Proceedings of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Data-driven fault detection with process topology for fault identification

Brian S. Lindner*. Lidia Auret**

** Department of Process Engineering, Stellenbosch University, Private Bag X1, MATIELAND, 7602, South Africa*
*** Department of Process Engineering, Stellenbosch University, Private Bag X1, MATIELAND, 7602, South Africa (e-mail: lauret@sun.ac.za)*

Abstract: In this paper a fault diagnosis framework based on detection with feature extraction methods and identification based on data-driven process topology methods was investigated. A simulation of a simple system consisting of two tanks with heat exchangers was used to generate data for normal operating conditions and a number of faults. Fault detection methods included principal component analysis and kernel principal component analysis feature extraction with Shewhart, cumulative sum and exponentially weighted moving average monitoring charts. Process topology information was extracted with linear cross-correlation, partial cross-correlation and transfer entropy. Connectivity maps were constructed to identify possible fault propagation paths to aid root cause analysis and changes in connectivity structure due to faults were exploited for fault identification. Kernel principal component analysis with a CUSUM chart gave the best detection performance, while connectivity graphs based on partial correlation gave an accurate representation of the system and assisted fault identification.

Keywords: Fault Diagnosis, Process Monitoring, Feature Extraction, Connectivity, Root cause analysis

1. INTRODUCTION

Timely detection of abnormal behaviour caused by fault conditions in industrial process plants is essential to ensure optimal performance in a plant since it allows operators to take actions to correct the fault. These plants frequently record data from large amounts of variables that show a large degree of correlation. In this environment, data-based process monitoring methods can be used for fault detection and diagnosis. Feature extraction methods, such as principal component analysis (PCA), have been widely applied for fault detection (e.g. Kano et al., 2002). One of the possible limitations of PCA is that it is a linear method, so it may give inadequate results when applied to data from processes that show significant non-linear behaviour. Kernel principal component analysis (KPCA) is a non-linear feature extraction method that has also been widely applied for fault detection (e.g. Lee et al., 2004). KPCA works by first mapping the input space nonlinearly into a higher dimensional feature space, where the data is more likely to show linear behaviour, and then performing linear PCA in that feature space (Lee et al., 2004). Although these methods perform well for the detection of faults, their utility in specific fault diagnosis is limited. The high degree of interconnectivity in the data means that simple faults often propagate throughout the system. Information about process connectivity (also referred to as process topology) can be used to aid fault detection and isolation, since topology is altered by the presence of faults (Chiang and Braatz, 2003), and propagation paths for the fault can be identified and traced back to their origin (e.g. Bauer et al., 2007). Methods for capturing topology from process data include linear cross-correlation (LC) (Bauer and Thornhill, 2008), partial cross-correlation (PC) (Yang et al., 2011) and transfer entropy (TE) (Bauer et al., 2004).

This paper presents a framework for fault diagnosis using feature extraction based methods for detection combined with process topology methods for fault identification.

2. FAULT DETECTION

2.1 Feature extraction

Data from processes typically consist of a large amount of variables that are highly correlated. This correlation means that most of the variation in the data can be characterized by a small number of underlying features. Feature extraction methods identify and isolate these features, allowing the dimensionality of the data to be reduced. Figure 1 demonstrates the principle. A feature extraction method (FEM) is applied to a data set **X** consisting of N samples and m variables.

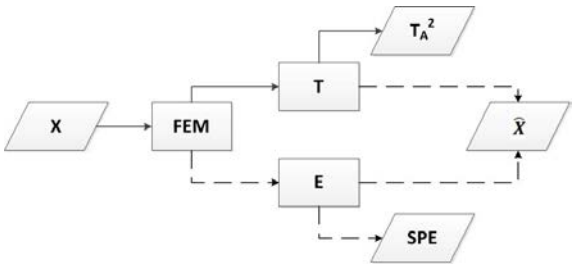

Fig. 1. Feature extraction (dashed lines do not apply to KPCA, since it does not allow reconstruction)

Two multivariate statistical feature extraction techniques are considered in this paper: PCA and KPCA, summarized in Table 1. The model is trained on normal operating conditions (NOC) data.

Table 1. Feature extraction methods (FEM's)

	PCA	KPCA
Reference	(Kourti and MacGregor, 1995)	(Lee et al., 2004).
Mapping	Projection onto components yielding maximum variation obtained from eigenvectors of covariance matrix. $\mathbf{t}_{i,new} = \mathbf{p}_i\mathbf{X}_{new}$	Intrinsic mapping to higher dimensional nonlinear feature space through use of kernel function k (e.g. Gaussian kernel); linear PCA performed in that space. $t_{i,k,new} = \frac{1}{\sqrt{\lambda^k}}\sum_{j=1}^{N}\alpha_j^k k(\mathbf{x}_j, \mathbf{x}_i)$
Feature space diagnostic metric	$(T_A^2)_j = \sum_{i=1}^{A}\frac{t_{i,j}^2}{\lambda_i^2}$	$(T_A^2)_j = \sum_{i=1}^{A}\frac{t_{i,j}^2}{\lambda_i^2}$
Residual space diagnostic metric	$(SPE)_j = \sum_{i=1}^{m}(X_{test,i,j} - \hat{X}_{test,i,j})^2$	$(SPE)_j = \sum_{i=1}^{n} t_{i,j}^2 - \sum_{i=1}^{A} t_{i,j}^2$

Table 2. Topology extraction methods (TEM's)

	Linear Correlation	Partial Correlation	Transfer entropy
Reference	(Bauer and Thornhill, 2008)	(Yang et al., 2011)	(Bauer et al., 2007)
Calculation	$\hat{\phi}(x,y) = \frac{1}{N-k}\sum_{i=1}^{N-k}\frac{(x_i-\mu_x)(y_{i+k}-\mu_y)}{\sigma_x\sigma_y}$	Linear correlation between two variables while conditioning on any number of the remaining variables	$t(x\|y) = \sum p(x_{i+h}, x_i, y_i)\log\left(\frac{p(x_{i+h}\|x_i, y_i)}{p(x_{i+h}\|x_i)}\right)$ $t_{x\to y} = t(y\|x) - t(x\|y)$

New data points being tested are then projected onto the feature space giving the scores, **T** (N samples by A features, where A is the dimensionality of the reduced feature space). The data can then be reconstructed from the features allowing the residuals, **E**, to be calculated. The modified Hotelling's T_A^2 diagnostic statistic allows detection of behaviour that results in deviation from the centre of the normal operating conditions in the feature space. The SPE diagnostic statistic allows detection of behaviour that results in deviation from the NOC manifold, which arises when the relationships between the process variables changes compared to the NOC case.

2.2 Monitoring charts

The above mentioned T_A^2 and SPE statistics can be monitored by plotting the value obtained at each sample point over time. This is known as a Shewhart chart. This method only uses information from the current observation and is therefore insensitive to slight shifts in the process. In order to account for previous observations in the historical data and thereby improve the ability to detect small shifts the cumulative sum (CUSUM) (Bin Shams et al., 2011) and exponentially weighted moving average (EWMA) (Prabhu and Runger, 1997) based charts have been applied for process monitoring. As the name implies, CUSUM calculates a cumulative sum of past values, usually summing the difference of the observations from the in-control mean, μ: $C_i^z = \mathrm{Max}(0, C_{i-1}^z + z_i - \mu_z)$ where z may represent either the T_A^2 or the SPE. EWMA sums past values, but gives progressively less weight to older data: $EWMA_i = rz_i + (1-r)EWMA_{i-1}$ where r indicates the weighting. A value of 0.1 is typically used

3. FAULT IDENTIFICATION

3.1. Use of connectivity for fault identification

Variables in a process are connected to each other through material or information flow. Information about how all the variables in a process are connected is known as the topology or connectivity of the process. This information can be useful for the isolation of faults since the presence of faults alters the connectivity structure and faults will propagate along paths that follow the material and/or information flow.

3.2. Topology extraction methods

Data-based techniques exist to extract topology information from process data. The connections between two variables can be inferred from historical process data by estimating the lags between their time series (Bauer and Thornhill, 2008) which results in maximum connectivity, where connectivity is some measure of linear or nonlinear correlation. Three methods were considered in this paper: linear correlation (LC), partial correlation (PC) and transfer entropy (TE). Table 2 gives a summary of the calculations used for these three methods. These are used to find significant connections between all pairs of variables in the data, giving connectivity matrices. These matrices can then be used to construct a connectivity map, where nodes represent variables and edges between them represent connections.

4. FAULT DIAGNOSIS FRAMEWORK

The proposed fault diagnosis framework is demonstrated in Figure 2.

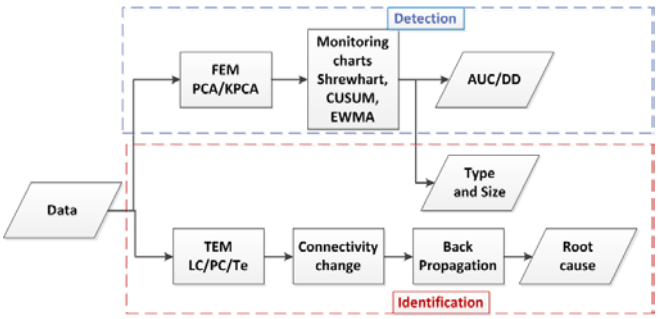

Fig. 2. Proposed fault diagnosis framework

A feature extraction method (FEM) is first applied to the data, giving the T_A^2 and SPE statistics. The monitoring charts are used to determine whether the fault was detected using the metrics described above. The SPE statistic can then be used to analyse the type of fault, i.e. whether the SPE follows

a ramp or a step trend, and determine a relative magnitude for the ramp or step. The topology extraction method (TEM) can then be applied, giving a connectivity map for normal operating conditions and the change in connectivity due to the presence of the fault. The variables most affected by the fault can be determined from the connectivity change and can be flagged as symptom nodes. The connectivity map can then be used to trace a path backwards from the symptom nodes to the root nodes.

In order to gauge the performance of the fault detection methods several metrics, shown in Table 3, are employed. The performance of the fault identification methodology depends on the ability to locate the area in the plant where the fault has occurred.

5. CASE STUDY DESCRIPTION

The chosen example consisting of two tanks was simulated in Simulink. A diagram of the system is shown in Figure 3.

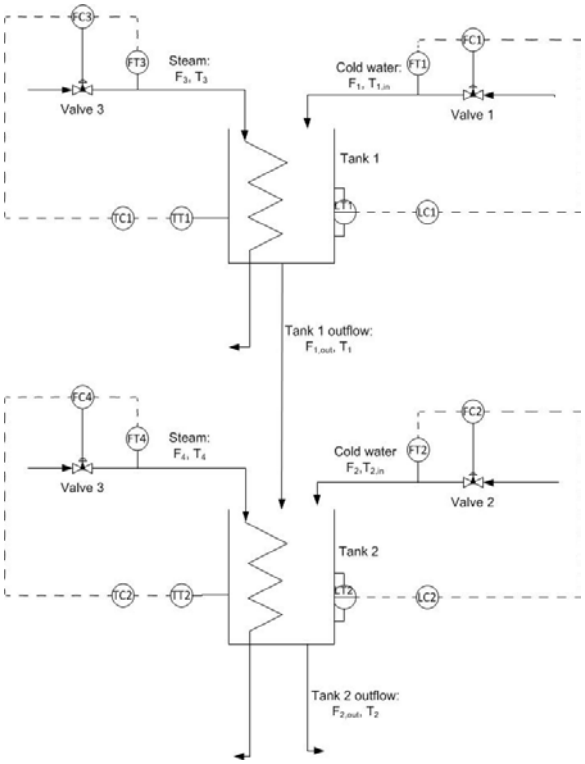

Fig. 3. Simulated two-tank system

The outlet flow from both tanks is proportional to the square root of the level in each tank. The outlet from the first tank flows into the second tank. Each tank has its own supply of cold water with a control valve to control the level of each tank. The temperature in the tanks is controlled using the control valves on the steam lines. The main variables of interest are the flow rate of the inlet streams to the tanks, F1 and F2, the flow rates of the steam in the heating coils in both tanks, F3 and F4, the levels of both tanks, L1 and L2, and the temperatures of both tanks, T1 and T2. The controllers used are simple proportional integral derivative (PID) controllers that change the values of the MVs according to the deviation of the controlled variables (CVs) from their set-points (SPs). The methodology was implemented on data from this simulation generated in closed loop.

The simulation was used to generate closed loop normal operating conditions (NOC) data as well as closed loop fault data. Four faults were simulated in this system. Fault 1: Step disturbance in temperature T1in. Fault 2: Ramp disturbance in temperature T1in. Fault 3: Step disturbance in temperature T2in. Fault 4: Fouling in both heat exchange coils. Fouling was simulated by changing the heat exchange constant. This disturbance was simulated as a ramp disturbance since fouling would accumulate over time in a real system.

6. METHODOLOGY

1) Data was divided into four subsets: NOC training data, NOC validation data, NOC testing data and fault data
2) FEM's (PCA and KPCA) were trained on NOC training data. Retention of components was based on an explained variance of 90%: three features for both PCA and KPCA. Cross validation was performed to determine that a Gaussian kernel width c = 0.5 gave the lowest SPE.
3) The NOC training SPE and T_A^2 were calculated. EWMA and CUSUM calculations were performed.
4) The trained FEM was applied to NOC validation data. Significance thresholds for SPE and T_A^2 were based on the 99th percentile of the diagnostic values NOC validation data.
5) The trained FEM was applied to NOC testing data and fault data, giving the T_A^2 and SPE.
6) FAR, MAR, DD and AUC were calculated for Shewhart, CUSUM, and EWMA for both diagnostics. The McNemar test (McNemar, 1947) for correlated proportions was used to compare the FARs and MARs for KPCA and PCA.
7) The SPE was analysed to determine the type and size of the fault.
8) TEM's were applied on the NOC and fault data, resulting in connectivity matrices.
9) The connectivity matrices resulting from the TEM's for NOC and fault data were compared to identify changed connections. For each node the number of changed connections was counted and the three nodes that showed the largest number of changes were selected as the symptom nodes for the input into the antecedent root cause method.
10) The NOC connectivity map was used to find the root nodes associated with the symptom nodes identified from the connectivity change. The ancestor nodes for each symptom node were determined, and the overall root node was identified as the common antecedent of all ancestor notes. The overall root node was identified as the most likely location for the origin of the fault.
11) All of the above steps were repeated for each fault and all fault sizes.

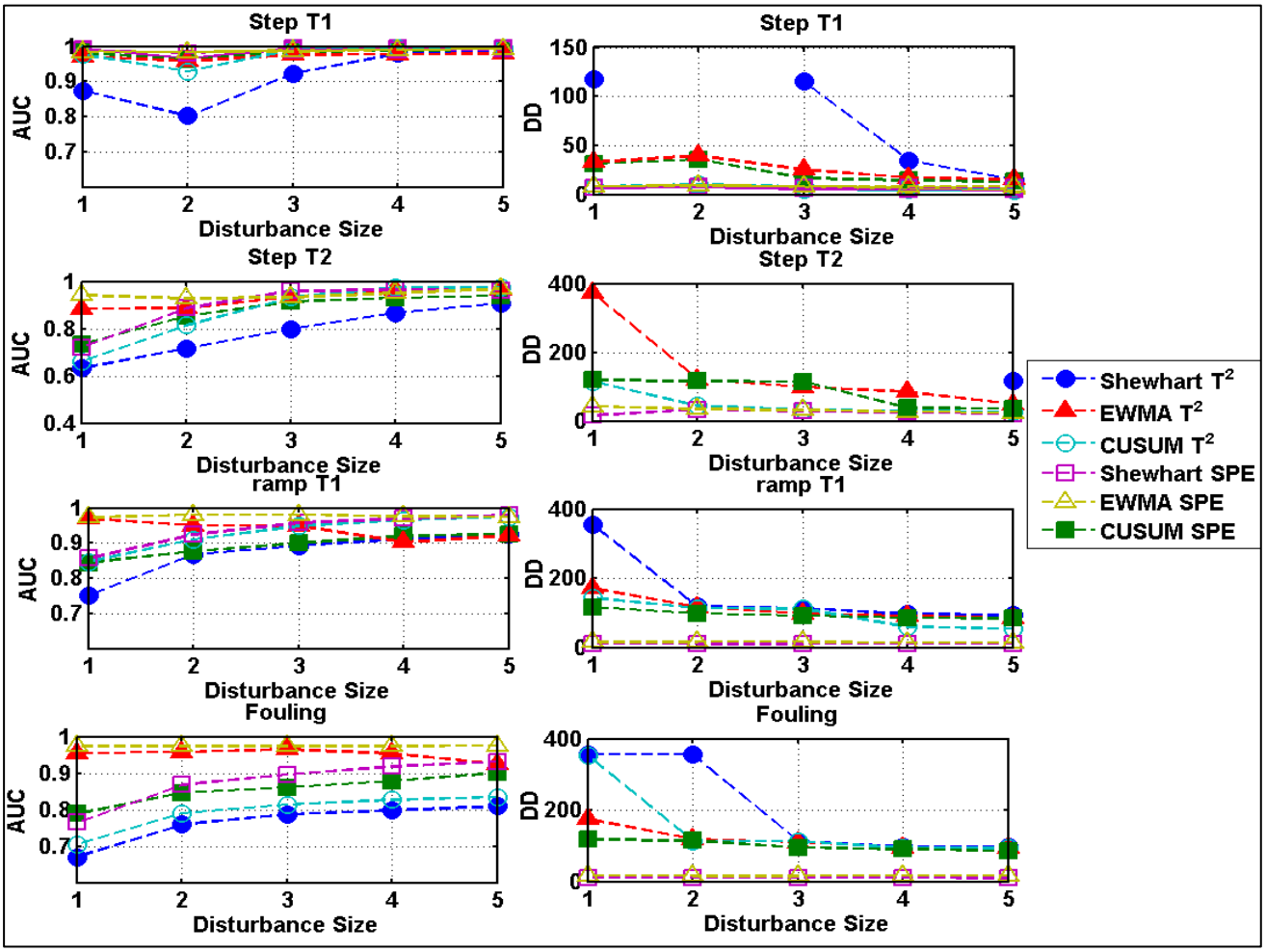

Fig. 4. AUCs (left) and detection delays (right) for different faults, fault sizes and different monitoring methods. Only KPCA results shown since it significantly outperformed PCA

7. RESULTS AND DISCUSSION

7.1 Detection

It was found that KPCA generally performed better than PCA for all the faults observed; showing much lower detection delays and higher AUCs. The McNemar test was used to compare the FAR and MARs for both methods and it was found that this difference in performance was statistically significant. Therefore only the results for KPCA are shown in Figure 4. For both the T_A^2 and SPE statistic the CUSUM chart performs the best for fault detection, with a higher AUC. The EWMA chart performs better than the Shewhart chart, but slightly worse than the CUSUM chart. However, the detection delays are generally larger for these charts in comparison with the Shewart chart, so the improved detection ability comes at the expense of detection speed.

7.2 Connectivity

Consider the NOC connectivity diagram derived by means of the partial correlation method, as shown in Figure 5. A sanity check verifies the calculated connections between the different variables: A change in F1 would cause a change in T1 due to energy conservation principals. A change in T1 would activate the first tank temperature controller, causing a change in steam flow F3. A change in steam flow F3 would affect T1, and through energy conservation in the first tank, affect T2. However, seemingly spurious connections are also identified: A connection between F4 and T1 is indicated, with implied causality from F4 to T1. This causality direction is impossible through either mass or energy conservation, or control action. Some expected connections are also omitted, for example the causality due to controller action between L2 and F2.

7.3 Fault identification

For the different faults and fault sizes, the fault identification method was applied, with the results shown in Table 4. The second column shows whether or not the fault identification method was able to correctly identify the type of the fault (step or ramp). The third column shows whether or not the size of the SPE change increased with increasing fault size. The method failed to detect that the fouling fault was a ramp type, but accurately identified the types of faults for the rest of the faults.

To illustrate the use of connectivity to find the root cause of a fault, consider the NOC and fault connectivity maps for the step change in T1in given in Figure 5. The top three symptom nodes for this fault were identified from the fault connectivity map as F4, T2 and T1, since these nodes were associated with the largest connectivity changes between the fault and NOC maps. These symptom nodes may imply a fault related to the

energy balances of the tanks, since these symptom nodes are the tank temperatures and steam flow to the second tank. By tracing the antecedent nodes, the variables associated with the root cause of the fault are identified as F1 and L1. These root causes do not correlate well with the true fault, that of a step change in T1in. In fact, the symptom nodes correlated better with the true cause.

In many of the cases the LC method did not show any change in connectivity in the presence of the faults, suggesting that this system contains nonlinear relations between variables (as is the case). For this reason the LC method was unable to identify the location of the fault. For the TE method in most cases it identified multiple possible root causes, but none of these identified were connected to all of the symptom nodes, so it was unable to narrow down the true cause. However, for both the LC and TE methods, variables associated with stream 1 were identified (specifically F1 and L1), suggesting that the fault cause is located in stream 1. The PC method performed well, being able to narrow the fault down to two possible root causes. Although none of the connectivity methods identified the exact causal variable, the vicinity of the fault (e.g. stream 1) could be gauged.

8. CONCLUSIONS

For the two-tank case study considered, both PCA and KPCA worked well for fault detection, with KPCA displaying the best detection performance. The CUSUM chart showed the best robustness in terms of false alarm / missing alarm trade-off, at the expense of a larger detection delay.

The NOC connectivity maps derived using partial correlation showed generally good correspondence to causal flow due to mass and energy conservation as well as control loops. However, some spurious connections were present, and some expected connections were absent.

In terms of fault identification, the change-in-connectivity approach could identify sensible symptom nodes (especially for the partial correlation connectivities), but the antecedent-root-cause approach could not identify sensible root cause nodes. The benefit of connectivity maps over contribution plots lies in the extracted causality flow: from symptom nodes to a causal root node.

In general, connectivity maps can be considered a useful tool to visualize process causality; however, any interpretations should be supported by the consideration of process knowledge beyond data-derived correlations.

Recommendations for future work include refining of the antecedent-root-cause approach by including contribution plot information, as well as estimating fault sizes based on calculations on the time-series of variables identified by the connectivity and contribution approaches.

REFERENCES

Bauer, M., Cox, J.W., Caveness, M.H., Downs, J.J., Thornhill, N.F., 2007. Finding the Direction of Disturbance Propagation in a Chemical Process Using Transfer Entropy. Control Syst. Technol. IEEE Trans. On 15, 12–21.

Bauer, M., Thornhill, N.F., 2008. A practical method for identifying the propagation path of plant-wide disturbances. J. Process Control 18, 707–719.

Bauer, M., Thornhill, N.F., Meaburn, A., 2004. Specifying the directionality of fault propagation paths using transfer entropy. Proc 7th Int. Symp. Dyn. Control Process Syst. 203–208.

Bin Shams, M.A., Budman, H.M., Duever, T.A., 2011. Fault detection, identification and diagnosis using CUSUM based PCA. Chem. Eng. Sci. 66, 4488–4498.

Chiang, L.H., Braatz, R.D., 2003. Process monitoring using causal map and multivariate statistics: fault detection and identification. Chemom. Intell. Lab. Syst. 65, 159–178.

Kano, M., Nagao, K., Hasebe, S., Hashimoto, I., Ohno, H., Strauss, R., Bakshi, B.R., 2002. Comparison of multivariate statistical process monitoring methods with applications to the Eastman challenge problem. Comput. Chem. Eng. 26, 161–174.

Kourti, T., MacGregor, J.F., 1995. Process analysis, monitoring and diagnosis, using multivariate projection methods. Chemom. Intell. Lab. Syst. 28, 3–21.

Lee, J.-M., Yoo, C., Choi, S.W., Vanrolleghem, P.A., Lee, I.-B., 2004. Nonlinear process monitoring using kernel principal component analysis. Chem. Eng. Sci. 59, 223–234.

McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12, 153–157.

Prabhu, S.S., Runger, G.C., 1997. Designing a multivariate EWMA control chart. J. Qual. Technol. 29, 8–15.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., Yin, K., 2003. A review of process fault detection and diagnosis: Part III: Process history based methods. Comput. Chem. Eng. 27, 327–346.

Yang, J., Li, L., Wang, A., 2011. A partial correlation-based Bayesian network structure learning algorithm under linear SEM. Knowl.-Based Syst. 24, 963–976.

Table 4. Fault Identification results for four different disturbances

	PCA		KPCA		Location node identified		
Fault	**Type**	**Size Increase**	**Type**	**Size Increase**	**LC**	**PC**	**TE**
Fault 1: Step T1in	5/5	yes	5/5	yes	none	5/5 F1,L1	none
Fault 2: Ramp T1in	4/5	yes	4/5	yes	3/5 F1	5/5 F1,L1	3/5 F1
Fault 3: Step T2in	4/5	yes	5/5	yes	none	5/5 F1,L1	none
Fault 4: Fouling	2/5	no	5/5	yes	none	5/5 F1,L1	1/5 F1

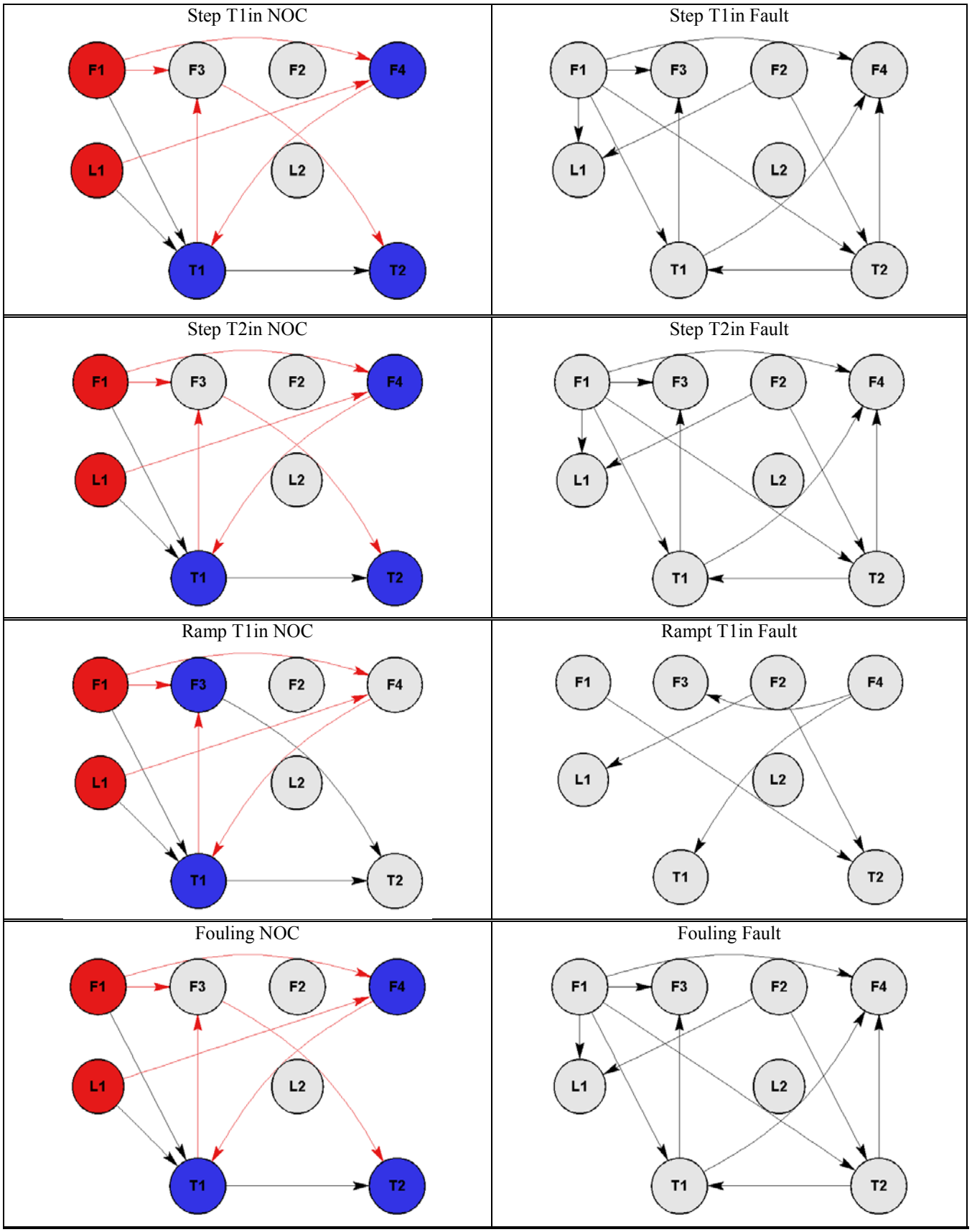

Fig. 5. Connectivity maps obtained using the PC method for NOC (left) and fault (right) conditions for each fault. Blue nodes indicate symptom nodes, red nodes indicate root nodes. Red lines indicate paths from root to symptom nodes