Sparse Adjacency Forecasting and Its Application to Efficient Root Cause Diagnosis of Process Faults

Pengyu Song, Chunhui Zhao*

**State Key Laboratory of Industrial Control Technology, College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China*
(Corresponding author, email: pysong@zju.edu.cn, chhzhao@zju.edu.cn)*

Abstract: Efficient fault root cause diagnosis is essential to ensure the production safety of industrial processes. The existing root cause diagnosis models can be summarized as linear methods and nonlinear methods. Linear methods cannot handle nonlinear processes well, while nonlinear methods usually require pairwise calculations between variables, which are complex and difficult to apply in real time. To address the above issues, a method for root cause diagnosis of nonlinear processes, termed sparse adjacency forecasting (SAF), is proposed in this paper. SAF is a causal inference method based on the idea of Granger causality. While forecasting time series, it constructs an adjacency matrix to synthesize the process information and the interaction of different variables. By adding sparse constraints to the adjacency matrix, the predictive effects between variables are reflected, and the causality is captured. This method only needs to model once to obtain the causal relationship between all variables, which avoids multiple modeling and improves diagnosis efficiency. Besides, in order to solve the nonlinear problem, multiple nonlinear random feature nodes are introduced for time series prediction. Two cases are adopted to verify the causal inference and root cause diagnosis performance of the proposed method, including a numerical case and the Tennessee Eastman benchmark process.

Keywords: root cause diagnosis; nonlinear process; causal inference; sparse adjacency forecasting; nonlinear random feature

1. INTRODUCTION

Fault detection and diagnosis are of great significance to ensure the production safety of industrial processes. In recent years, with the rise of big data and intelligent manufacturing, research on data-driven fault detection and diagnosis has increasingly emerged (Li, Zhao, & Gao, 2017; Zhao, & Huang, 2018; Zhao, Sun, & Tian, 2019; Yu, & Zhao, 2020; Chai, & Zhao, 2020; Feng, & Zhao, 2020). However, the current research on fault diagnosis focuses more on fault isolation, that is, to find the process variables that are significantly affected by the fault, but the root cause diagnosis of the fault has not been discussed and studied in depth.

The root cause diagnosis of process faults can be regarded as a causal inference task. The process data collected within a period of time after the occurrence of the fault are used to infer the causal relationship between the variables, so that the fault propagation pathways can be reflected. Granger causality (Granger, 1969) is a typical bivariate causality analysis method. The basic idea is that given two time series **x** and **y**, if the introduction of previous information of **x** can significantly enhance the prediction accuracy of **y**, then **x** is said to be the Granger cause of **y**. To date, Granger causality has been applied to root cause diagnosis of industrial processes. Yuan et al. (2014) used Granger causality to discover the root causes of plant-wide oscillations. Chen et al. (2018) proposed the GPR-Granger method, which uses gaussian process regression to replace the AR model in conventional Granger causality for time series prediction to overcome nonlinear and non-stationary problems. Liu et al. (2020) combined Maximum Spanning Tree and Granger causality to simplify causal inference results and facilitate the root cause diagnosis.

In addition to Granger causality, Transfer entropy (TE) (Williams, & Rasmussen, 1996) is another representative method for root cause diagnosis (Duan et al., 2013; Lee et al., 2020). TE is a causal inference method based on information theory, which is a nonlinear root cause diagnosis method. However, it requires the stationarity of the process variables. Besides, the high computational complexity also brings challenges to its application in real-time fault diagnosis. Except for Granger causality and TE, dynamic Bayesian network (DBN) is another root cause diagnosis method (Gharahbagheri, Imtiaz, & Khan, 2017). Similar to TE, dynamic Bayesian network is also limited by computational complexity.

The methods described above can be divided into two categories, including linear methods and nonlinear methods. The causal inference method represented by Granger causality is simple and efficient, and has good real-time performance, but it cannot tackle with nonlinear situations. Although nonlinear methods represented by GPR-Granger and TE can capture causality from a nonlinear perspective, while they are usually complex and require pairwise calculations between process variables, which brings a huge amount of calculation and significantly reduces the efficiency of root cause diagnosis. How to maintain high computational efficiency while extracting nonlinear causality is an urgent problem for real-time root cause diagnosis.

To solve the above-mentioned problems, in this paper, a novel causal inference method, termed sparse adjacency forecasting (SAF) is proposed and applied for the root cause diagnosis of nonlinear processes. SAF is essentially a nonlinear time series prediction method. Drawing on the idea of Granger causality, SAF aims to select other variables that can significantly improve the prediction accuracy of each variable, so as to extract causality. In SAF, a variable adjacency matrix used to characterize causality is constructed. Each variable performs weighted information fusion according to this adjacency matrix, and then performs time series prediction. By adding sparsity constraints to the adjacency matrix, the adjacency relationship that has a significant contribution to the prediction is selected, so that the causal relationship is captured. Since SAF directly obtains the adjacency matrix that characterizes the causality, complicated pairwise calculations are avoided, thus improving the efficiency of root cause diagnosis. The main contributions of this paper are summarized as follows.

(1) A SAF model is proposed for causal inference of multivariate time series. The causality is characterized by the adjacency relationship matrix between variables, which avoids pairwise calculations.

(2) An efficient root cause diagnosis strategy is proposed based on the SAF model. By analyzing the adjacency relationship between process variables, the propagation path of the fault information can be obtained, thereby discovering the root cause of the fault.

The remainder of this paper is organized as follows. The Granger causality, Lasso Granger causality methods are revisited in Section 2. Section 3 presents the proposed SAF model in detail. In Section 4, experimental results are presented to illustrate the performance of the proposed method. Finally, the conclusion is summarized in Section 5.

2. PRELMINARIES

2.1 Granger causality

Granger causality is a typical causal inference method. Given two time series **x** and **y**, if **x** is the cause of **y**, then the past information of **x** can assist in the prediction of **y**, and vice versa. Mathematically, two different autoregressive (AR) models are established, one of which is a bivariate model:

$$y(t) = \sum_{p=1}^{h} a_{1,p} x(t-p) + \sum_{p=1}^{h} a_{2,p} y(t-p) + \varepsilon(t) \tag{1}$$

and the other is a reduced model:

$$y(t) = \sum_{p=1}^{h} b_p y(t-p) + \varepsilon_r(t) \tag{2}$$

where $a_{i,p}$ and b_p are the parameters of the two AR models, ε and ε_r represent the prediction errors of the bivariate model and the reduced model, respectively, and h is the time lag. If the prediction residual of the bivariate model (ε) is significantly less than that of the reduced model (ε_r), it means that the past value of x is valid for the prediction of **y**. The following F statistic is constructed:

$$F = \frac{(RSS_0 - RSS_1)/h}{RSS_1/(N-2h-1)} \sim F(h, N-2h-1) \tag{3}$$

where RSS_0 and RSS_1 are the sums of squared residuals of the reduced model and the bivariate model, and N is the sample size. The null hypothesis is that the introduction of **x** cannot improve the prediction accuracy of **y**. If the null hypothesis is rejected with a confidence level α, then **x** is determined to be the Granger cause of **y**.

2.2 Lasso Granger

Despite its simplicity, conventional Granger causality cannot well tackle the causal analysis of multivariate time series. For a multivariate time series with J variables, if simply perform Granger causality test for each pair of variables, then the total number of operations is as high as $J(J-1)$. Besides, this method does not take the interaction among multiple variables into account, which may result in false causality. The Lasso Granger method (Arnold, Liu, & Abe, 2007) overcomes the above disadvantages through Lasso regression. Given a multivariate time series $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_{2,} \dots, \mathbf{x}_J]$, Lasso Granger build a prediction model for each variable xi as follows:

$$x_i(t) = \sum_{j=1}^{J} \sum_{p=1}^{h} \omega_{ji,p} x_j(t-p) + \lambda \sum_{i=1}^{J} \sum_{p=1}^{h} \| \omega_{ji,p} \|_1 \tag{4}$$

where $\omega_{ji,p}$ denotes the regression coefficient of $x_j(t\text{-}p)$ on $x_i(t)$, and λ is the L1 penalty factor.

The Lasso regularization method has the effect of sparse variable selection (Tibshirani, 1996), which tends to give larger regression coefficients to key variables, while the regression coefficients of other variables are 0. Hence, if $\omega_{ji,p}$ is equal to zero for each p, it means that x_j has no significant effect on the prediction of x_i, that is, x_j is not a Granger cause of x_i. Conversely, if there exists a p, such that $\omega_{ji,p}$ is not equal to zero, then x_j "Granger cause" x_i. The Lasso Granger method establishes a regression model for each variable, rather than for each pair of variables, which significantly improve the computing efficiency.

3. METHODOLOGY

In this section, the proposed SAF model and its corresponding root cause diagnosis strategy are introduced in detail. SAF is a nonlinear method, and it can infer the causal relationship between all variables only by modeling once.

3.1 The proposed SAF model

The schematic of the SAF method is shown in Fig. 1. The input of the SAF model is a multivariate time series **X** with dimensions (N×T), where N is the number of variables and T is the number of sampling points.

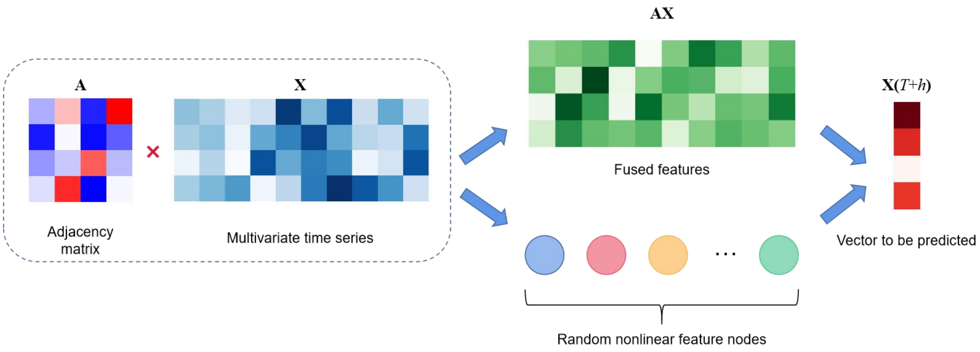


Fig. 1. The schematic of the proposed SAF method

SAF first uses an adjacency matrix $\mathbf{A}$ to linearly transform $\mathbf{X}$, which is calculated as:

$$\mathbf{X}_f = \mathbf{AX} \tag{5}$$

where $\mathbf{X}_f$ is the fused feature matrix, matrix $\mathbf{X}$ and $\mathbf{A}$ are respectively denoted as:

$$\mathbf{X} = \left[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\right]^{\mathrm{T}} = \begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(T) \\ x_2(1) & x_2(2) & \cdots & x_2(T) \\ \vdots & \vdots & \ddots & \vdots \\ x_N(1) & x_N(2) & \cdots & x_N(T) \end{bmatrix}_{N\times T} \tag{6}$$

$$\mathbf{A} = \left[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N\right]^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}_{N\times N} \tag{7}$$

Then, the i-th row vector of $\mathbf{X}_f$ can be calculated as:

$$\mathbf{x}_i^f = \left[\sum_{j=1}^{N} a_{ij} x_j(1), \sum_{j=1}^{N} a_{ij} x_j(2), \ldots, \sum_{j=1}^{N} a_{ij} x_j(T)\right]^{\mathrm{T}} \tag{8}$$

It can be found that this linear mapping is essentially a weighted summation of each variable in $\mathbf{X}$. The row vector of the i-th row in the $\mathbf{X}_f$ matrix represents the result of the weighted fusion of the i-th variable with other variables, which is defined as the feature vector of the i-th variable.

In the next step, SAF uses each row vector $\mathbf{x}_f^{(i)}$ in $\mathbf{X}_f$ to predict the future value $x_i(T+h)$ of the i-th variable $\mathbf{x}_i$. For a simple linear case, linear mapping is often used for prediction. Construct a linear weight matrix $\mathbf{W}$ with dimensions ($T\times N$), the prediction method is as follows：

$$\mathbf{X}_p = \mathbf{X}_f \mathbf{W} \tag{9}$$

Then the i-th element x_i^p on the diagonal of $\mathbf{X}_p$ can be calculated as:

$$x_i^p = \sum_{j=1}^{T} w_{ji} x_i^f(j) \tag{10}$$

It can be seen that x_i^p is only calculated by each element of the feature vector of the i-th variable, and does not contain information from the features of other variables. Take the vector of diagonal elements of $\mathbf{X}_p$ as the prediction result $\hat{\mathbf{x}}(T+h)$, where h is the horizon parameter that is set according to the time delay of the process.

For nonlinear processes, linear mapping is not applicable. In this paper, drawing on the idea of broad learning system (BLS) (Chen, & Liu, 2017), the nonlinearity is introduced by adding random mapping nodes. For the fused feature matrix $\mathbf{X}_f$, the i-th mapping node matrix is calculated as:

$$\mathbf{Z}_i = g(\mathbf{X}_f \mathbf{W}_i + \mathbf{b}_i) \tag{11}$$

where $\mathbf{W}_i$ and $\mathbf{b}_i$ are randomly generated mapping matrices and bias vectors, respectively, and g is a nonlinear mapping function, its expression is:

$$g(x) = 1/(1+e^{-x}) \tag{12}$$

These random mapping matrices and the original fusion feature matrix $\mathbf{X}_f$ are concatenated to obtain a new matrix with nonlinear features, that is:

$$\mathbf{X}_n = \left[\mathbf{X}_f \,|\, \mathbf{Z}_1 | \ldots | \mathbf{Z}_m\right] \tag{13}$$

where m is the number of mapping nodes. Then the future value of the time series is predicted according to equation (9), so that the nonlinear relationship is considered.

So far, it can be found that in both linear and nonlinear versions of SAF, the prediction of each variable needs to fuse information of other variables according to the adjacency matrix $\mathbf{A}$. According to equation (8), if the value of the element a_{ij} in the i-th row and j-th column of matrix $\mathbf{A}$ is relatively small, variable $\mathbf{x}_j$ will not have a significant effect to the prediction of $\mathbf{x}_i$. Similar to the Lasso Granger method, if the sparsity constraint can be added to the elements in the $\mathbf{A}$ matrix, the variables with significant predictive contributions can be selected to represent the causality. Hence, the optimization objective function of SAF is designed as follow:

$$\min_{\mathbf{A},\mathbf{W}} \left\|\mathbf{x}(t+h) - \hat{\mathbf{x}}(t+h)\right\|_2^2 + \alpha \sum_{i=1}^{N} \sum_{j=1}^{N} \left|a_{ij}\right| \tag{14}$$

where α is the sparse penalty coefficient.

In this way, if the variable x_j has a significant contribution to the prediction of x_i, the a_{ij} coefficient will be relatively large, otherwise, a_{ij} will be small or even zero. Therefore, a causal inference strategy can be designed based on the relative size of a_{ij}. For a trained SAF model, calculate the mean value of the absolute value of each row of the adjacency matrix **A**:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N}\left|a_{ij}\right| \tag{15}$$

where i represents the i-th row. For the variable $\mathbf{x}_i$, if the coefficient a_{ij} in the matrix **A** with the j-th variable $\mathbf{x}_j$ is greater than μ_i, it means that the variable $\mathbf{x}_j$ markedly contributes to the prediction of $\mathbf{x}_i$, that means, $\mathbf{x}_j$ is the cause of $\mathbf{x}_i$, and vice versa. According to this rule, the causal relationship between two variables can be divided into four cases, which are listed in Table 1.

Table 1. Causal inference strategy of SAF method

Case	$\left\|a_{ij}\right\| > \mu_i$	$\left\|a_{ji}\right\| > \mu_j$	Inference result
1	√	×	$\mathbf{x}_i$ is the cause of $\mathbf{x}_j$.
2	×	√	$\mathbf{x}_j$ is the cause of $\mathbf{x}_i$.
3	√	√	$\mathbf{x}_i$ and $\mathbf{x}_j$ are mutually causal.
4	×	×	There is no causality between $\mathbf{x}_i$ and $\mathbf{x}_j$.

3.2 SAF-based root cause diagnosis strategy

According to the prediction mechanism and causal inference strategy of the proposed SAF method, the causal relationship between process variables can be extracted. The main steps of SAF-based root cause diagnosis method are summarized as follows.

Pretep: Establish fault detection and fault variable isolation models for the process.

Step 1: Divide the collected fault process data with fixed time lag and horizon to generate training samples. The time lag can be selected according to the Akaike information criterion (AIC) (Akaike, 1974).

Step 2: Train the SAF model and obtain the **A** matrix. Before model training, the z-score method as shown below is usually used for standardization:

$$\tilde{\mathbf{x}} = \frac{\mathbf{x} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{16}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ represent the mean vector and the standard deviation vector of the training samples, respectively.

Step 3: Infer the causality between process variables. According to the trained **A** matrix, the mean value of the absolute value of each row of the adjacency matrix **A** is calculated, and the causal relationships are determined according to the inference strategy in Table 1.

Step 4: Construct the causal map and draw a conclusion regarding the root cause.

4. CASE STUDY

In this section, the root cause diagnosis performance of the proposed SAF method is illustrated with a numerical example and Tennessee Eastman process (TEP). The Lasso Granger (LG) method and dynamic Bayesian network (DBN) are used as representatives of linear and nonlinear methods to provide comparisons.

4.1 Numerical example

In this case, a nonlinear process with 7 variables with known causality is constructed as follows:

$$\begin{aligned}
&t_1(k) = \begin{cases} 0 & k = 1 \\ 0.9t_1(k-1) + 0.1\varepsilon_0(k) & otherwise \end{cases} \\
&x_1(k) = 2t_1(k-3) + 1 + \varepsilon_1(k) \\
&x_2(k) = -0.2t_1(k-2)^2 + x_1(k-2) + \varepsilon_2(k) \\
&x_3(k) = \exp(x_1(k-2)) + \varepsilon_3(k) \\
&x_4(k) = x_3(k-1)^2 + \varepsilon_4(k) \\
&x_5(k) = x_3(k-2) + x_4(k-2) + \varepsilon_5(k) \\
&x_6(k) = x_4(k-2)^3 + \varepsilon_6(k)
\end{aligned} \tag{17}$$

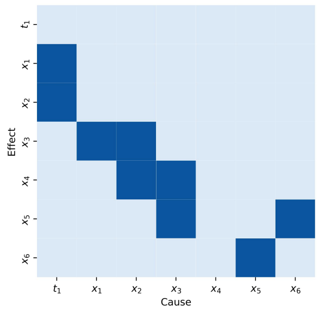
where $\varepsilon_0(k)$ is the independent random sample from the standard uniform distribution U (0,1) at the k-th sampling point, and ε_i (i=1, …, 6) are the independent Gaussian noises sampled from the standard normal distribution N (0,1). Obviously, t_1 variable is the root cause.

The causal inference results of the proposed SAF method with nonlinearity are presented in Fig. 2. It can be seen that SAF correctly diagnosed the root cause t_1, and the propagation path of the information is basically consistent with the real situation. The inference results of the two comparison methods, LG and DBN, are shown in Fig. 3 and Fig. 4, respectively. As shown in the figures, these two methods cannot correctly diagnose the root cause. The LG method is linear and cannot handle the nonlinear process in this case well. Although the DBN method can tackle with nonlinearity, it cannot overcome the non-stationary problem, which causes it to mistake x_6 as the root cause variable.

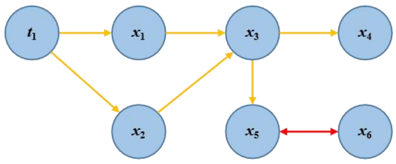
4.2 Tennessee Eastman process (TEP)

The TEP is a simulation system created based on an actual chemical process of Eastman Chemical Company. In this paper, IDV (1) fault case collected from the TEP is used to verify the performance of the proposed method. This fault leads to significant nonlinear and non-stationary characteristics of the process. The root cause variables of the fault are x_1 and x_{44}, and the specific propagation mechanism of the fault can be found in (Chen, Yan, Yao, Huang, & Wong, 2018). Refer to their previous work, the key fault variables include $\{x_1, x_4, x_7, x_{13}, x_{16}, x_{18}, x_{19}, x_{44}, x_{50}\}$.

The causal inference results of the proposed SAF method with nonlinearity, LG method, and BN method are presented in Fig. 5, Fig. 6 and Fig. 7, respectively. In the experimental results, there are some confusing circular causality structures in the causal map of the LG method, which makes it difficult to determine the root cause. Besides, the DBN method misjudges the root cause as the variable x_7. In contrast, due to its better adaptability to nonlinearity and non-stationarity, the SAF method correctly infers the root cause variables x_1 and x_{44}. What is more, the proposed SAF method more clearly reflects the fault propagation path, which further verifies its good performance in root cause diagnosis of nonlinear processes.



(a)



(b)

Fig. 2. The causal inference results of the proposed SAF for the numerical example. (a) causality matrix. (b) causality map.

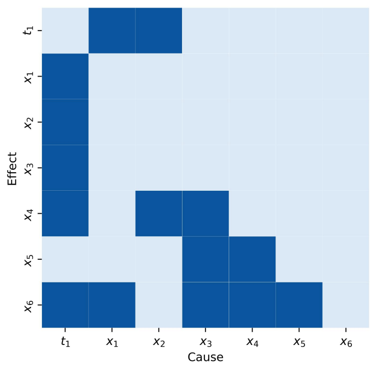


Fig. 3. The causal inference results of LG method for the numerical example.

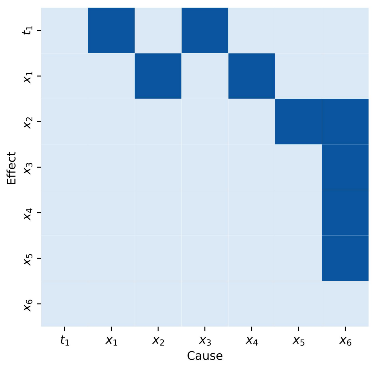
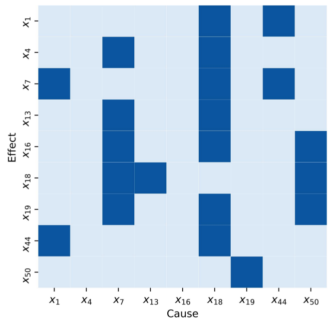
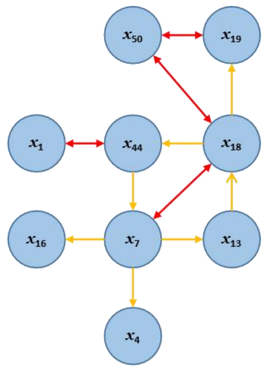


Fig. 4. The causal inference results of DBN method for the numerical example.

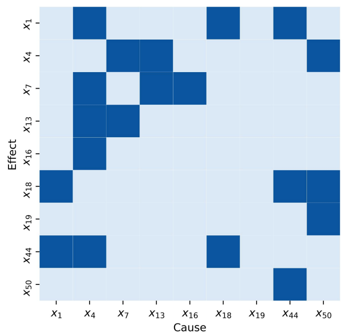


(a)

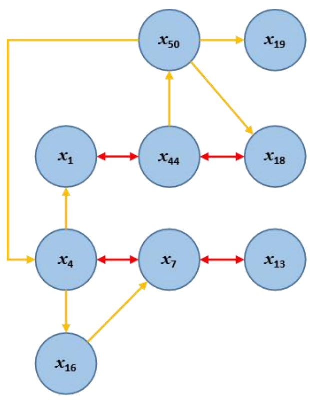


(b)

Fig. 5. The causal inference results of the proposed SAF for the TEP example. (a) causality matrix. (b) causality map.

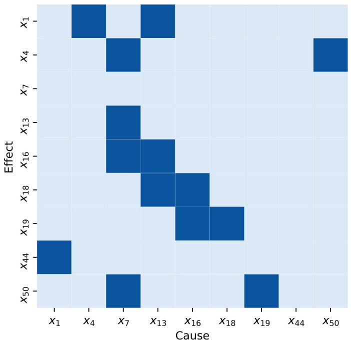


(a)

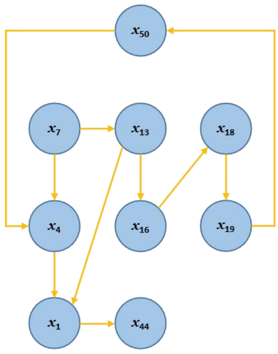


(b)

Fig. 6. The causal inference results of the LG method for the TEP example. (a) causality matrix. (b) causality map.



(a)



(b)

Fig. 7. The causal inference results of the DBN method for the TEP example. (a) causality matrix. (b) causality map.

5. CONCLUSIONS

In this paper, a nonlinear causal inference method, termed SAF is proposed and applied to the efficient root cause diagnosis of industrial process faults. The performance and efficiency of causal inference and root cause diagnosis are improved through a sparse constrained adjacency matrix mechanism and the introduction of nonlinear random features. A numerical example, and the Tennessee Eastman benchmark process are adopted to verify the performance of the proposed method. Experimental results indicate that the proposed SAF method can not only accurately extract the causal relationship between nonlinear process variables and reflect the root cause, but also improve the efficiency of nonlinear causal analysis and root cause diagnosis.

ACKNOWLEDGMENT

This work is supported by NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (No. U1709211), Zhejiang Key Research and Development Project (2019C01048).

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.

Arnold, A., Liu, Y., & Abe, N. (2007, August). Temporal causal modeling with graphical granger methods. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 66-75).

Chai, Z., & Zhao, C. (2020). A fine-grained adversarial network method for cross-domain industrial fault diagnosis. *IEEE Transactions on Automation Science and Engineering*, 17(3), 1432-1442.

Chen, C. P., & Liu, Z. (2017). Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE transactions on neural networks and learning systems*, 29(1), 10-24.

Chen, H. S., Yan, Z., Yao, Y., Huang, T. B., & Wong, Y. S. (2018). Systematic procedure for Granger-causality-based root cause diagnosis of chemical process faults. *Industrial & Engineering Chemistry Research*, 57(29), 9500-9512.

Duan, P., Yang, F., Chen, T., & Shah, S. L. (2013). Direct causality detection via the transfer entropy approach. *IEEE transactions on control systems technology*, 21(6), 2052-2066.

Feng, L., & Zhao, C. (2020). Fault description based attribute transfer for zero-sample industrial fault diagnosis. IEEE Transactions on Industrial Informatics, 17(3), 1852-1862.

Gharahbagheri, H., Imtiaz, S. A., & Khan, F. (2017). Root cause diagnosis of process fault using KPCA and Bayesian network. *Industrial & Engineering Chemistry Research*, 56(8), 2054-2070.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424-438.

Lee, H., Kim, C., Lim, S., & Lee, J. M. (2020). Data-driven fault diagnosis for chemical processes using transfer entropy and graphical lasso. *Computers & Chemical Engineering*, *142*, 107064.

Li, W., Zhao, C., & Gao, F. (2017). Linearity evaluation and variable subset partition based hierarchical process modeling and monitoring. *IEEE Transactions on Industrial Electronics*, 65(3), 2683-2692.

Liu, Y., Chen, H. S., Wu, H., Dai, Y., Yao, Y., & Yan, Z. (2020). Simplified Granger causality map for data-driven root cause diagnosis of process disturbances. *Journal of Process Control*, 95, 45-54.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2), 461.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

Yu, W., & Zhao, C. (2020) Robust Monitoring and Fault Isolation of Nonlinear Industrial Processes using Denoising Autoencoder and Elastic Net. *IEEE Transactions on Control Systems Technology*, 28(3), 1083-1091.

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

Zhao, C., Chen, J., & Jing, H. (2020). Condition-Driven Data Analytics and Monitoring for Wide-Range Nonstationary and Transient Continuous Processes. *IEEE Transactions on Automation Science and Engineering*.

Zhao, C., Sun, H., & Tian, F. (2019). Total variable decomposition based on sparse cointegration analysis for distributed monitoring of nonstationary industrial processes. *IEEE Transactions on Control Systems Technology*, 28(4), 1542-1549.