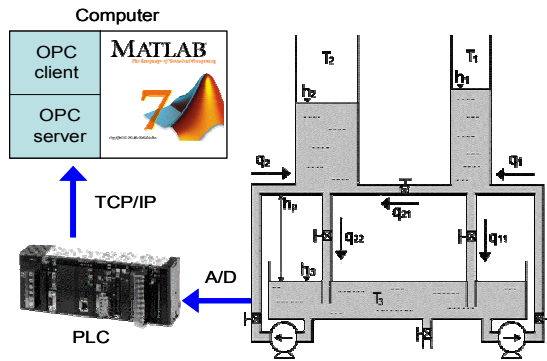


Figure 3. Three-tank laboratory plant and data acquisition.
Slika 3: Laboratorijski model treh posod in zajem podatkov

$$\begin{aligned} A_1 \frac{dh_1}{dt} &= q_1 - q_{21} - q_{11} \\ A_2 \frac{dh_2}{dt} &= q_2 + q_{21} - q_{22} \\ A_3 \frac{dh_3}{dt} &= q_{22} + q_{11} - q_1 - q_2 \end{aligned} \quad (10)$$

Where A_i denotes cross-section of the tank, h_i level in the tank and q_{ij} tank volume inflow or outflow, respectively. The medium in the tanks is fluid, which is taken as ideal and incompressible, therefore the specific density of the medium can be neglected (V denotes volume, g denotes gravity constant).

$$q_{in} - q_{out} = \frac{dV}{dt} = A \frac{dh}{dt} \quad (11)$$

The outlet of the tank is defined by:

$$q_{ij} = S_{vi} \cdot k_{vi} \cdot \text{sign}(h_i - h_j) \cdot \sqrt{2 \cdot g \cdot |h_i - h_j|} \quad (12)$$

From Eq. (12) it is obvious that the process is of a non-linear nature and (10) describes the non-linear model of the laboratory plant.

Although the process may not represent a full production-type industrial plant, it exhibits

characteristics that are common in nature (non-linearities, multiple inputs and outputs, noisy measurements, etc).

For both schemes the table of possible faults that could be brought about in the process, were put down (Table 1). The faults had to be modelled properly and

Fault	Fault description	Group
f_{h1}	Bias of the level 1 (h_1) sensor	Sensor fault
f_{h2}	Bias of the level 2 (h_2) sensor	Sensor fault
f_{v1}	Clog of the V1 valve	Component fault
f_{v21}	Clog of the V21 valve	Component fault
f_{v2}	Clog of the V2 valve	Component fault
f_{Q1}	Leakage of tank 1	Component fault
f_{Q2}	Leakage of tank 2	Component fault
f_{Q3}	Leakage of tank 3	Component fault
f_{P1}	Clog in pump/pipeline 1	Actuator fault
f_{P2}	Clog in pump/pipeline 2	Actuator fault

Table 1. List of possible faults in the process.

Tabela 1: Seznam možnih napak v procesu

then included in model equations. In the presented paper only some of them were used to test functionality and performance of FDI schemes.

The D_f matrix takes into account that the third state variable (level h_3) is not being measured but calculated from the mass (volume) balance of the system.

$$h_3 = \frac{V - h_1 \cdot A_1 - h_2 \cdot A_2}{A_3} \quad (13)$$

From (13) follows that the »measured« value of h_3 , h_{3m} is indeed affected by f_{h1} and f_{h2} :

$$h_{3m} = h_3 - \frac{A_1}{A_3} \cdot f_{h1} - \frac{A_2}{A_3} \cdot f_{h2} \quad (14)$$

The obtained residuals are functions of described faults:

$$\begin{aligned} r_1 &= r_1(f_{h1}, f_{h2}, f_{P1}) \\ r_2 &= r_2(f_{h1}, f_{h2}, f_{P2}) \\ r_3 &= r_3(f_{h1}, f_{h2}, f_{P1}, f_{P2}) \end{aligned} \quad (15)$$

Where f_{h_i} denotes the faults of level sensors and f_{P_i} denotes the faults of the pumps, respectively. In the end, one must form a weight matrix W to obtain a complete isolability of the faults. The following dependencies to the faults were chosen:

$$\begin{aligned} r_1' &= r_1'(f_{h1}, f_{P1}, f_{P2}) \\ r_2' &= r_2'(f_{h2}, f_{P1}, f_{P2}) \\ r_3' &= r_3'(f_{h1}, f_{h2}, f_{P1}) \end{aligned} \quad (16)$$

Where $r' = W \cdot r$ and finally (16) gives the following isolation matrix:

	r_1'	r_2'	r_3'
f_{h1}	1	0	1
f_{h2}	0	1	1
f_{P1}	1	1	1
f_{P2}	1	1	0

Table 2. Isolation (incidence) matrix.
Tabela 2: Izolacijska (incidenčna) matrika

4.1 Data acquisition

By implementing FDI methods to a real process one must consider that results highly depend on quality of data acquisition and data extraction from the noise correlated signals. In order to set up an optimally modern real industrial environment, an OPC standard (OLE for Process Control) was used. The laboratory model was controlled locally by a PLC and touch-screen display, while the process variables (inputs and outputs of the model) and FDI schemes were processed in Matlab/Simulink software. Communication between PLC and PC (Matlab) was realized by the TCP/IP protocol, PLC's OPC server and Matlab's OPC client.

The sampling time was due to limitations of the OPC server limited to 100 mS. In order to achieve soft real-time, the Simulink's time was locked to the OPC data acquisition time of 100 mS by a real-time simulation block capable of monitoring and reserving the CPU time for execution of the FDI scheme. The simulation time was then identical to the real-process operation time. [9]

4.2 Detection of faults using derived FDI schemes

The performance of the FDI schemes was evaluated by several fault cases introduced to the three-tank laboratory plant. The following faults were introduced: f_{h1} and f_{h2} – displacement of the level sensors (they were

separately displaced for approximately 2-4%, and f_{P1} and f_{P2} – pipeline of the pumps P_1 and P_2 were partially clogged. All tested faults were abruptly brought about

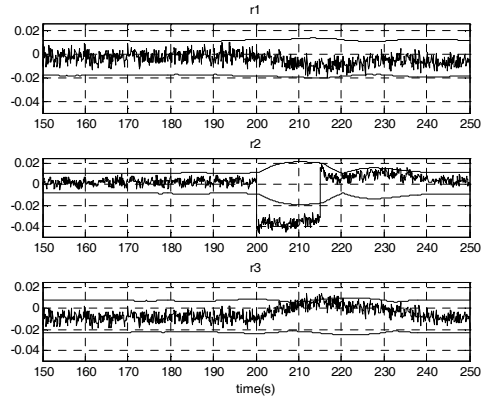


Figure 5. Residuals at small displacement of the level 2 sensor between 200 and 215 seconds.

Slika 5: Potek residuumov pri majhni prestavitvi senzorja nivoja v drugi posodi med 200 in 215 sekundami

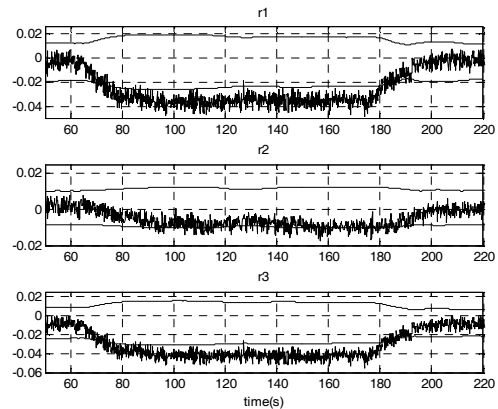


Figure 6. Pipeline of the pump P_1 was partially clogged between 60 and 175 seconds.

Slika 6: Potek residuumov pri zamašitvi dovoda v prvo posodo med 60 in 175 sekundami

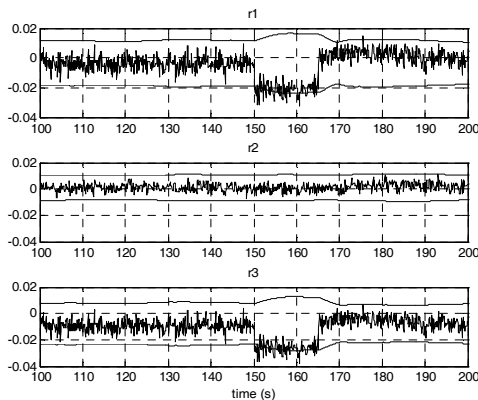


Figure 4. Residuals at small displacement of the level 1 sensor between 150 and 165 seconds.

Slika 4: Potek residuumov pri majhni prestavitvi senzorja nivoja v prvi posodi med 150 in 165 sekundami

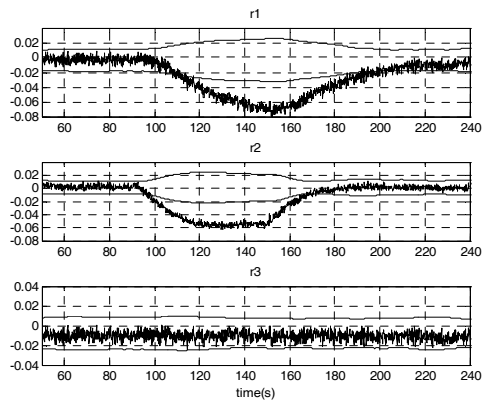


Figure 7. Pipeline of the pump P_2 was partially clogged between 90 and 150 seconds.

Slika 7: Potek residuumov pri zamašitvi dovoda v drugo posodo med 90 in 150 sekundami

and no multiple faults were predicted or tested. The situation in Figure 4 is clear as level h_1 (being actually a faulty signal) enters all three equations of the model and thus generates residuals, derived from that model. When the same fault is introduced to the level sensor 2, a similar result appears. That means that in this case those two failures can be uniquely identified without acquiring additional symptoms of the system. The matrix W causes the residuals r_2 and r_1 to be unaffected by the faults f_{h1} and f_{h2} respectively.

5 Conclusion

In this paper the use and implementation of a non-linear and linear FDI scheme is presented on a laboratory process plant. The schemes were realized in Matlab/ Simulink by Ethernet TCP/IP communication, which is becoming more and more used in industrial environments and together with the OPC standard present a powerful communication solution.

Firstly, the FDI scheme was performed with no linearization, as the extended Luenberger observer enables a simple approach not requiring a lot of processing time. Next, the implemented linear fault sensitive filters were used, though rarely used in practice, they proved well and offered good performance. According to the observed laboratory plant specifics, the non-linear approach showed better results on the plant inputs hence the mathematical model was not completely known (un-modelled dynamics of the pumps). The linear approach, on the other hand, proved well with the output faults where mathematical equations adequately described behaviour of the plant outputs.

Still, due to the noise and un-modelled dynamics of the models, an adaptive threshold function was used which was derived so as to be insensitive to process inputs. This means that only faults introduced to the plant caused the residuals to cross the threshold limit. By using the adaptive instead of the fixed threshold, more reliable fault detection can be achieved with approximately 6-7% of the false alarm rate if the thresholds are set properly.

Both approaches in combination with the adaptive threshold gave good results and a detection of only 2% of the abrupt fault was achieved at the outputs. From Figures 4-7 it is obvious that the residuals respond according to the predefined isolation matrix (Table 2).

Literature

- [1] Basseville, M., & Nikiforov, I. V., "Detection of abrupt changes: theory and applications", Englewood Cliffs, NJ: Prentice Hall.
- [2] Chiang, L. H., Russell E. L., Braatz R. D., "Fault Detection and Diagnosis in Industrial Control", New York: Springer-Verlag, 2001.

- [3] Patton, R., Frank, P., Clark, R. (Ed), "Fault Diagnosis in Dynamic Systems", Prentice Hall, New York, 1989.
- [4] Höfling, T., "Methoden zur Fehlererkennung mit Parameterschätzung und Paritätsgleichungen", Reihe 8: Meß-, Steuerungs und Regelungstechnik, Nr. 546, VDI-Verlag, Düsseldorf 1996.
- [5] Korbitz, J.K., Kościelny J.M., Kowalczyk, Z., Cholewa W., "Fault Diagnosis", Springer-Verlag 2004.
- [6] Rakar A, Juričić Đ., Balle P., "Transferable Belief Model in Fault Diagnosis", Engineering Applications of Artificial Intelligence No. 12 (5), p.p. 555-567, 1999.
- [7] Valh, D., "Fault Detection of Industrial Processes", Master Thesis, Polytechnic Nova Gorica, 2000.
- [8] Venkatasubramanian, V., Raghunathan, R. Yin, K., Kavuri, S., "A review of process fault detection and diagnosis", Computers and Chemical Engineering, 2002, Part I, Part II, Part III.
- [9] Peršin, S., Tovornik B., "Real-time implementation of fault diagnosis to a heat exchanger", Control Engineering Practice, No. 13, p.p. 1061-1069, 2005
- [10] Valh, D., Bratina, B., Tovornik, B., "Real-time implementation of fault sensitive filters to a three-tank laboratory plant", Proceedings of the IEEE EUROCON 2005 Conference, p.p. 1562-1565, Belgrade, Serbia & Montenegro, 2005.
- [11] Valh, D., Bratina, B., Tovornik, B., "Fault detection using non-linear model-based approach", Proceedings of the IEEE International Symposium on Industrial Electronics, p.p. 211-216, Dubrovnik, Croatia, 2005

Drago Valh was born in 1973 in Maribor, Slovenia. In 1997 he graduated from the Faculty of Electrical Engineering of the University of Maribor and in 2000 he received his M.Sc. degree from the Polytechnics, Nova Gorica. His field of research interests includes Fault Detection and Accommodation in Industrial Processes.

Božidar Bratina was born in 1978 in Celje, Slovenia. In 2004 he graduated from the Faculty of Electrical Engineering of the University of Maribor and is now working towards his Ph.D. degree. His field of research interests includes Fault Detection and Isolation, Fault Diagnosis and Fault Recovery.

Boris Tovornik was born in 1947 in Maribor, Slovenia. In 1974 he graduated from the University of Ljubljana. In 1984 and 1991 he received his M.Sc. and Ph.D. degrees in Electrical Engineering from the University of Maribor where he is currently an Associate Professor. His field of research interests includes Computer Control of Industrial Processes, Modelling and Process Identification, Fuzzy Control, Intelligent Systems, Fault Detection, Supervision and Safety.