Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Optimized Sensor Network Design for Process Monitoring based on Independent Component Analysis

Karim Salahshoor*, Fariborz Kiasi*

**Department of Automation and Instrumentation,*
Petroleum University of Technology, Tehran, Iran (email: salahshoor@ put.ac.ir, fkiasi@put.ac.ir)

Abstract: In this paper, a new sensor network design methodology has been proposed to improve the performance of ICA-based process monitoring approaches. Design procedure incorporates sensor cost, fault detectability and fault detection rate in the design formulation. The design problem has been transformed into an optimization problem. A genetic algorithm (GA) solver has been employed to yield optimal sensor locations for improved process monitoring. The proposed design methodology is evaluated on the Tennessee Eastman (TE) challenge benchmark problem. The simulation results demonstrate the effectiveness of the design procedure to enhance the process monitoring tasks with the less number of sensors for ICA approach.

Keywords: sensor network design, process monitoring, fault detection, independent component analysis (ICA), dynamic independent component analysis, GA

1. INTRODUCTION

Statistical process monitoring provides an efficient data analysis approach to identify any unusual conditions in chemical process operation and control. Process plants now routinely have large volumes of historical data. The exploitation of these data is a critical task in successful operation of any industrial process over long term. For more than a decade, considerable research efforts have been focused on developing statistical monitoring methods which can extract original information sources inherent in the multivariate measured data. Independent component analysis (ICA) is a recently developed statistical method for revealing independent latent variables that drive a process but are not directly measurable. Therefore, it is natural to infer that monitoring based on ICA model can give efficient process monitoring. On the other hand, whenever a process encounters a fault, the effect of this fault is propagated to all or some of the process variables. Efficient observation of these fault symptoms through a measurement system to determine the root cause for the observed behaviour necessitates a well-designed sensor network. Thus, the efficiency of the diagnostic monitoring system depends critically upon the location of the sensors to monitor the most fault relevant process variables. With hundreds of process variables available for measurement in any chemical plant, selection of crucial and optimum sensor positions poses a unique problem. A number of applications of the ICA-based approaches have been reported in the literature. Lee et al. (2003) proposed a new statistical monitoring method that uses ICA methodology. The method is, however, based on the assumption of an already available sensor network in place. To the best of out knowledge, the relationship between sensor location and ICA-based monitoring performance has not addressed yet in the literature. The emphasis of most of the works on the fault monitoring has been more focused on procedures to perform fault detection and diagnosis for a given set of sensors and less on the actual location of sensors for efficient monitoring and identification. Wang et al. (2002) has utilized the digraph-based approach proposed by Raghuraj et al. (1999) to optimize sensor locations for improved PCA-based monitoring in fault detection and isolation. Recently, Musulin et al. (2004) have presented a design method for sensor placement to improve fault detection monitoring based on principle component analysis (PCA) approach. Their design method formulates an optimization problem to minimize the sum of sensor cost and cost due to the relative impact of faults to locate the sensors. The proposed method, however, does not consider the cost due to the detection rates of faults. In this paper, a sensor network design method is presented for ICA-based monitoring approaches. The proposed method incorporates the sensor cost, the cost due to the relative impact of different faults and the cost due to the fault detection rates in the criteria of locating the sensors. In this way, design procedure is formulated in terms of an optimization problem. The inclusion of fault detection rates in the design procedure enhances the detectability of faults in the designed sensor network. The optimization problem is then solved using a genetic algorithm (GA). Finally, the effectiveness of the proposed design method is validated on the Tennessee Eastman (TE) challenge process.

2. ICA PROCESS MONITORING

ICA is a statistical technique to reveal hidden independent components (ICs) that underlie sets of process measurements. In this technique, it is assumed that the observed d-dimensional measured data at time k, i.e. $x(k)=[x_1(k),...,x_d(k)]^T$, can be expressed as a linear combination of m $m \le d$ unknown ICs, i.e. $s(k)=[s_1(k),...,s_m(k)]^T$, given by the following model:

$$x(k) = As(k) + e(k) \tag{1}$$

Where $A \in R^{d \times m}$ is the unknown mixing matrix and *e(k)* denotes the residual model error. The objective of ICA is to find a demising matrix W so that the components of the reconstructed data vector $\hat{s}(k) = Wx(k)$ become independent of each other. W can be calculated by $W = B^T Q$, where B is an orthogonal matrix (i.e., $BB^T = I$) and Q is the whitening matrix, given by $Q = D^{-1/2}V^T$, where V is the orthogonal matrix of eigenvectors of the data covariance matrix $R_x = E(x(k)x^T(k))$ as its columns, and D is the diagonal matrix of its corresponding eigenvalues. Lee et al. (2003) proposed three statistic measures (I^2, I_e^2, SPE) for ICA-based process monitoring, defined as follows:

$$I^2(k) = \hat{s}_{newd}(k)^T \hat{s}_{newd}(k) \tag{2}$$

$$I_e^2(k) = \hat{s}_{newe}(k)^T \hat{s}_{newe}(k) \tag{3}$$

$$SPE(k) = (x_{new}(k) - \hat{x}_{newd}(k))^T (x_{new}(k) - \hat{x}_{newd}(k)) \tag{4}$$

Where:

$$\hat{x}_{newd}(k) = Q^{-1}B_d \hat{s}_{newd}(k) = Q^{-1}B_d W_d x_{new}(k) \tag{5}$$

Where W_d represents dominant part of W and $B_d = (W_d Q^{-1})^T$ while W_e and $B_e = (W_e Q^{-1})^T$ indicate their excluded counterparts. Thus, the independent data vectors, $\hat{s}_{newd}(k)$ and $\hat{s}_{newe}(k)$, can be obtained when new data, $x_{new}(k)$, becomes available at sample time k through the demixing matrices W_d and W_e as follows:

$$\hat{s}_{newd}(k) = W_d x_{new}(k) \tag{6}$$

$$\hat{s}_{newe}(k) = W_e x_{new}(k) \tag{7}$$

Similar to the above reasoning for the *SPE* statistic, another new statistic measure can be proposed as follows to take care of monitoring the excluded part of the ICs:

$$SPE_e(k) = (x_{new}(k) - \hat{x}_{newe}(k))^T (x_{new}(k) - \hat{x}_{newe}(k)) \tag{8}$$

Where $\hat{x}_{newe}(k) = Q^{-1}B_e W_e x_{new}(k)$. Once the ICA model has been developed in terms of the four statistics (I^2, I_e^2, SPE and SPE_e), any departure from the process normal operating condition can be detected using corresponding confidence limit values, adjusted similar to the method adopted by Chiang et al. (2001). In this wok, FastICA algorithm, presented by Hyväarinen (1999), is used to perform ICA which entails maximizing the negentropy under the constraint of $\|b_i\| = 1$ (b_i is the ith column of matrix B*)*.

3. DERIVATION OF FAULT DETECTION LIMITS FOR ICA MODEL

Wang et al. (2002) derived sufficient conditions so that a fault is detectable in PCA-based monitoring approach. Muslin et al. (2004) utilized the same results in their sensor network design method for PCA-based monitoring application. The same reasoning is employed in this paper to derive the required conditions so that a fault can be detected in ICA-based monitoring approach. Suppose that there exists a set of J important faults, denoted by $\{F_j\}_{j=1}^J$, in the process to be monitored. Each fault can be described in a fault subspace ($F_j \in R^n$) by a set of orthogonal basis represented by $\theta_j \in R^{n \times l_j}$. Where l_j represents the number of process variables affected by the fault F_j. Thus, when a fault (F_j) occurs, the vector of measured deviated process variables x' can be expressed by:

$$x' = x_0 + \theta_j f_j \tag{9}$$

Where x_0 denotes the process measurements under normal operating conditions and $\theta_j f_j$ represents the induced fault deviation. In this formulation, each column of θ_j is zero except for the affected process variable which is 1 or -1 depending on the sign of deviation. $f_j \in R^{l_j}$ indicates the magnitude of the derivations caused by F_j in the corresponding process variable. Projecting every normalized measured data vector $x'(k)$ in the ICs space, leads to the ICA statistic model, defined by the following measures:

$$I^2(k) = \|\hat{s}_d(k)\|^2 = \|W_d x'(k)\|^2 \tag{10}$$

$$I_e^2(k) = \|\hat{s}_e(k)\|^2 = \|W_e x'(k)\|^2 \tag{11}$$

$$SPE(k) = \|x'(k) - \hat{x}'_d(k)\|^2 = \left\|x'(k) - Q^{-1}B_d W_d x'(k)\right\|^2 \tag{12}$$

$$SPE_e(k) = \|x'(k) - \hat{x}'_e(k)\|^2 = \left\|x'(k) - Q^{-1}B_e W_e x'(k)\right\|^2 \tag{13}$$

Consequently, the corresponding detection limit for I^2 can be evaluated using the following expression:

$$I^2 = \|W_d x'\|^2 = \|W_d(x_0 + \theta_j f_j)\|^2$$
$$= \|W_d x_0 + W_d \theta_j f_j\|^2 \tag{14}$$

Using the following inequality expression:

$$\|W_d x'\| \ge \|W_d x_0\| - \|W_d \theta_j f_j\| \tag{15}$$

and the expression $\|W_d x_0\| \le \delta_{I^2}$ which is valid for normal operating condition (where δ_{I^2} is the control confidence limit of I^2 statistic), the following sufficient condition for detectability of fault F_j based on I^2 measure can be concluded:

$$\|W_d x'\| > 2\delta_{I^2} \tag{16}$$

Using the foregoing inequality, it can be written:

$$\|W_d \theta_j\|.\|f_j\| \ge \|W_d \theta_j f_j\| \ge 2\delta_{I^2} \tag{17}$$

The result can be expressed in the following restrictive sufficient condition corresponding to the fault magnitude f_j in the I^2 statistic monitoring measure:

$$\|f_j\| \ge \left\|f_{I_j^2}\right\| = 2\sigma_{\max}^{-1}(W_d \theta_j) \times \delta_{I^2} \tag{18}$$

Following the same reasoning for the other statistics in the ICA-monitoring, leads to:

$$\|f_j\| \ge \left\|f_{I_{e_j}^2}\right\| = 2\sigma_{\max}^{-1}(W_e \theta_j) \times \delta_{I_e^2} \tag{19}$$

$$\|f_j\| \ge \left\|f_{SPE_j}\right\| = 2\sigma_{\max}^{-1}((I - Q^{-1} B_d W_d)\theta_j) \times \delta_{SPE} \tag{20}$$

$$\|f_j\| \ge \left\|f_{SPEe_j}\right\| = 2\sigma_{\max}^{-1}((I - Q^{-1} B_e W_e)\theta_j) \times \delta_{SPEe} \tag{21}$$

Where $\left\|f_{I_j^2}\right\|, \left\|f_{I_{e_j}^2}\right\|, \left\|f_{SPE_j}\right\|$ and $\left\|f_{SPEe_j}\right\|$ are the critical fault magnitudes (CFMs). That is, they represent the minimum fault magnitude $\|f_j\|$ detectable by the four statistic measures (I^2, I_e^2, SPE and SPE_e) for ICA-based process monitoring. $\delta_{I^2}, \delta_{I_e^2}, \delta_{SPE}$ and δ_{SPEe} represent the control confidence or threshold limits corresponding to the individual ICA statistic measures for normal operation.

4. FORMULATION OF SENSOR NETWORK DESIGN FOR IMPROVED ICA-BASED PROCESS MONITORING

The problem of sensor design location for improving the efficiency of fault detection in ICA-based process monitoring can be formulated in forms of an optimization problem. For this purpose, the following important design factors are considered in this work:

1- The sensor cost
2- The critical fault detectability condition determined by CFMs.
3- The fault detection rates determined by the ICA statistical charts.

Consider a process with N number of process variables to be measured. A binary vector Q_i is used to define ith sensor network design candidate, as follows:

$$Q_i = [q_1, q_2, ..., q_N] \tag{22}$$

Each element of Q_i will be either 0 or 1 to represent the absence or presence of sensor. Similarly, the cost of ith sensor network design can be defined by the following vector:

$$C_i = [c_{i1,} c_{i2}, ..., c_{iN}] \tag{23}$$

Where c_{ik} $(k = 1, ..., N)$ indicates the cost of kth selected sensor in the ith sensor network design. Thus, the sensor network cost (SNC_i) for the ith design solution can be obtained as follows:

$$SNC_i = \sum_{k=1}^{N} q_k c_{ik} \tag{24}$$

To evaluate the performance of the ith sensor network design candidate, the following minimum critical fault magnitude (MCFM) concept is defined for the jth fault:

$$MCFM_{i,j} = \min\left\{\left\|f_{I_{i,j}^2}\right\|, \left\|f_{I_{e_{i,j}}^2}\right\|, \left\|f_{SPE_{i,j}}\right\|, \left\|f_{SPEe_{i,j}}\right\|\right\} \tag{25}$$

Where $\left\|f_{I_{i,j}^2}\right\|, \left\|f_{I_{e_{i,j}}^2}\right\|, \left\|f_{SPE_{i,j}}\right\|$ and $\left\|f_{SPEe_{i,j}}\right\|$ represent the CFMs corresponding to jth fault (F_j) when ICA monitoring is performed based on the measured variables dictated by the ith sensor network design (Q_i). It is evident that $MCFM_{i,j}$ indicates the lower fault magnitude of F_j to be detected by any of the ICA statistic charts (I^2, I_e^2, SPE and SPE_e). A fault size penalization FSP_i is defined to infer the performance of the candidate sensor network (Q_i), as follows:

$$FSP_i = \sum_{j=1}^{J} FSP_{i,j} = \sum_{j=1}^{J} w_{i,j} MCFM_{i,j} \tag{26}$$

Where $w_{i,j}$ denotes a weighting factor which reflects the importance of fault F_j in process monitoring. The size of F_j varies between two limit bounds. The upper bound

$f_{\sup_{i,j}}$ determines the compulsory level at which the *j*th fault has to be detected by sensor network Q_i. Otherwise, the process encounters with shutdowns or hazardous situations while, the lower bound $f_{\inf_{i,j}}$ is considered as the lowest significant level to be monitored for fault detection. Thus, the $FSP_{i,j}$ can be modelled by the following expression:

$$FSP_{i,j} = \begin{cases} 0 & MCFM_{i,j} < f_{\inf_{i,j}} \\ w_{i,j}(MCFM_{i,j} - f_{\inf_{i,j}}) & MCFM_{i,j} > f_{\inf_{i,j}} \\ & \& \ MCFM_{i,j} < f_{\sup_{i,j}} \\ +\infty & MCFM_{i,j} > f_{\sup_{i,j}} \end{cases} \tag{27}$$

To calculate $f_{\sup_{i,j}}$, the process data matrices x_j should be inspected to determine the process data vectors when any measured process variable exceeds the well-known "3δ edit rules" (Chiang et al, 2003) as the normal pre-specified limits. Thus, considering x_j^0 as the process data vectors under such fault situation, its normalized version represents the relevant fault parameter vectors (f_j^0) at which the fault should be compulsory detected. Therefore, $f_{\sup_{i,j}}$ can be calculated by:

$$f_{\sup_{i,j}} = \left\| f_j^0 * Q_i \right\| \tag{28}$$

Where * indicates an element by element multiplication while, the lower limit $f_{\inf_{i,j}}$ can be obtained by considering the variance of individual columns of normalized data matrix under normal operating condition as follows:

$$f_{\inf_{i,j}} = \left\| [\sigma_1, ..., \sigma_N] * Q_i \right\| \tag{29}$$

Where σ_k denotes the variance of the *k*th column of normalized data matrix. So (31) is transformed to:

$$f_{\inf_{i,j}} = \left\| [1, ..., 1] * Q_i \right\| \tag{30}$$

Noting that each fault (F_j) should have a constant maximum fault size penalization ($MFSP_j$) which is defined as the fault size penalization ($FSP_{i,j}$) assigned to Q_i when $MCFM_{i,j} = f_{\sup_{i,j}}$. Therefore, the penalization weights can be defined as follows:

$$w_{i,j} = \frac{MFSP_j}{f_{\sup_{i,j}} - f_{\inf_{i,j}}} \tag{31}$$

Another important factor to be considered in the design procedure is due to the fault detection rate penalization ($FDRP_{i,j}$) which is defined as follows for the *j*th fault (F_j) in terms of an economic penalization term:

$$FDRP_{i,j} = [1 - \max\{I^2, I_e^2, SPE, SPE_e\}] \times \cos t_{d_{i,j}} \tag{32}$$

Where $\cos t_{d_{i,j}}$ represents the cost penalization of F_j with zero fault detection rate in all the ICA statistic measures provided by the candidate sensor network Q_i. Therefore, the sensor network design problem can be translated into the following objective function:

$$J_i = FSP_i + SNC_i + FDRP_i \tag{33}$$

in which $MSFP_j$ should be expressed in terms of an equivalent economic penalization. Thus, the sensor network design for improved ICA-based process monitoring is formulated as the following optimization problem:

$$\underset{i}{Minimize} \ \{J_i\} \tag{34}$$

5. OPTIMAL SENSOR NETWORK DESIGN USING GA ALGORITHM

A GA algorithm is used in this work to solve the optimization design problem, formulated by (36). In GA, the potential solution or individual is coded as a vector, called as chromosome. In the proposed algorithm, each chromosome denotes a sensor network design which is codified as Q_i. The individuals are then evaluated based on the penalty index (J_i). Thus, the individuals with lower J_i are selected as the most fitted solution in each generation. First, a set of N_{ind} candidate solutions or population are generated randomly. The goodness of each candidate solution or individual is evaluated by J_i as the fitness criterion. Then, a new generation of individuals is created from the most fitted chromosomes using the roulette-wheel selection, crossover and mutation operators (Goldberg, 1989). The GA algorithm can be stopped when the number of generations reaches a predefined maximum value or when the current population does not give sufficient improvement compared with the previous generation.

6. SIMULATION CASE STUDY

6.1 ICA-based process monitoring study

The proposed sensor network design approach for improving the ICA-based process monitoring is used in this simulation case study to investigate its effectiveness to monitor the Tennessee Eastman (TE) challenge process, shown in Fig.1, as a well-known plant-wide benchmark problem. The process consists of five major unit operations; a continuous stirred tank reactor (CSTR), a condenser, a compressor, a flash drum separator and a stripper. It involved the production of two main products, G and H and an undesired by-product F from four reactants A, C, D and E. More details on the process description are well explained in a book of Chiang et al. (2001). In this research study, the same simulation data generated by Chiang et al. (2001) has been used. A total of 33 variables including 11 manipulated variables and 22 measured variables have been selected to be used as monitoring variables (as listed in Table 1). Table 2 summarizes an estimation of the costs ($MFSP_j$)considered for all 18 TE faults excluding the faults 3, 9 and 15, charging more the faults which can lead to shut down situation. Table 3 lists the estimated cost for the process and the manipulated instruments except for the sensors that have already been used for control purposes which are considered of null cost.

The GA algorithm in the MATLAB GA Toolbox has been used in which the initialization parameters have been set as follows:

- Population size= 10*33=330
- Population type: Bit string
- Scaling function: Rank
- Selection function: Roulette-wheel
- Reproduction elite count= 120, crossover fraction=80%
- Mutation function: uniform with rate=0.02
- Crossover function: scattered
- Stall generations=5
- Function tolerance=$1
- Stall time limit=1e4 seconds
- Number of generations=50

Table 4 (left section) summarizes the ICA-based fault monitoring results for the designed sensor network with 19 selected sensors as specified in Table 5. The cost penalization of all faults ($\cos t_{d_{i,j}}$) has been assumed to be equal to $2*1e+6. The obtained losses for this design sensor network are 1.05842e+07 due to fault penalization, $13800 for sensor cost and $4554375 for detection rates. Comparing the obtained detection rates for all 18 TE faults in Table 4 with those of Lee et al. (2004) with 33 sensors indicates a significant achievement in fault detection with only 19 sensors. Fig. 2. demonstrates how the sensor network design has made the resulting monitoring approach to detect fault 5 clearly.

6.2 Filtered ICA-based (f-ICA) monitoring

The presence of noise in the data matrix often induces undesirable deviations in the measurements which might be attributed to the process faults in ICA-monitoring. To alleviate the noise influence, the following filtered versions of the original measured signals, stored in the normalized data matrices, can be employed:

$$x''(t) = mean[x'(t), x'(t-T), x'(t-2T), ..., x'(t-5T)] \quad (35)$$

The previous simulation study in subsection 6.1 was repeated for this filtered-ICA (f-ICA) version under the same settings. Table 4 (right section) summarizes the fault detection rates for the designed sensor network with 16 selected sensors as specified in Table 5. Evaluating the obtained results indicates much more improvement in the fault detection rates with respect to the ICA-based approach. Whereas, this result has been achieved with only 16 sensors compared to 19 sensors in the previous study. The estimate losses for this designed sensor network are $1.0317e+07 due to fault penalization, $10200 due to the sensor cost and $4409375 for fault detection rates. These outcomes indicate the capability of the f-ICA technique to yield better results with lower number of sensors and fewer losses. Fig. 3. illustrates the efficiency of the resulting monitoring system to detect fault 10 better than the previous test study.

7. CONCLUSIONS

The design of sensor location issue has not yet been paid enough attention in most process monitoring applications. To the best of our knowledge, this is the first contribution paper which addresses this important issue for the ICA-based process monitoring approaches. A new design methodology has been proposed which incorporates fault detection rate factor with both the sensor cost and fault detectability conditions. The design problem has been expressed in form of an optimization formulation. Then, a standard genetic algorithm (GA) has been employed to yield an optimal sensor network design solution. The simulation studies demonstrate the effectiveness of the proposed design scheme to improve the capabilities of the resulting ICA and f-ICA-based monitoring methods to detect TE faults with more efficiency and lower number of required sensors.

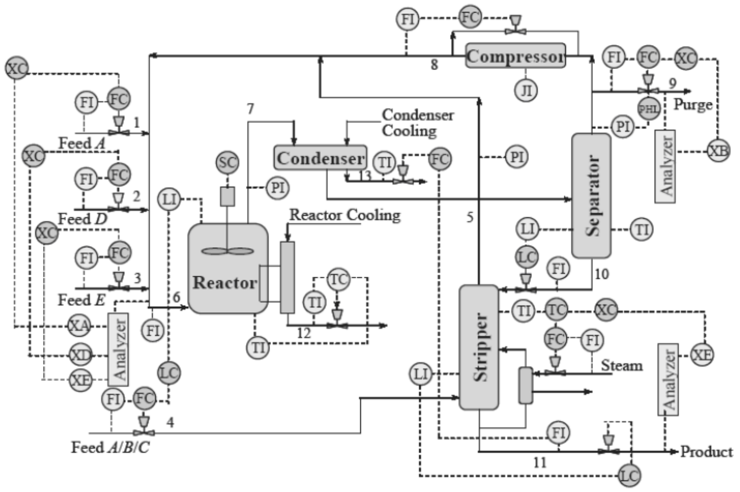

Fig. 1 Control system for the Tennessee Eastman process

Table 1. Process monitoring variables in TE Process

No.	Process Measurements	No.	Manipulated Variables
1	A feed (str.1)	23	D feed flow valve (str.2)
2	D feed (str.2)	24	E feed flow valve (str.3)
3	E feed (str.3)	25	A feed flow valve (str.1)
4	Total feed (str.4)	26	Total feed flow valve (str.4)
5	Recycle flow (str.8)	27	Compressor recycle valve
6	Reactor feed rate (str.6)	28	Purge valve (str.9)
7	Reactor pressure	29	Sep. pot underflow valve (str.10)
8	Reactor level	30	Stripper under flow valve (str.11)
9	Reactor temperature	31	Stripper steam valve
10	Purge rate (str.9)	32	Reactor Cooling water flow
11	Product sep. temp.	33	Condenser cooling water valve
12	Product sep. level		
13	Product sep. pressure		
14	Product sep. underflow (str.10)		
15	Stripper level		
16	Stripper pressure		
17	Stripper underflow (str.11)		
18	Stripper temperature		
19	Stripper steam flow		
20	Compressor work		
21	Reactor Cooling water outlet temp.		
22	Sep. cooling water outlet temp.		

Table 2. Estimation of the costs considered for all 18 TE faults

No	Fault	Type	Cost
1	A/C feed ratio, B composition constant (str.4)	Step	$7500
2	B composition, A/C feed ratio constant (str.4)	Step	$25000
3	D feed temp. (str.2)	Step	0
4	Reactor Cooling water inlet temp.	Step	$2500
5	Condenser cooling water inlet temp.	Step	$10000
6	A feed loss (str.1)	Step	$500000
7	C header press. Loss-reduced availability (str.4)	Step	$500000
8	A,B,C feed co position (str.4)	Random variation	$15000
9	D feed temp. (str.2)	Random variation	0
10	C feed temp. (str.4)	Random variation	$5000
11	Reactor Cooling water inlet temp.	Random variation	$15000
12	Condenser cooling water inlet temp.	Random variation	$35000
13	Reaction kinetics	Slow drift	$1000
14	Reactor cooling water valve	Sticking	$3750
15	Condenser cooling water valve	Sticking	0
16	Unknown		$24000
17	Unknown		$6000
18	Unknown		$500000
19	Unknown		$40000
20	Unknown		$500000
21	Unknown		$2500

Table 3. List of estimated cost of instruments

No	Process Meas.	Cost	No	MV	Cost
1	A feed (str.1)	0	23	D feed flow valve (str.2)	$700
2	D feed (str.2)	0	24	E feed flow valve (str.3)	$700
3	E feed (str.3)	0	25	A feed flow valve (str.1)	$700
4	Total feed (str.4)	0	26	Total feed flow valve (str.4)	$700
5	Recycle flow (str.8)	0	27	Compressor recycle valve	$700
6	Reactor feed rate (str.6)	$2100	28	Purge valve (str.9)	$700
7	Reactor pressure	$1600	29	Sep. pot underflow valve (str.10)	$700
8	Reactor level	0	30	Stripper under flow valve (str.11)	$700
9	Reactor temperature	0	31	Stripper steam valve	$700
10	Purge rate (str.9)	0	32	Reactor Cooling water flow	$700
11	Product sep. temp.	$600	33	Condenser cooling water valve	$700
12	Product sep. level	0			
13	Product sep. pressure	$1600			
14	Product sep. underflow (str.10)	$2100			
15	Stripper level	0			
16	Stripper pressure	$1600			
17	Stripper underflow (str.11)	0			
18	Stripper temperature	0			
19	Stripper steam flow	0			
20	Compressor work	$3000			
21	Reactor Cooling water outlet temp.	0			
22	Sep. cooling water outlet temp.	$600			

Table 4. Detection rates for ICA and ICA with filtered data matrices (f-ICA)

Faults	ICA I^2	ICA I_e^2	ICA SPE	ICA SPE_e	f-ICA I^2	f-ICA I_e^2	f-ICA SPE	f-ICA SPE_e
1	100	99.75	99.75	99.75	99.625	99.75	99.625	99.5
2	95.50	96.75	95.125	94.625	95.5	96.25	95	94
3	1.875	3.5	4.375	1.875	1.375	2.125	0.875	0.375
4	100	29.375	27.25	95.375	98.25	99.25	94.75	75.125
5	99.875	99.875	99.875	99.875	99.875	99.75	99.625	99.875
6	100	100	100	100	100	100	100	100
7	100	100	93	100	100	100	100	98.625
8	96.625	96.75	95.5	92.625	96.75	92.75	88.75	92.375
9	1.375	2.375	2.625	2.125	1.625	3.5	2.25	0.5
10	40.625	48.5	36.625	29.625	83.125	61.375	49.5	49.875
11	61.125	28.625	26.25	52.375	64	74.375	69	53.75
12	99.875	99.75	98.75	96.5	99.5	98.625	96	95.125
13	95	95	94.375	94.25	94.875	94.25	92.125	94.5
14	100	99.875	99.875	100	80.625	87.625	87.5	79.875
15	4	11.125	10.125	2.625	7.25	3.375	1.625	5.125
16	77.25	80.125	69.625	65	83.375	60.375	48.25	44.875
17	87.5	82	75.5	74.125	65.375	65	63.25	57.5
18	89.75	89.5	89	89.25	89.5	89.875	89.25	88.375
19	83.625	72.875	60.375	61.875	60.25	78.875	56	21.375
20	81.125	64.625	53.125	50.125	81	70.625	59.875	52.625
21	60.5	52	42.875	39.75	60.25	35.625	23.5	32.625
False alarm	1	1.125	1.125	1.125	1.125	1	1.125	1.125

Table 5. Sensor locations for ICA and ICA with filtered data matrices (ICA-f)

Sensor No.	**ICA**	**ICA-f**	**Sensor No.**	**ICA**	**ICA-f**
1	true	true	18	false	true
2	true	false	19	true	false
3	false	true	20	true	false
4	false	false	21	true	false
5	true	true	22	false	false
6	false	true	23	true	false
7	false	true	24	false	false
8	true	false	25	true	true
9	false	true	26	true	true
10	false	false	27	true	true
11	true	false	28	false	false
12	false	false	29	false	false
13	true	false	30	false	true
14	true	false	31	true	true
15	false	false	32	true	true
16	true	true	33	true	true
17	true	true			

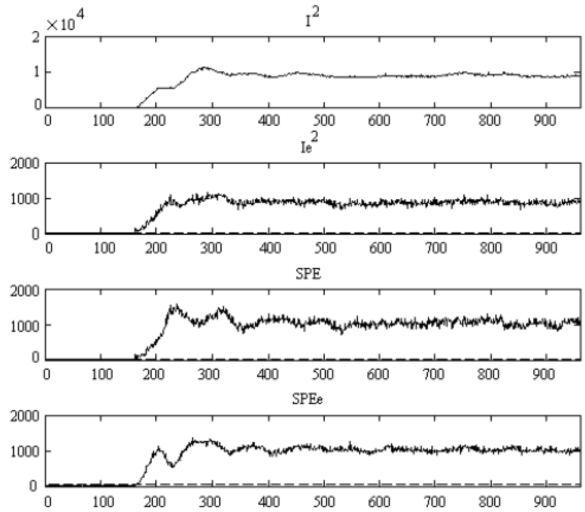

Fig. 2 Detection of fault 5 by ICA

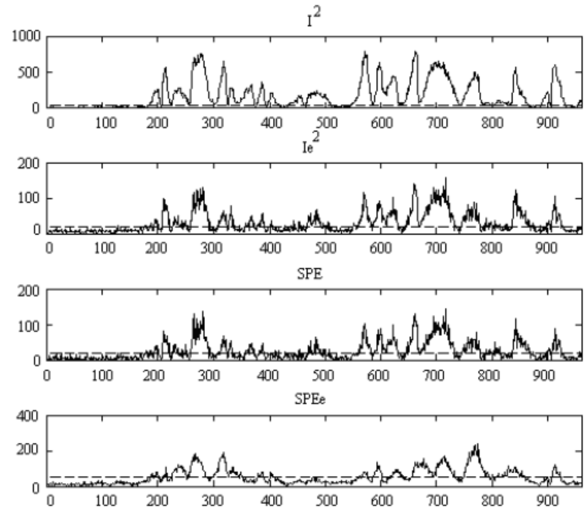

Fig. 3 Detection of fault 10 by f- ICA (filtered version)

REFRENCES

Chiang, L.H., E.L. Russell and R.D. Braatz (2001). *Fault Detection and Diagnosis in Industrial Systems.* Springer, London.

Chiang, L.H., R.D. Braatz (2003). Process monitoring using causal map and multivariate statistics: fault detection and identification. *Chemometrics and Intelligent Laboratory Systems*, **65**, 159–178.

Cristianini, N, J. Shawe-Taylor (2000). An introduction to support vector machines and other kernel-based learning methods. First edition (Cambridge: Cambridge university press)

Hyväarinen, A.(1999). Fast and robust Fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, **10**, 626–634.

Goldberg, D. (1989). *Genetic algorithms in search, optimization, and machine learning*; Addison-Wesley.

Lee, J.-M., C.K Yoo and I.-B Lee (2003). Statistical process monitoring with independent component analysis. *Journal of Process Control*, **14 (5)**, 467–485.

Lee, J.-M., C.K Yoo and I.-B Lee (2004). Statistical monitoring of dynamic Processes based on dynamic independent component analysis. *Chemical Engineering Science*, **59**, 2995 – 3006

Musulin E, M. Bagajewicz, J.M. Nougues, L. Puigjaner (2004). Instrumentation design and upgrade for principle components analysis monitoring. *Ind. Chem. Res*, **43**, 2150-2159.

Raghuraj, R., M. Bhushan, R.Rengaswamy (1999). Locating sensors in complex chemical plants based on fault diagnosis observability criteria. *AIChE J.*, **45**, 310-322.

Wang, H., Z. Song, H. Wang (2002) Statistical process monitoring using improved PCA with optimized sensor locations. *J. Process Control*, **12**, 735-744.