Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

Root Cause Diagnosis of Process Faults Using Conditional Granger Causality Analysis and Maximum Spanning Tree

Han-Sheng Chen*, Zhengbing Yan**, Xuelei Zhang***, Yi Liu****, Yuan Yao*****

*Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
(e-mail: az122115@gmail.com)
**College of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China
(e-mail: zbyan@wzu.edu.cn)
***Shanghai Entry-Exit Inspection and Quarantine Bureau, Shanghai 200135, China
(e-mail: zhangxl@shciq.gov.cn)
****Institute of Process Equipment and Control Engineering, Zhejiang University of Technology, Hangzhou 310014, China
(e-mail: yliuzju@zjut.edu.cn)
*****Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
(Corresponding author. e-mail: yyao@mx.nthu.edu.tw)

Abstract:

In industrial processes, various types of faults often propagate from one unit to another along information and material flows. In severe cases, fault propagation can eventually affect the entire plant, leading to the reduction in product quality and productivity, and even causing damages. In order to avoid these issues, effective root cause diagnosis is desired because the correct identification of the sources of process abnormalities is critically important for restoring the system to its normal condition in a timely manner. In recent years, the data-driven causality analysis method, such as Granger causality (GC) test, has been adopted to identify the causes of process faults. However, the conventional pairwise GC only considers the causal relationship between a pair of time series. In multivariate cases, repeated pairwise analyses are often conducted, which yet often give over-complex and misleading results. To solve this problem, in this research, the multivariate GC technique, which measures the conditional dependence between time series, is utilized to construct the causal map between process variables. In addition, the obtained causal map is further simplified by finding its maximum spanning tree, facilitating the identification of the root cause. The feasibility of the proposed method is illustrated by case studies.

Keywords: root cause diagnosis, fault diagnosis, causality analysis, Granger causality, maximum spanning tree.

1. INTRODUCTION

Due to the rapid development of technology, the scale and complexity of many industrial plants is continuously increased. In order to ensure process safety and productivity, a large number of sensors are installed for data collection and process monitoring. In recent decades, data-driven multivariate statistical process monitoring (MSPM) techniques have been widely applied to both continuous and batch processes (Ge et al., 2013).

According to (Chiang et al., 2004), a complete procedure of process monitoring involves four main steps: fault detection, fault isolation, fault diagnosis, and process recovery. Significant research efforts have been devoted to fault detection and isolation, where the purpose of fault detection is to determine when abnormal process behavior has occurred, and fault isolation aims to identify the process variables critical to the detected fault. However, root cause diagnosis is still a difficult task because of fault propagation across different process units. A large-scale industrial process often consists of various interconnected process units, such as chemical reactors, heat exchangers, distillation towers, etc. Thus a fault can easily propagate from one unit to another along information and material flows. The utilization of the feedback control loops makes the propagation mechanism more difficult to analyze.

In recent years, causality analysis has received increasing attention in the research field of fault diagnosis, because a causal map can provide an intuitive way of representing the fault propagation pathways and revealing the root causes (Duan et al., 2014). Usually, the construction of a causal map requires the knowledge from a plant engineer (Chiang and Braatz, 2003; Chiang et al., 2015). In some applications, such information may be insufficient or too complicated to construct a causal map that is practical in root cause diagnosis. Therefore, data-driven causality analysis methods, such as Bayesian networks (Weidl et al., 2005), transfer entropy (Bauer et al., 2007), Granger causality (GC) analysis (Yuan and Qin, 2014), and dynamic time warping (DTW) (Li et al., 2016), have received increasing attention. It is noted that none of the methods outperforms the others in all the situations. Just as the famous saying goes: "all models are wrong but some are useful" (Box, 1979) in specific applications. This work mainly focuses on GC analysis which investigates the flow of information between time series using a statistical hypothesis test.

However, the conventional GC (Granger, 1969) only considers the causal relationship between a pair of time series, while industrial processes are multivariate in nature. For constructing the entire causal map, repeated pairwise analyses can be conducted. Nevertheless, pairwise GC is defined based on two principles (Granger, 1980): the cause happens prior to its effect, and the cause contains unique information about the future values of its effect over the period of analysis. In industrial processes, the second principle is often violated due to the correlation among variables. As a result, pairwise GC often gives over-complex and misleading results.

In this research, the multivariate conditional GC technique (Barnett and Seth, 2014; Geweke, 1984) is adopted to overcome the above-mentioned problem and generate a more compact causal map. In addition, the obtained causal map is further simplified by finding its maximum spanning tree, facilitating the identification of the root cause.

2. GRANGER CAUSALITY ANALYSIS

2.1 Pairwise Granger Causality

The concept of Granger causality was proposed in 1969 (Granger, 1969). As a tool for investigating the causal relationship between the two time series, pairwise GC test has been widely used in various fields, based on the following two assumptions.

I. The future can be affected by the past and present, not the other way around.

II. The cause set contains no redundant information. In other words, the cause variables cannot be correlated.

Consider two time series: $X_1 = (x_1(1), x_1(2), ..., x_1(n))$ and $X_2 = (x_2(1), x_2(2), \ldots, x_2(n))$. If X_1 Granger-causes X_2, the past and present values of X_1 can help to forecast the future values of X_2. To investigate such cause-effect relationship, a full auto-regressive (AR) model and a corresponding reduced model are built:

$$x_1(t) = \sum_{l=1}^{p} a_{11,l} x_1(t-l) + \sum_{l=1}^{p} a_{12,l} x_2(t-l) + \varepsilon_1(t) \,, \tag{1}$$

$$x_1(t) = \sum_{l=1}^{p} b_{11,l} x_1(t-l) + \varepsilon_{1(2)}(t) \,. \tag{2}$$

In these equations, $a_{1j,l}$ and $b_{11,l}$ are the model coefficients, ε_1 contains the residuals (i.e. prediction errors) of the full model, and $\varepsilon_{1(2)}$ contains the residuals of the reduced model. Obviously, the reduced model predicts the values of X_1 by excluding the influence of X_2 from the model. p is the model order defining the time lags included in the models, which can be specified by maximizing the Akaike Information Criterion (AIC) (Akaike, 1974) or the Bayesian Information Criterion (BIC) (Schwarz, 1978).

An improvement in the prediction is observed when the variability of ε_1 is significantly less than $\varepsilon_{1(2)}$, which can be qualified by conducting an F-test:

$$F_{X_2 \to X_1} = \ln \frac{\mathrm{var}(\varepsilon_{1(2)})}{\mathrm{var}(\varepsilon_1)} \,. \tag{3}$$

Test the null hypothesis H_0: $F_{X_2 \to X_1} = 0$ against the alternative hypothesis H_1: $F_{X_j \to X_i} > 0$. A rejection of the null implies that there is Granger causality from X_1 to X_2. The hypothesis can be tested via an F-statistic defined as

$$F_{\mathrm{statistic}} = \frac{(\mathrm{RSS}_0 - \mathrm{RSS}_1)/p}{\mathrm{RSS}_1/(N-2p-1)} \sim F(p, N-2p-1) \,, \tag{4}$$

where RSS_0 is the residual sum of squares (RSS) of the reduced model, RSS_1 is the RSS of the full model, and N is the total number of observations used to estimate the models. The null hypothesis is rejected if the F-statistic is greater than the confidence limit corresponding to a desired false-rejection probability.

In the cases of root cause diagnosis, X_1 and X_2 represent the time series of two different process variables. Because the industrial processes are inherently multivariate, repeated pairwise analyses are often required to construct the entire causal map. However, the correlation among the process variables breaks the second assumption of the pairwise GC analysis, affecting the interpretation of the results. The multivariate GC analysis technique is better suited to such situations.

2.2 Multivariate Conditional Granger Causality

Conditional GC (Geweke, 1984) is a multivariate version of GC, which includes simultaneously all measured variables into the AR models. In conditional GC analysis, the full model has a form of

$$x_1(t) = \sum_{l=1}^{p} a_{11,l} x_1(t-l) + \sum_{l=1}^{p} a_{12,l} x_2(t-l) + \sum_{j=3}^{J} \sum_{l=1}^{p} a_{1j,l} x_j(t-l) + \varepsilon_1(t), \tag{5}$$

while the reduced model is defined as

$$x_1(t) = \sum_{l=1}^{p} b_{11,l} x_1(t-l) + \sum_{j=3}^{J} \sum_{l=1}^{p} b_{1j,l} x_j(t-l) + \varepsilon_{1(2)}(t) \,. \tag{6}$$

In the above equations, J is the total number of the process variables under investigation, and x_j ($j = 3, \ldots, J$) include the time series of all variables except X_1 and X_2.

By fitting the models in (5) and (6) and conducting the hypothesis test with the F-statistic defined in (4), the direct GC relationship can be discovered, while the indirect GC conditional on other variables does not lead to the rejection of the null hypothesis. As a result, the causal map obtained by the conditional GC test is largely simplified and more meaningful comparing the results of the conventional pairwise GC, based on which it is easier to find the root cause of the process fault in diagnosis.

3. MAXIMUM SPANNING TREE

Although the utilization of the conditional GC technique is helpful in root cause diagnosis of process faults, the type I error, i.e. the incorrect rejection of a true null hypothesis, may still occur and make the results difficult to read. In addition, it is noted that the causal links identified by conditional GC do not necessarily correspond to the path of fault propagation. When the candidate set of the process variables is improperly selected for diagnosis, the resulting causal map may be misleading. Specifically, when redundant variables are included in the candidate set, the causal map identified by conditional GC often contains causal links irrelevant to the fault propagation and becomes unnecessarily complex, which is unfavorable for root cause diagnosis. Furthermore, the loops in the causal map make it more difficult to identify the root cause. In such situations, it is desired to further simplify the causal map and highlight the root cause variable. Here, the maximum spanning tree is introduced to solve the above-mentioned problems.

As known, a causal map is a directed graph that represents the cause-effect relations. For root cause diagnosis, each node in the causal map represents a candidate process variable, while the edge with an arrow between a pair of nodes indicates a causal relationship between two variables. In this research, a weight is assigned to each arrow to indicate the strength of causality, which is equal to the F-statistic used in the conditional GC test. In doing this, the causal map has a form of the weighted directed graph.

According to the graph theory, such a graph can be transformed into a spanning tree. The concept of spanning tree is originally developed for undirected graphs. By definition, a spanning tree T of an undirected graph G is a subgraph of G, which is a tree (i.e. a graph in which any two nodes are connected by exactly one path) including all of the notes of G with minimum possible number of edges. The idea of spanning tree can be generalized to directed graphs, e.g. (Gabow et al., 1986).

In this research, the causal map discovered by the conditional GC test is simplified by finding its maximum spanning tree, where a maximum spanning tree of a weighted graph is a spanning tree with a maximum total weight among all the spanning trees. In a maximum spanning tree, each node has one incoming edge except the root (i.e. the starting point) and at least one outgoing edge except the endpoints. For a causal map, maximizing the total weight in the spanning tree is equivalent to retaining the most significant causal relationships during the simplification procedure. Therefore, it is expected that all uncritical links can be removed from the causal map, while the note corresponding to the root cause variable can be highlighted because it appears as the root of the tree which has only an outgoing edge and no incoming edge.

The Chu-Liu/Edmonds' algorithm (Chu and Liu, 1965; Edmonds, 1967) is modified to find the maximum spanning tree. The detailed steps are as follows.

1. Choose an arbitrary node as the root.

2. Discard the incoming edges of the root.

3. For each node except the root, keep the incoming edge with the largest weight and discard all other incoming edges. Totally (J - 1) edges are retained.

4. If these edges do not form any cycle, a maximum spanning tree is found. Go to Step 9. Otherwise, go to Step 5.

5. "Contract" all the notes in each cycle into a single pseudo-node and generate a new graph accordingly.

6. In the new graph, the weight of the edge between two nodes u and v is defined as follows.

1) If u is a node outside the cycle and v is a node in the cycle, then a weight $w(u,v_c)$ is assigned to the incoming edge from u to the pseudo-node denoted as v_c.

$$w(u, v_c) = w(u, v) - w(\pi(v), v) , \tag{7}$$

where $\pi(v)$ denotes the note which is the source of the incoming edge to v in the original graph.

2) If u is a node in the cycle and v is a node outside the cycle, then a weight $w(v_c,v)$ is assigned to the outgoing edge from the pseudo-node v_c to u.

$$w(v_c, v) = w(u, v) . \tag{8}$$

3) If neither u or v is in the cycle, the weight of the edge between them is equal to $w(u,v)$, which is same as the corresponding weight in the original graph.

7. For each pseudo-node, select the incoming edge with the largest weight. Without loss of generality, suppose this edge connects a node (u_m) outside the cycle and a node (v_m) in the cycle. Then, replace the original incoming edge of v_m (selected in Step 3) by the edge from u_m to v_m. In doing this, the cycle is eliminated.

8. Go to Step 3.

9. Choose another node as the root and then go back to Step 2, until each node has been selected as the root once.

10. Compare the total weights of the generated maximum spanning trees and choose that with the maximum total weight as the final result.

4. CASE STUDIES

In this section, the Tennessee Eastman (TE) process (Downs and Vogel, 1993) is utilized to illustrate the feasibility of the proposed method. As a benchmark process, the TE process has been widely used for testing various monitoring and control algorithms. It consists of five main units, including a reactor, a condenser, a separation tower, a stripper and a compressor. There are eight components in the streams of the plant, namely four reactants A, C, D and E, two products G and H, a byproduct F and an inert component B. Totally 52 process variables, including 11 manipulated variables and 41 measured variables, are recorded in both normal operation and abnormal situations triggered by 20 different types of faults. The sampling interval is 3 min. In each scenario, the fault occurs at the 161st sampling time point. The flowchart of the TE process and the variable list can be found in the literature (Russell et al., 2000). In the following of this paper, the process variables are denoted as V_j, where j is the variable index.

For illustration, Fault 1 is considered, which is caused by a step change in the A/C feed ratio in Stream 4. In detail, the C feed is increased, while the A feed is decreased. The change of the A feed leads to a decrease of the amount of A in the recycle stream, i.e. Stream 5. Consequently, the composition of A in Stream 6 (V_{23}) is also decreased. In order to compensate the influence of such disturbance, the feedback control system increasing the A feed in Stream 1 (V_1), which eventually results in an increase of the A feed in Stream 6 (V_{44}). According to the process understanding, the root cause should be identified as V_{23}, although a large number of variables are affected by this fault.

For root cause diagnosis, a candidate variable set should be determined. Here, the principal component analysis (PCA) contribution plots (Westerhuis et al., 2000) are utilized for screening. Fig. 1 shows the average T^2 contribution of each process variable between the 161st and the 240th sampling time points. The variables with the first 20 largest contributions are selected for further analysis. Because of the smearing effect (Van den Kerkhof et al., 2013), these variables are all outside the corresponding control limits.

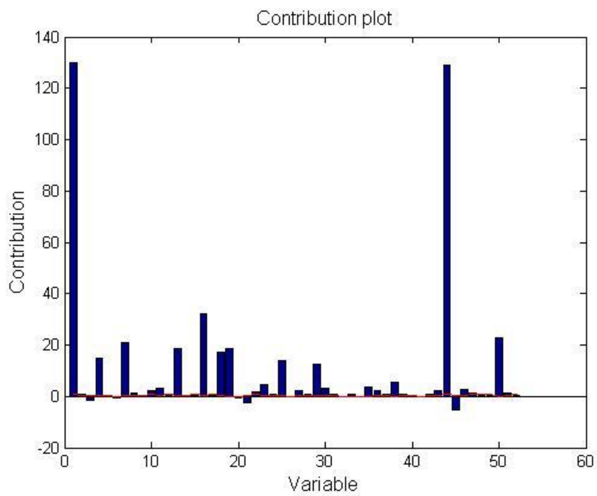

Fig. 1. Average T^2 contribution of each process variable (Fault 1)

Conditional GC test is conducted to discover the causal relationships among these 20 process variables. The result is shown in Fig. 2, which can be further transformed to a causal map (not shown here). In this figure, the x-axis denotes the indices of the cause variables, while the y-axis represents the indices of the effect variables. A solid block indicates that there is a Granger causality between the corresponding variable pair, which means the F-statistic is larger than the confidence limit corresponding to a false-rejection probability of 0.05. Therefore, a direct way to identify the root cause variable(s) is to look for the rows with no solid block. In Fig. 2, the variables correspond to the null rows are V_{30}, V_{36}, V_{43} and V_{44}. Obviously, such results are not correct, because the real root cause is V_{23}. Among the identified 4 variables, only V_{44} is directly affected by the root, i.e. V_{23}. The reason is multi-factorial. The inherent errors in statistical inference, the redundant candidate variables, the non-stationary and nonlinear characteristics contained in variable trajectories, etc. all may lead to misleading diagnosis results. Therefore, it is necessary to implement the maximum spanning tree.

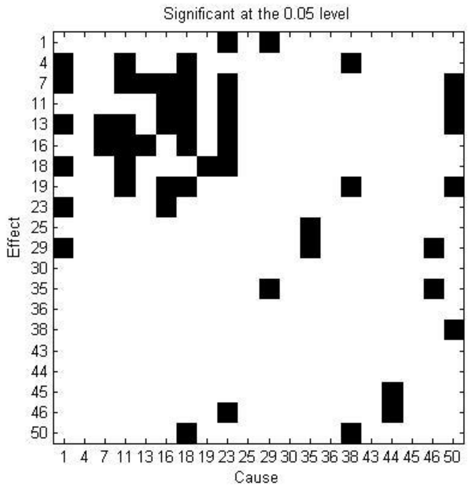

Fig. 2. Result of conditional GC test (Fault 1)

The result is shown in Fig. 3. It is clear that V_{23} is the root cause of Fault 1 because it has no incoming edge and appears as the root of the tree. Also, this variable affects V_1 and hence V_{44}. Such result is confirmed with the process understanding. Nevertheless, the maximum spanning tree does no retain the entire information of fault propagation, although it successfully highlights the root cause variable. The reason is that, in a tree, any two nodes are connected by exactly one path. However, in an industrial plant, recycle streams and feedback control loops widely exist. The real propagation path of a process fault is seldom describable with a tree. Therefore, we do not recommend tracing the fault propagation with the maximum spanning tree. This method is developed mainly for the easier identification of the root cause of a fault.

The same procedure is also used to diagnose Fault 7 which is related to the C header pressure loss. The feed flow of stream 4 sharply drops when the event occurs to the process. Consequently, the level of the reactor becomes lower than usual and many process variables are affected because of fault propagation. In order to maintain the reactor level, the feedback controller adjusts the flow rate of stream 4 (V_4) by manipulating the corresponding valve opening (V_{45}). Therefore, V_{45} and V_4 are the most likely root cause variables.

20 process variables are selected based on the contribution plot shown in Fig. 4, based on which conditional GC test is adopted for causality analysis. The result is plotted in Fig. 5, according to which V_{27} and V_{30} are mistakenly chosen to be the root cause. The implementation of the maximum spanning tree largely improves the result. As shown in Fig. 6, V_{45} is the root of the tree, while V_4 is the variable most close to the root. Such a result is reasonable. Again, it is emphasized that the path between the nodes on the tree is not necessarily in accordance with the path of fault propagation, especially for the nodes far from the root.

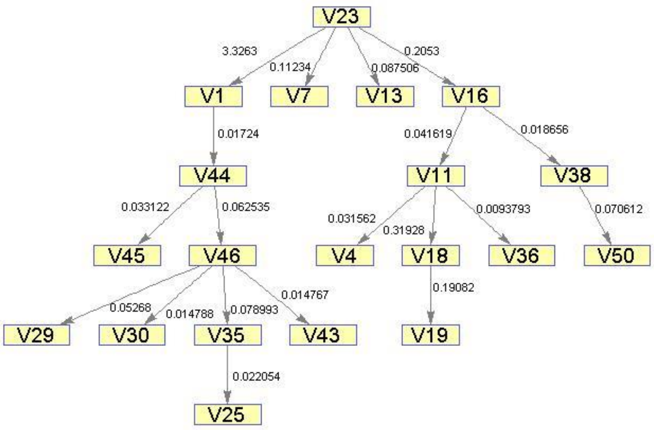

Fig. 3. Result of maximum spanning tree (Fault 1)

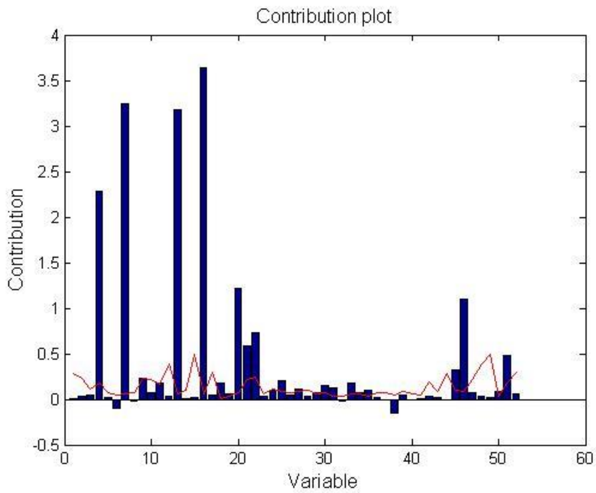

Fig. 4. Average T^2 contribution of each process variable (Fault 7)

6. CONCLUSIONS

In recent years, the causality analysis technique, GC, is utilized in root cause diagnosis of process faults. However, its performance may be affected by the correlation among process variables. In this research, multivariate conditional GC is adopted to handle this issue. In addition, the diagnosis performance is further improved by using the maximum spanning tree. The feasibility of the proposed method is illustrated by the case studies on the TE process.

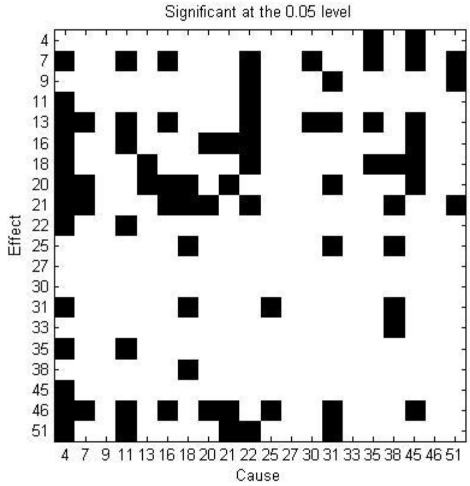

Fig. 5. Result of conditional GC test (Fault 7)

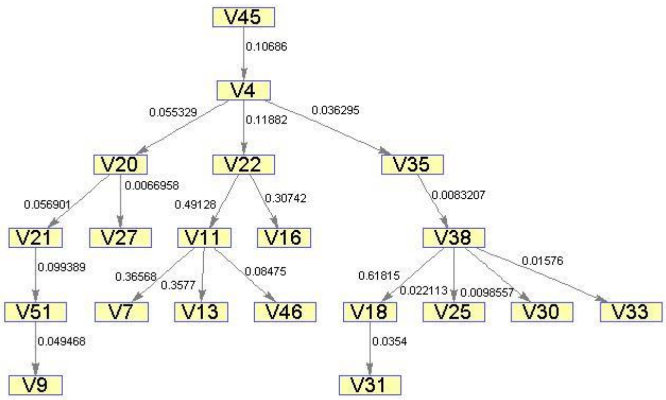

Fig. 6. Result of maximum spanning tree (Fault 7)

In the end of this paper, we would like to give some future perspectives on the related research topics. First, in a strict sense, GC test, as well as many other causality analysis tools, such as transfer entropy, is only suited to analyze the causality between stationary time series. In industrial processes, the trajectories of process variables often become non-stationary when abnormal events occur, violating the presupposition of GC test. This is an important issue that affects the performance of root cause diagnosis. The utilization of DTW may solve the problem in a certain sense (Li et al., 2016). However, DTW is only suited to the cases where the cause and effect time series have a similar shape. It is necessary to devote more research efforts on this topic. Second, the causality analysis results are sensitive to the selection of the candidate time series. In the context of process monitoring, fault isolation, i.e. the identification of faulty variables, is an important preparation step before conducting root cause diagnosis. It is recommended to use the isolation methods that can avoid the smearing effect (Kuang et al., 2015). Third, in practice, the diagnosis result is often affected by the selection of the time window in which the causality analysis is conducted. By using a method like GC, the task of causality analysis is similar to a model identification problem. During different time periods, the system may have different excitation signals which lead to different identification results, i.e. causal maps. Fourth, a related question is whether normal operation data should be used in causality analysis. As mentioned, the identified causal relationships do not necessary indicate the path of fault propagation. Therefore, it should be very careful in result interpretation. Usually, using fault data only may highlight the causality related to fault propagation, while incorporating normal operation data may provide a more complete view of the cause-effect relationships among process variables. Fifth, the parameter determination in the GC models is another issue to consider. For example, the conventional GC assigns a same value to the time lags of different variables. It is doubt whether it is a best way. In summary, it is of great need to bridge the gaps between the statistical theories of causality analysis and their industrial applications on root cause diagnosis.

ACKNOWLEDGMENT

This work was supported in part by Ministry of Science and Technology, ROC under Grant No. MOST 105-2628-E-007-013-MY2 and MOST 106-2622-8-007-017. Yi Liu was supported by the Foundation of Key Laboratory of Advanced Process Control for Light Industry (Jiangnan University), Ministry of Education, China (Grant No. APCLI1603)

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.

Barnett, L., and Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, 223, 50-68.

Bauer, M., Cox, J. W., Caveness, M. H., Downs, J. J., and Thornhill, N. F. (2007). Finding the direction of disturbance propagation in a chemical process using transfer entropy. *IEEE Transactions on Control Systems Technology*, 15, 12-21.

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer and G. N. Wilkinson (ed.), *Robustness in Statistics*, 201-236. Academic Press, New York.

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

Chiang, L. H., Jiang, B., Zhu, X., Huang, D., and Braatz, R. D. (2015). Diagnosis of multiple and unknown faults using the causal map and multivariate statistics. *Journal of Process Control*, 28, 27-39.

Chiang, L. H., Kotanchek, M. E., and Kordon, A. K. (2004). Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Computers & Chemical Engineering*, 28, 1389-1401.

Chu, Y.-J., and Liu, T.-H. (1965). On the shortest arborescence of a directed graph. *Science Sinica*, 14, 1396-1400.

Downs, J., and Vogel, E. (1993). A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 17, 245-255.

Duan, P., Chen, T., Shah, S. L., and Yang, F. (2014). Methods for root cause diagnosis of plant-wide oscillations. *AIChE Journal*, 60, 2019-2034.

Edmonds, J. (1967). Optimum branchings. *Journal of Research of the National Bureau of Standards - B. Mathematics and Mathematical Physics*, 71B, 233-240.

Gabow, H. N., Galil, Z., Spencer, T., and Tarjan, R. E. (1986). Efficient algorithms for finding minimum spanning trees in undirected and directed graphs. *Combinatorica*, 6, 109-122.

Ge, Z., Song, Z., and Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 52, 3543-3562.

Geweke, J. F. (1984). Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association*, 79, 907-915.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424-438.

Granger, C. W. J. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control*, 2, 329-352.

Kuang, T.-H., Yan, Z., and Yao, Y. (2015). Multivariate fault isolation via variable selection in discriminant analysis. *Journal of Process Control*, 35, 30-40.

Li, G., Qin, S. J., and Yuan, T. (2016). Data-driven root cause diagnosis of faults in process industries. *Chemometrics and Intelligent Laboratory Systems*, 159, 1-11.

Russell, E. L., Chiang, L. H., and Braatz, R. D. (2000). Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 51, 81-93.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464.

Van den Kerkhof, P., Vanlaer, J., Gins, G., and Van Impe, J. F. M. (2013). Analysis of smearing-out in contribution plot based fault isolation for statistical process control. *Chemical Engineering Science*, 104, 285-293.

Weidl, G., Madsen, A. L., and Israelson, S. (2005). Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. *Computers & Chemical Engineering*, 29, 1996-2009.

Westerhuis, J., Gurden, S., and Smilde, A. (2000). Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 51, 95-114.

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