Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

Dynamic time warping based causality analysis for root-cause diagnosis of nonstationary fault processes

Gang Li [*], **Tao Yuan** [**], **S. Joe Qin** [***], **Tianyou Chai** [****]

[*] *Gang Li is with Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California 90089, USA. (Email: gli816@usc.edu)*
[**] *Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA. (Email: tyuan@usc.edu)*
[***] *Corresponding author. The Chinese University of Hong Kong, Shenzhen, 2001 Longxiang Blvd., Longgang, Shenzhen, Guangdong, China. On leave from The University of Southern California, Los Angeles, CA 90089, USA. (Email: sqin@usc.edu)*
[****] *State Key Laboratory of Synthetical Automation for Industrial Processes, Northeastern University, Shenyang, P. R. China. (Email: tychai@mail.neu.eud.cn)*

Abstract: It is very important to diagnose abnormal events in industrial processes. Based on normal operating data in a dynamic process, dynamic latent variable model provides a clear view of separating dynamic and static variations. Recent work by Li et al. (2014a) has shown an effective diagnosis in faulty variables with multidirectional reconstruction based contributions. Their further work took Granger causality analysis into accounts to explore the casual relations instead of only correlations. Although Granger causality is a widely used method for many applications, it needs time series to be stationary to calculate the causality index, which is not applicable for nonstationary fault processes. In this paper, a new causality analysis index based on dynamic time warping is proposed to determine the causal direction between pairs of faulty variables. The case study on the Tennessee Eastman process with a step fault shows the effectiveness of the proposed approach.

Keywords: Root cause diagnosis, causality analysis, dynamic latent variable model, multi-directional reconstruction based contribution, dynamic time warping, wavelet denoising

1. INTRODUCTION

In the area of industrial manufacturing, it has been a great issue to monitor industrial processes for a long time. Because of wide use of decentralized controller, it is difficult and expensive to obtain a plant-wide physical model. However, with the development of statistics and computer science, it has become easier to grasp the latent structure hidden under the huge data in both normal or abnormal situations. One of the most popular and fruitful area is multivariate statistical process monitoring, including principal component analysis (PCA) and partial least squares (PLS) (Qin, 2012). These data-driven modeling and monitoring techniques benefit greatly from the development in machine learning, distributed cloud computation, advanced sensors, large scale data collection and storage, which can be seen as the application of big data technology in traditional manufacturing monitoring.

Static PCA models are not enough to describe dynamic processes. In order to deal with auto-correlations among variables, dynamic PCA (DPCA) is proposed by Ku et al. (1995), which performs PCA on an augmented matrix of measurements. However, as DPCA uses variables at different time, it is difficult to interpret. Recently, Li et al. (2014b) proposed a new method of dynamic latent variable modeling for dynamic processes, which extracts dynamic factors explicitly and then static relations. To identify faulty variables, one of the most popular methods for root diagnosis is contribution plots (MacGregor et al., 1994) and reconstruction based contribution (RBC) (Alcala and Qin, 2009). In order to limit the number of candidates

⋆ This work was supported by the members of Texas-Wisconsin-California Control Consortium (TWCCC). It was also supported by NSFC under grants 61020106003, 61273173 and 61490700, and the SAPI fundamental Research of Northeastern University of China(2013ZCX02-01).

of faulty variables, Li et al. (2014a) extended RBC to multidirectional RBC.

However, as statistical models only describe correlations rather than causality, it is difficult to determine the causal direction between pairs of faulty variables. More recently, Granger causality has been used for identifying causality between different variables responsible for undesirable oscillations (Yuan and Qin, 2014). However, the above method is suitable for only stationary faulty process. For nonstationary faults such as step and drift faults, Granger analysis may lose its power.

In this work, a dynamic latent variable (DLV) model is used as the model for dynamic processes. Once a fault is detected, multi-directional RBC method is implemented to search a compact set of faulty variables. Then, wavelet based denoising is used to extract the trend of each time series from the candidates and k-means clustering is used for grouping these trends. Within a cluster, a dynamic time warping (DTW) based causality analysis index is proposed to determine the causality direction pairwise, which manages to locate the root cause and uncover the fault propagation path. Finally, the case study on the Tennessee Eastman process with a step fault is used to demonstrate the effectiveness of the proposed framework.

2. MODELING AND MONITORING WITH DLV MODEL

2.1 Dynamic latent variable modeling

Let vector $\mathbf{x}(k) \in \mathbb{R}^m$ represent m sensor measurements sampled at time k. Due to the nature of dynamic processes, samples at different time are correlated. To capture auto-correlations and cross correlations separately, a dynamic latent variable model is constructed, which optimizes the following objective (Li et al., 2014b):

$$\begin{aligned} &\max_{\mathbf{w},\boldsymbol{\beta}} \ (\boldsymbol{\beta}\otimes\mathbf{w})^T\mathbf{Z}^T\mathbf{Z}(\boldsymbol{\beta}\otimes\mathbf{w}) \\ &s.t. \quad \|\mathbf{w}\|=1, \|\boldsymbol{\beta}\|=1 \end{aligned} \tag{1}$$

where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{x}(q) & \mathbf{x}(q+1) & \dots & \mathbf{x}(n+q-1) \\ \mathbf{x}(q-1) & \mathbf{x}(q) & \dots & \mathbf{x}(n+q-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}(1) & \mathbf{x}(2) & \dots & \mathbf{x}(n) \end{bmatrix}^T \tag{2}$$

and $\boldsymbol{\beta}\otimes\mathbf{w} = [\beta_0\mathbf{w}^T, \beta_1\mathbf{w}^T, \dots, \beta_{q-1}\mathbf{w}^T]^T$ is the Kronecker product. The DLV model can be expressed as follows,

$$\begin{cases} \mathbf{t}(k) = \sum_{j=1}^{p} \boldsymbol{\alpha}_j \mathbf{t}(k-j) + \mathbf{v}(k) \\ \mathbf{x}(k) = \mathbf{P}\mathbf{t}(k) + \mathbf{P}_s\mathbf{t}_s(k) + \mathbf{e}_r(k) \end{cases} \tag{3}$$

and from a new sample, each part of DLV model can be calculated as:

$$\begin{aligned} \mathbf{t}(k) &= \mathbf{R}^T\mathbf{x}(k) \\ \mathbf{v}(k) &= \mathbf{t}(k) - \sum_{j=1}^{p} \boldsymbol{\alpha}_j \mathbf{t}(k-j) \\ \mathbf{t}_s(k) &= \mathbf{P}_s^T(\mathbf{I}-\mathbf{P}\mathbf{R}^T)\mathbf{x}(k) \\ \mathbf{e}_r(k) &= (\mathbf{I}-\mathbf{P}_s\mathbf{P}_s^T)(\mathbf{I}-\mathbf{P}\mathbf{R}^T)\mathbf{x}(k) \end{aligned} \tag{4}$$

where $\mathbf{t}(k)$ and $\mathbf{v}(k)$ represent the score and residual in dynamic variations, and $\mathbf{t}_s(k)$ and $\mathbf{e}_r(k)$ represent the score and residual in static variations. $\mathbf{P}, \mathbf{P}_s$ are the loading matrices corresponding to dynamic and static components. $\mathbf{R}$ is the weighting matrix, used for obtaining dynamic scores directly from original $\mathbf{x}$. $\boldsymbol{\alpha}_i$ is the parameters in the auto-regression model for $\mathbf{t}(k)$.

2.2 Monitoring based on the DLV model

DLV decomposes original measurement space into three subspaces with different meanings, which are supposed to be monitored separately. Therefore, three indices are constructed to monitor the process based on DLV model which are listed in the Table 1. Different indices have different

Table 1. Fault detection indices

Statistics	Calculation	Control limit
T_d^2	$\mathbf{v}^T\Lambda_v^{-1}\mathbf{v}$	$\delta_d = \frac{A(n^2-1)}{n(n-A)}F_{A,n-A,\alpha}$
T_s^2	$\mathbf{t}_s^T\Lambda_s^{-1}\mathbf{t}_s$	$\delta_s = \frac{A_s(n^2-1)}{n(n-A_s)}F_{A_s,n-A_s,\alpha}$
Q_r	$\|\mathbf{e}_r\|^2$	$\delta_r = g\chi^2_{h,\alpha}$

n: number of training samples, A: number of dynamic principal components; A_s: number of static principal components; $\Lambda_d = \frac{1}{n-1}\mathbf{V}^T\mathbf{V}$; $\Lambda_s = \frac{1}{n-1}\mathbf{T}_s^T\mathbf{T}_s$; $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_n]^T$; For Q_r, $g = \sigma^2/2\mu$, $h = 2\mu^2/\sigma^2$, μ is the sample mean of Q_r, and σ^2 is the sample variance of Q_r

meaning for monitoring. On one hand, T_d^2 monitors the innovation process for dynamic variations. On the other hand, T_s^2 reflects static variations, while Q_r measures the noise level and modeling error. In practice, it is preferred to combine T_s^2 and Q_r together into a static index to detect faults in the original space:

$$\phi_s(k) = \frac{T_s^2}{\delta_s} + \frac{Q_r}{\delta_r} = \mathbf{x}^T(k)\Phi\mathbf{x}(k) \tag{5}$$

where Φ can be derived easily.

2.3 Diagnosis based on multi-directional RBC

It is critical to identify faulty variables responsible for the detected fault. Alcala and Qin proposed reconstruction based contribution method to improve contribution a lot. For simple faults, the faulty sensor can be identified effectively by the RBC method. Nevertheless, for complex faults, RBC can be extended to a generalized form (Li et al., 2011):

$$\mathbf{RBC}_{\Xi} = \mathbf{x}^T\Phi\Xi(\Xi^T\Phi\Xi)^{+}\Xi^T\Phi\mathbf{x} \tag{6}$$

where $(\cdot)^+$ denotes the Moore-Penrose pseudoinverse of a matrix, $\boldsymbol{\Xi}$ is the fault direction matrix which can be extracted from faulty data. For a new fault, fault direction information is usually unavailable. Therefore, searching for a set of faulty variables responsible for the fault is an alternative. Li et al. (2014a) extended generalized RBC to multi-directional RBC, which searches the candidates among all sensors. A brief algorithm is given in Appendix A. After obtaining the candidate set S_f, it is necessary to calculate the contributions of selected candidates to the whole RBC:

$$\begin{cases} \mathbf{RBC} = \sum_{j=1}^{l} \mathbf{Cont}_j \\ \mathbf{Cont}_j = \{\xi_j^T[(\Xi^T\mathbf{M}\Xi)^+]^{\frac{1}{2}}\Xi^T\mathbf{M}\mathbf{x}\}^2 \end{cases} \quad (7)$$

where $(j = 1, ..., l)$ is the index number in S_f, and ξ_j is jth column of identity matrix $\mathbf{I}_l$. For convenience, contribution to $\mathbf{RBC}$ is normalized by $\mathbf{RBC}$. Subsequently, a criterion for candidate variable is defined over all samples:

$$\mathbf{TR}_j = \frac{1}{n_f}\sum_{k=1}^{n_f}\mathbf{Cont}_j(k)(j = 1, \ldots, m) \quad (8)$$

where $\mathbf{Cont}_j(k)$ means the $\mathbf{RBC}$ for the j-th variable of the k-th sample. As $\sum_{j=1}^{m}\mathbf{Cont}_j(k) = 1$ holds for normalized RBC contribution each sample, it is evident to see $\sum_{j=1}^{m}\mathbf{TR}_j = 1$. If $\mathbf{TR}_j > \frac{1}{m}$, i.e. the average contribution of a variable, x_i should be selected into the set of candidates. Notice that j in (7) denotes the j^{th} element in S_f, while j in (8) denotes the j^{th} variable.

3. CAUSALITY ANALYSIS FOR NONSTATIONARY FAULTS

The above multi-directional RBC can provide more compact candidate set than RBC. For static processes, candidates at different time may not change a lot. However, for dynamic processes, the candidates could change over time, for the reason that the influence of fault will propagate from one variable to another. Therefore, it is informative whether a variable has a causal impact on another.

Granger causality (GC) uses a statistical hypothesis test to judge whether a time series is causally affected by another time series based on predictability. Recently, Yuan and Qin (2014) proposed a method based on Granger causality for the diagnosis of oscillation sources and propagation. Li et al. then proposed an effective framework for root cause diagnosis of random faults in dynamic processes. However, it is not so effective to use Granger causality analysis for nonstationary fault process, because the model residual will not be stabilized so that the method of statistical inference fails. In this section, faulty data series are directly analyzed in the form of shape. For a pair of similar series in shape, an index is defined to measure the time order of them, which indirectly determine the causality between them. In order to remove disturbance from noise, only low-frequency part of a series is captured by using a wavelet based denoising procedure.

3.1 causality analysis based on DTW

In the area of time series, dynamic time warping is widely used as an algorithm for measuring the similarity between two temporal sequences which may vary in time or speed (Berndt and Clifford, 1994). Generally, DTW calculates an optimal match between two given time series with some restrictions. The series are warped in the time dimension to determine a measure of their similarity.

Given two time series $\mathbf{x}_i, \mathbf{x}_j$ in our case, the two sequences could be arranged on the sides of a grid, with $\mathbf{x}_i$ on the bottom and $\mathbf{x}_j$ the left hand side. Inside each cell, a distance measure can be calculated for the corresponding elements of the two time series. The objective is to find route with the minimum sum of distances between the start point and the ending point in all possible routes through the grid. The DTW algorithm adopts dynamic programming to keep track of the best path at each point as follows (Berndt and Clifford, 1994):

$$\gamma(u, v) = d(u, v)+\min\{\gamma(u-1, v-1), \gamma(u, v-1), \gamma(u-1, v)\} \quad (9)$$

where boundary conditions are given as $\gamma(1, 1) = d(1, 1)$, $\gamma(u, 1) = d(u, 1)+\gamma(u-1, 1)$, $\gamma(1, v) = d(1, v)+\gamma(1, v-1)$.

Remark 1. $d(u, v) = \|x_i(u) - x_j(v)\|$ is the distance measure for each pair of points in the grid, i.e. u point in x_i and v point in x_j. $\gamma(u, v)$ is the cost from point (1,1) to (u,v). From $\gamma(u, v)$ in all cells, a path with shortest cost can be found effectively by searching from (n,n) to (1,1) along only the bottom,left, bottom-left directions.

As DTW is sensitive to the basis and scale of the series, series should be preprocessed in a standard form. In this paper, time series is shifted firstly so that the minimum point reaches zero point. Secondly, the series data is zoomed by a factor so that the maximum point is set to one. At last, DTW value is divided by the larger length of two series. After the modification, DTW will be focused on the shape of curves, regardless of the basis, scale, and length of a time series.

Denote the DTW distance of two series $\mathbf{x}_i, \mathbf{x}_j$ with shift $(0 \le w \le w_{max})$ as

$$\mathbf{D}_{ij}(w) = \mathbf{DTW}(\mathbf{x}_i(1 : n - w), \mathbf{x}_j(w + 1 : n)) \quad (10)$$

Subsequently, a new DTW based causality index (DCI) can be defined as follows:

$$\mathbf{DCI}_{ij} = \frac{\min D_{ij}(w)}{D_{ij}(0)} \quad (11)$$

$\mathbf{DCI}_{ij}$ measures whether a pattern appears earlier in x_i than in x_j, i.e. whether x_i precedes x_j. Therefore, if $\mathbf{DCI}_{ij} < \alpha$, it can be concluded that there is a significant causality from $\mathbf{x}_i$ to $\mathbf{x}_j$, where α is the significance level. If $\mathbf{DCI}_{ij} \ge \alpha$, it indicates the two series are causal to each other. Notice that if two series are not similar in the shape, it is not proper to use this index for indicating the causality. Therefore, it is necessary to cluster these series according to DTW distance.

3.2 k-means clustering with DTW distance

The task of clustering is to classify a set of samples so that the within-group similarity is minimized and the between-group dissimilarity is maximized. k means clustering is popular for cluster analysis in data mining.

(k-means clustering with the DTW distance)

(1) Select the proper number of clusters, k. Initialize k centers for all clusters randomly, $\mathbf{c}_j, (j = 1, \ldots, k)$.
(2) For each time series $\mathbf{x}_i$, calculate $\mathbf{DTW}(\mathbf{x}_i, \mathbf{c}_j)$ and assign each sample to the class with the maximum of DTW distance,i.e. $class(\mathbf{x}_i) = \mathbf{argmax}_j \mathbf{DTW}(\mathbf{x}_i, \mathbf{c}_j)$.
(3) When all the objects have been assigned, recalculate the center of each cluster, i.e. $\mathbf{c}_j = \frac{1}{n_j} \sum_{\mathbf{x}_i \in S_j} \mathbf{x}_i$.
(4) Go to Step 2 until $\mathbf{c}_j$ converges for all j.

Given a set of series samples $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$, k-means clustering aims to partition m observations into k sets $\{S_1, S_2, \ldots, S_k\}$ (Kanungo et al., 2002). In this paper, the Euclidean distance is replaced by DTW distance, which is more effective for a time series. The algorithm is summarized as follows: After clustering the time series, it is clear which series are similar in shape and share similar dynamic response. It is reasonable to use **DCI** for causality analysis of series in a group. With the above clustering and causality analysis approaches, the root cause of each group is explored, which indicates the real source of the fault ultimately.

Remark 2. Once faulty variables are clustered into several groups, the relations among the variables within one group will be restricted into three types. Takes pairs (x_i, x_j) as two variables in a group. The possible cases are $x_i \to x_j$, $x_j \leftarrow x_i$, and $x_i \leftrightarrow x_j$. The case that there is no causality between x_i and x_j will not appear, because they are similar in the shape and therefore highly correlated.

However, it is also notable that the proposed approach only considers the causality that are reflected in a direct shape matching, covering linear and stretching transformation of two series. Those nonlinear and seriously distorted effects between two nodes with a strong causality, may be ignored considerably in this framework. The whole procedure is summarized in the Figure 1.

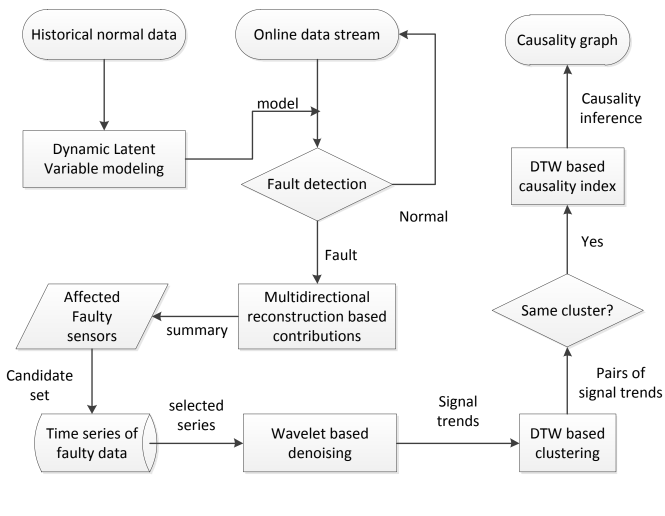

Fig. 1. The whole procedure for root diagnosis of dynamic processes

4. CASE STUDY ON TE PROCESS

In this section, the effectiveness of the proposed causality analysis methods is demonstrated by the Tennessee Eastman (TE) process, which was created as the benchmark for evaluating process control and monitoring approaches (Chiang et al., 2001). TE process benchmark provided 12 manipulated variables and 41 measurements for analysis, including 19 quality indices as product qualities, as shown in Table 2. In this paper, XMEAS(1-22) and XMV(1-11), are chosen as $\mathbf{X}$.

To perform process monitoring for the TE process, it is necessary to build a dynamic process model with normal historical data first. After 480 normal samples are centered to zero mean and scaled to unit variance, they are used to construct the DLV model with model parameters $q = 2$, $A = 6$, $p = 3$ and $A_s = 16$ (Li et al., 2014b). There are 15 known faults, including seven step faults, five random faults and three sticking and slow change faults. Due to page limitation, only the first type of fault denoted as IDV(1), is taken as an example to show the effectiveness of the method. When IDV (1) is introduced to the testing data at the 160th sample, a step change is induced in the A/C feed ratio in Stream 4, which causes an increase in the C feed and a decrease in the A feed in Stream 4. The description of IDV(1) implies variable 4 is directly affected and thus becomes the source for the fault. The disturbance in variable 4 will spread to the manipulated variable (26) quickly and then leads to influence in other variables. This fault is detected at 161th sample as shown in Fig. 2. Fig. 3 depicts the multi-directional RBC result for diagnosis. The area with darker color represents a higher contribution to RBC value, which should be more responsible for the fault. According to the criterion in 8, the candidates are selected as variables $\{2, 4, 18, 19, 20, 21, 23, 24, 26, 27, 31\}$. Figures 4-5 show all series from candidates before and after denoising, respectively, reflecting denoising procedure is beneficial to the trend extraction of series. The denoising is implemented with Haar wavelet in 5 layers of decomposition, and a soft threshold. After that, k-means clustering with DTW distance is performed on these trend signals. Although the initial number of clusters are 3, the result of clustering is two groups. One group consists of variables $\{2, 4, 20, 21, 23, 24, 26, 27\}$, the other contains $\{18, 19, 31\}$. The clustering result is convincing by observing the pattern in Fig. 5. Finally, $\mathbf{DCI}_{ij}$ is calculated for each group with a significance $\alpha = 0.99$ shown in Figs 6-7, where a dashed line represents a bi-directional causality and an arrowed line represents a unidirectional causality from the tail to the head. From Fig. 6, it can be seen that variable 4 and 26 are the most possible causes over other variables in Group 1, which also discloses where this fault originates. Group 2 has less correlation with group 1, and they response to the fault nearly simultaneously according to Fig. 7. The diagnosis results are consistent with the description of IDV(1), which shows a great improvement compared to existing RBC based methods.

Table 2. Process variables and quality variable

Var index	Description	Type	Var index	Description	Type
1	A feed (stream 1)	Measured	18	stripper temperature	Measured
2	D feed (stream 2)	Measured	19	stripper steam flow	Measured
3	E feed (stream 3)	Measured	20	compressor work	Measured
4	total Feed (stream 4)	Measured	21	reactor cooling water outlet temp	Measured
5	recycle flow (stream 8)	Measured	22	condenser cooling water outlet temp	Measured
6	reactor feed rate (stream 6)	Measured	23	D feed flow (stream 2)	Manipulated
7	reactor pressure	Measured	24	E feed flow (stream 3)	Manipulated
8	reactor level	Measured	25	A feed flow (stream 1)	Manipulated
9	reactor temperature	Measured	26	total feed flow (stream 4)	Manipulated
10	purge rate (stream 9)	Measured	27	compressor recycle valve	Manipulated
11	separator temperature	Measured	28	purge valve (stream 9)	Manipulated
12	separator level	Measured	29	separator pot liquid flow (stream 10)	Manipulated
13	separator pressure	Measured	30	stripper liquid product flow	Manipulated
14	separator underflow (stream 10)	Measured	31	stripper steam valve	Manipulated
15	stripper level	Measured	32	reactor cooling water flow	Manipulated
16	stripper pressure	Measured	33	condenser cooling water flow	Manipulated
17	stripper underflow (stream 11)	Measured			

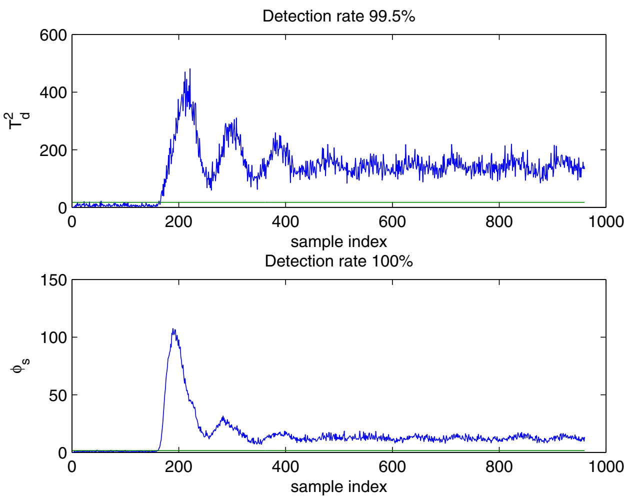

Fig. 2. Fault detection for IDV(1) with T_d^2 and ϕ based on DLV models

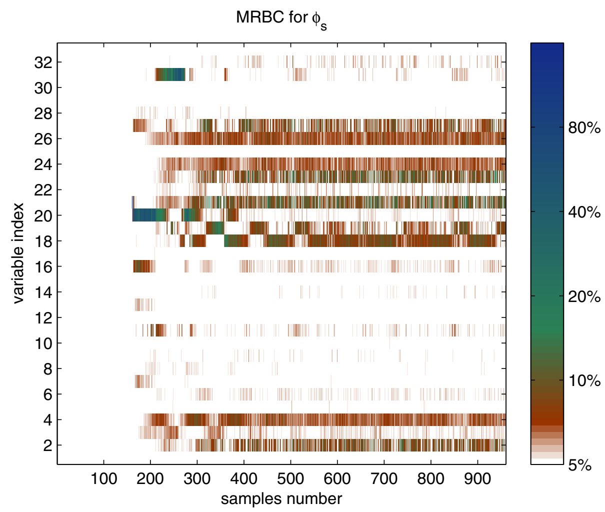

Fig. 3. Faulty variables identification for IDV(1) based on MRBC for ϕ_s

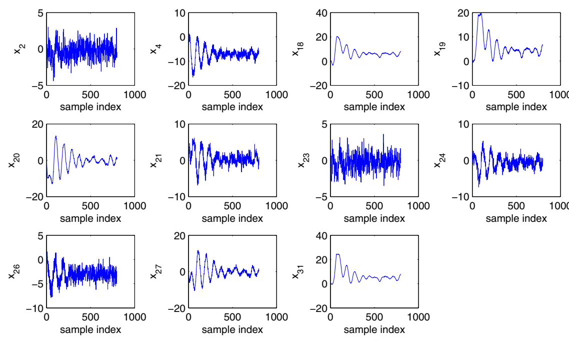

Fig. 4. Original measurements of selected faulty variables

5. CONCLUSION

In this paper, a new root diagnosis tool is proposed based on dynamic time warping with a dynamic latent variable model. When there is a fault detected by DLV model, Multi-directional RBC is used for initial diagnosis of fault related variables. After extracting the trend signal from the original faulty series, fault related variables can be clustered into groups with k-means clustering. Finally, the causality directions can be revealed with the proposed DTW based causality index. The case study on TE process shows a success in locating where the fault comes from. The method can be applied to other nonstationary faults, such as a drift fault.

Appendix A. MULTI-DIRECTIONAL RBC

For each faulty sample, the following algorithm can search for a minimized set of faulty related candidates:

(1) Initialize $S_f = \emptyset$ as the set of faulty variables and $l = 0$ is the number of elements in S_f .

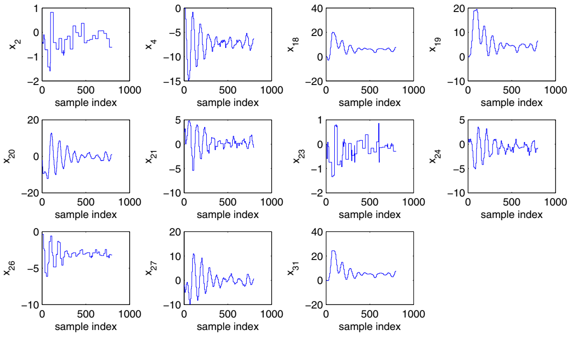

Fig. 5. Measurements of selected faulty variables after wavelet denoising

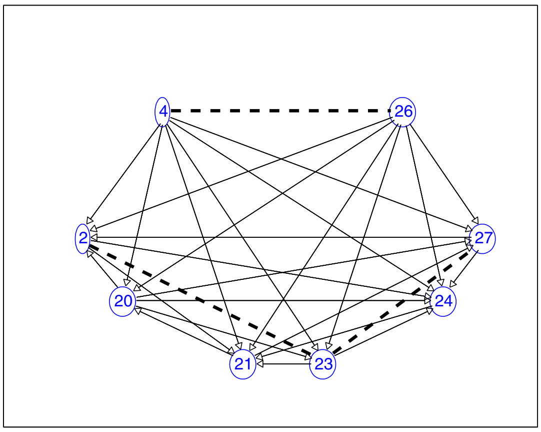

Fig. 6. Causality analysis for the 1st group of variables

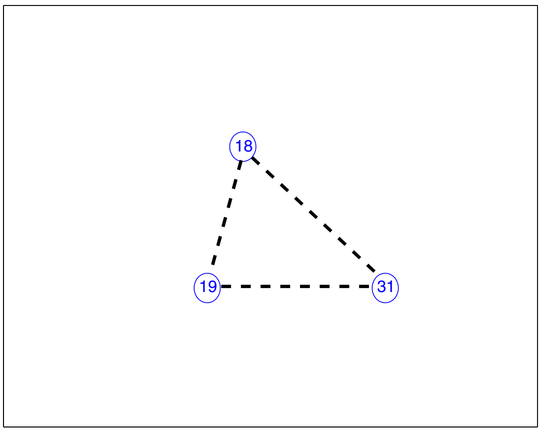

Fig. 7. Causality analysis for the 2nd group of variables

(2) for $i = 1 : m - l$, construct $\mathbf{\Xi}_i = \mathbf{I}(f_1, \ldots, f_l, g_i)$, where $f_l \in S_f$, $g_i \notin S_f$, $\mathbf{I}$ is the identity matrix, $\mathbf{I}(\cdot)$ consists of indexed columns of $\mathbf{I}$. Calculate **RBC** for $\mathbf{\Xi}_i$ with Eq. (6).

(3) Insert $f_{l+1} = \arg\max_{g_i} \mathbf{RBC}$ into S_f.

(4) $l = l + 1$, go back to Step 2, until reconstructed index $\mathbf{Index}_{recont} = \mathbf{Index} - \mathbf{RBC}_{\Xi}$ is below the control limit.

REFERENCES

Alcala, C. and Qin, S.J. (2009). Reconstruction-based Contribution for Process Monitoring. *Automatica*, 45(7), 1593–1600.

Berndt, D.J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, 359–370. Seattle, WA.

Chiang, L.H., Russell, E., and Braatz, R.D. (2001). *Fault detection and diagnosis in industrial systems*. Springer, London.

Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., and Wu, A.Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), 881–892.

Ku, W., Storer, R., and Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and intelligent laboratory systems*, 30(1), 179–196.

Li, G., Alcala, C.F., Qin, S.J., and Zhou, D. (2011). Generalized reconstruction-based contributions for output-relevant fault diagnosis with application to the tennessee eastman process. *Control Systems Technology, IEEE Transactions on*, 19(5), 1114–1127.

Li, G., Qin, S.J., and Chai, T. (2014a). Multi-directional reconstruction based contributions for root-cause diagnosis of dynamic processes. In *American Control Conference (ACC), 2014*, 3500–3505. IEEE.

Li, G., Qin, S.J., and Zhou, D. (2014b). A new method of dynamic latent variable modeling for processes monitoring. *IEEE Transactions on Industrial Electronics*, 61, 6438 – 6445.

MacGregor, J.F., Jaeckle, C., Kiparissides, C., and Koutoudi, M. (1994). Process monitoring and diagnosis by multiblock PLS methods. *AIChE Journal*, 40(5), 826–838.

Qin, S.J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36, 220–234.

Yuan, T. and Qin, S.J. (2014). Root cause diagnosis of plant-wide oscillations using granger causality. *Journal of Process Control*, 24(2), 450–459.

