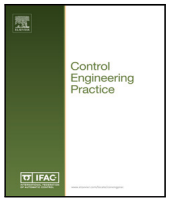
Control Engineering Practice 131 (2023) 105404

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac



A systematic nonstationary causality analysis framework for root cause diagnosis of faults in manufacturing processes

Hongjun Zhang [a,b], Kaixiang Peng [a], Liang Ma [a,*]

[a] Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, 100083, China
[b] Ansteel Group Corporation Limited, Anshan, 114000, Liaoning Province, China

ARTICLE INFO

Keywords:
Root cause diagnosis
Nonstationary causality analysis
Attention based gated recurrent unit
Gonzalo–Granger decomposition
Manufacturing processes

ABSTRACT

Due to production planning adjustments and equipment faults, most variables in the manufacturing processes exhibit nonstationary characteristics. The fault may be covered by the nonstationary trends of variables, making it difficult to identify fault propagation paths and locate root causes by means of traditional fault diagnosis methods. Thus, a systematic causality analysis framework is designed for root cause diagnosis of faults, which fully considers the nonstationarity of time series after faults and the propagation characteristics of faults in manufacturing processes. To be specific, the Kwiatkowski–Phillips–Schmidt–Shin and augmented Dickey–Fuller tests are firstly combined for assessing whether the industrial time series after faults are stationary or not. Then, for the nonstationary variables, the Gonzalo–Granger decomposition is used for extracting stationary factors. Furthermore, the stationary variables and stationary factors are merged into a new data set, an attention based gated recurrent unit is proposed for nonlinear and dynamic causality analysis aiming at root cause diagnosis of faults. Finally, a typical manufacturing process, the hot rolling process is used for verification, and the simulation and comparison results show the practicability and feasibility of the proposed scheme.

1. Introduction

Safe and reliable operation of manufacturing process is critical for the sustainable development and long-term profitability of modern process industries. There are thousands of process control variables and product quality indicators involved in the production process. An abnormal change in one measured variable will cause a chain reaction of other related variables, and meanwhile, the quality abnormality and performance degradation problems in the upstream processes will also be inherited to the downstream processes, resulting in the accumulation of quality and performance abnormal. Therefore, as an important step of fault diagnosis technology, root cause diagnosis, that can analyze causalities among fault variables, identify propagation path of faults, and locate root causes, plays a pivotal role in modern industrial processes. However, complex process characteristics, such as large-scale, nonlinear, nonstationary characteristics after faults, pose serious challenges to root cause diagnosis, which have become hot spots in industry and academia (Chen et al., 2021; Huang & Kadali, 2008; Jiang et al., 2021; Qin et al., 2020; Wang et al., 2022, 2020, 2016; Yu et al., 2022; Zhang, Zhang et al., 2022; Zhao et al., 2021; Zhao & Zhao, 2022).

Generally speaking, root cause diagnosis methods are divided into knowledge-based methods and data-based methods, as shown in Fig. 1. Knowledge-based methods are usually based on connectivity or correlation that use piping & instrumentation diagram (P&ID), expert experience knowledge, and algebraic equations (DAE), to construct a process topology for locating root causes and propagation path identification (Duan et al., 2012; Yang et al., 2014). Among them, the most representative methods are signed directed graph (SDG) (Maurya et al., 2004), adjacency matrix (AM) (Jiang et al., 2009), and fault tree (FT) (Kabir, 2017), which are often used in the engineering fields with simple operating mechanisms. However, those methods are difficult to apply in industrial processes lacking of prior knowledge. Meanwhile, the statistical information of real-time data are rarely involved, making it difficult to quantize the causalities among variables and adapt to the dynamic changes of the manufacturing processes.

With the rapid development of computer network, industrial communication, and advanced sensing technology, a large amount of data reflecting the operating status of manufacturing processes can be stored, making the data based root cause diagnosis methods become the mainstream direction. The widely used are cross correlation function (CCF) (Bauer & Thornhill, 2008), Granger causality (GC) (Bressler & Seth, 2010; Landman et al., 2014; Yuan & Qin, 2014), Bayesian network (BN) (Amin et al., 2019; Chen & Ge, 2020), and transfer entropy (TE) (Lindner et al., 2020; Naghoosi et al., 2013). Causal strengths can be frequently defined and the associated causalities can be captured

* Corresponding author.
E-mail address: liangma@ustb.edu.cn (L. Ma).

https://doi.org/10.1016/j.conengprac.2022.105404
Received 15 May 2022; Received in revised form 24 November 2022; Accepted 26 November 2022
Available online 9 December 2022

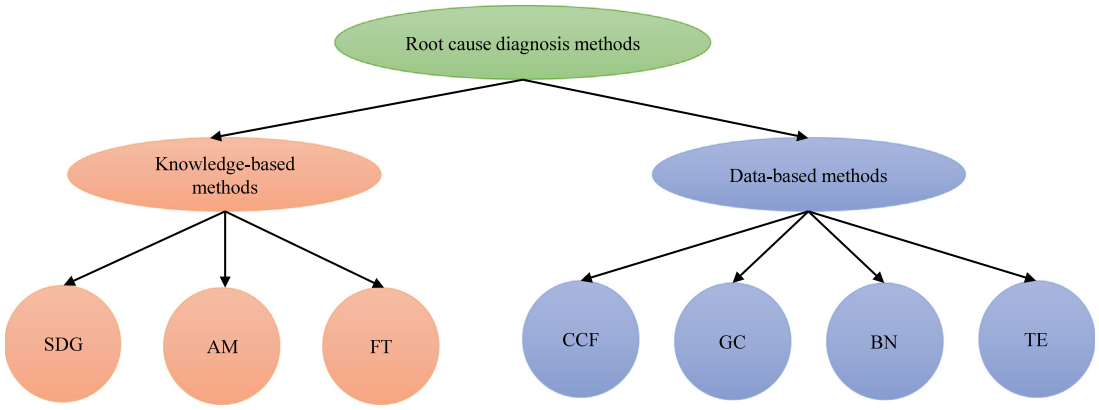


Fig. 1. Classical root cause diagnosis methods.

accordingly by those methods. For example, the CCF method can measure the correlation between two time series after time adjustment. The BN method can construct a directed acyclic graph (DAG) by means of learning conditional probability and inferring conditional correlation between variable nodes, which needs to assume that the diagnosed system is relatively stable within a certain time range. The GC analysis method examines the accuracy of one time series in forecasting another between two time series by multiple linear regression (MLR), which has been widely used because of its interpretability and easy implementation. However, most of the classical GC analysis methods are applicable for extracting the causalities of linear and stationary process measurements.

To obtain the nonlinear causalities, the information theory based asymmetric measurement method, TE, has received extensive attention. This method needs mass data to construct causal topology by calculating conditional probability function and designing reasonable directional measure. However, with the increase of data dimension, the computational complexity of joint probability density function (PDF) increases, which affects the construction efficiency of causal topology model. Meanwhile, once the production process is abnormal, the correct causalities among variables may be covered or distorted, making it less effective. Although some classical methods, such as Gaussian process regression (GPR) (Chen et al., 2018) and dynamic time warping (DTW) (Li et al., 2016) can overcome the limitation of nonstationary, lots of pairwise calculations are still required.

Alternatively, to handle the temporal relationship of high-dimensional data, deep learning methods show great advantages, which are widely used in various fields (Chen et al., 2022; Sun & Ge, 2021; Yuan et al., 2021; Zhao et al., 2019). Among them, the unique network structure design of recurrent neural network (RNN) makes it have certain advantages in time series modeling. Nevertheless, as the length of time series increases, long-term dependencies problem may arise, which may result in gradient explosion and vanishing in the process of model training. To address this issue, long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997; Zhang, Hu et al., 2022) and gated recurrent unit (GRU) (Cho et al., 2014) were further improved by introducing reasonable network structures into the basic RNN units. However, to the best of our knowledge, most of those methods focus on time series prediction, while the more challenging issues on propagation path identification and root cause location may not be deeply studied.

In the actual manufacturing processes, after fault occurs, some variables show nonstationary characteristics, while other variables present stationary characteristics, making the production process have mixed characteristics of stationary and nonstationary. How to fully explore the causalities of time series with coexistence of nonstationary and stationary before and after faults for propagation path identification and root cause location is a challenging problem. To this end, in this work, a systematic causality analysis framework is developed. The main contributions are shown as follows:

- proposing a new stationarity assessment method for industrial time series combined Kwiatkowski–Phillips–Schmidt–Shin (KPSS) and augmented Dickey–Fuller (ADF) tests;
- introducing the Gonzalo–Granger decomposition for extracting stationary factors from the nonstationary time series, and then an attention based GRU (AGRU) is developed for nonlinear and dynamic causality analysis aiming at propagation path identification and root cause location of faults;
- carrying out simulation and comparison experiments by a representative manufacturing process, the hot rolling process (HRP), indicating that the proposed method has outstanding root cause diagnosis accuracy compared with other competition methods.

The remainder of this paper is arranged as follows. Section 2 introduces the background and problem formulation. The proposed systematic nonstationary causality analysis framework is introduced in Section 3. Section 4 gives the simulations and discussions. Finally, the conclusion is given in Section 5.

2. Background and problem formulation

In this section, the basic knowledge of KPSS and ADF based stationary tests, and GRU are briefly described. Then, the problem formulation is given.

2.1. KPSS and ADF based stationary tests

Suppose that an observable time series x_t^*, $t = 1, 2, \ldots, T$, can be decomposed into

$$x_t^* = \xi_t + r_t + \varepsilon_t = \xi_t + r_{t-1} + u_t + \varepsilon_t \tag{1}$$

where ξ_t is a deterministic trend, r_t is a random walk, ε_t and u_t are stationary errors. Define

$$\sigma_\varepsilon^2 = \lim_{T\to\infty} \mathrm{E}[(\frac{\sum_{t=1}^{T} \varepsilon_t}{T})^2] \tag{2}$$

and

$$\sigma_u^2 = \lim_{T\to\infty} \mathrm{E}[(\frac{\sum_{t=1}^{T} u_t}{T})^2] \tag{3}$$

are long time variances, then, in the KPSS test, the stationarity hypothesis is simply $\sigma_u^2 = 0$ (Kwiatkowski et al., 1992).

In the ADF test, the observable time series x_t^* can be built by a regression model

$$x_t^* = \beta x_{t-1}^* + \sum_{k=1}^{l-1} \alpha_k \Delta x_{t-k}^* + \varepsilon_t \tag{4}$$

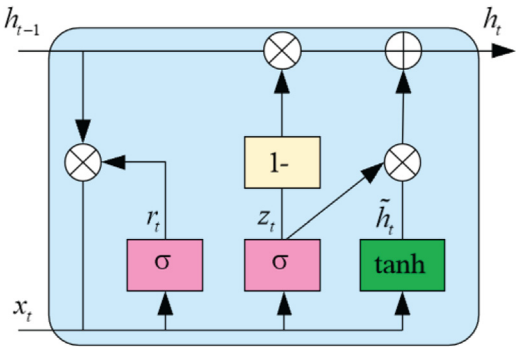


Fig. 2. Data flow and operations of GRU.

where $\Delta x^*_{t-k} = x^*_{t-k} - x^*_{t-k-1}$, l is the number of lags, α_k is the coefficient of Δx^*_{t-k}, β is an autoregression coefficient. In the ADF test, the stationarity hypothesis is simply $\beta = 0$ (Dickey & Fuller, 1981).

It can be seen that the KPSS test is a reversed test, the null hypothesis is that the observable time series is stationary, and the associated hypothesis is there exists unit root. Correspondingly, for the ADF test, the null hypothesis is the unit root presenting in the model, and the hypothesis is that the observable time series is stationary. In engineering practice, ideal stability test results may not be achieved by a single method. It is often recommended to combine both methods for reducing test error, which will provide basic stability test method in this work. For more details about the KPSS and ADF tests, please refer to the literatures (Dickey & Fuller, 1981; Kwiatkowski et al., 1992).

2.2. GRU-based nonlinear representation

GRU is an improved form of classical RNN, which contains two gates: reset gate r_t and update gate z_t, as presented in Fig. 2. At each time step t, r_t can be updated by

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{5}$$

where σ is an activation function, W_r and U_r are trained weights, x_t is the input at t, h_{t-1} is an activation at $t-1$, and b_r is the bias term. The candidate activation $\tilde{h}_t$ can be computed by

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{6}$$

where W_h and U_h are trained weights, $\odot$ is the Hadamard product between two vectors, and b_h is the bias term.

The update gate z_t can be computed as similar as r_t

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{7}$$

where W_z and U_z are the trained weights, and b_z is the bias term.

Then, the current activation h_t can be obtained by

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{8}$$

and the output is

$$o_t = \sigma(W_o h_t + b_o) \tag{9}$$

where W_o is trained weight, and b_o is the bias term (Cho et al., 2014).

It can be seen that long-term dependent and short-term important information can be properly saved and applied by the GRU method, making it more suitable for nonlinear and dynamic causality analysis. However, due to external disturbances and equipment faults, the actual manufacturing process has obvious nonstationary characteristics. If the traditional causality analysis method is still used, a large number of misdiagnosis may occur. Moreover, the faults are easily covered by the nonstationary trends of variables, making it difficult to identify fault propagation paths and locate root causes. As a result, extracting stationary features from nonstationary time series, and then using GRU method for realizing causality analysis, so as to diagnose root causes of faults is a challenging topic.

3. The proposed systematic nonstationary causality analysis framework

In this section, the proposed systematic nonstationary causality analysis method for root cause diagnosis of faults is explained.

3.1. The Gonzalo–Granger decomposition based stationary factor extraction

Suppose that the process variable set $\mathbf{X}^* \in \mathbb{R}^{T\times M}$ has been divided into a stationary variable set $\mathbf{X}^*_S = [x_{S1}, x_{S2}, \dots, x_{SM_S}]^{\mathrm{T}} \in \mathbb{R}^{T\times M_s}$ and a nonstationary variable set $\mathbf{X}^*_N = [x_{N1}, x_{N2}, \dots, x_{NM_N}]^{\mathrm{T}} \in \mathbb{R}^{T\times M_N}$ by combination of KPSS and ADF tests, where N is the number of samples, M_s and M_N are the number of stationary and nonstationary variables satisfying $M_s + M_N = M$, $\mathbf{x}_{Si} = (x_{S1}, x_{S2}, \dots, x_{ST})^{\mathrm{T}}$, $x_{Ni} = (x_{N1}, x_{N2}, \dots, x_{NT})^{\mathrm{T}}$, $i = 1, 2, \dots, T$. For the nonstationary time series x_{Nt}, a vector autoregressive (VAR) model can be built by

$$x_{Nt} = \sum_{i=1}^{p} \mathbf{C}_i x_{N(t-i)} + \mu + \delta_t \tag{10}$$

where $\mathbf{C}_i \in \mathbb{R}^{M_N\times M_N}$ is the coefficient matrix, μ is the intercept, δ_t is the white noise vector, p is the lag order.

The vector error correction (VEC) model can be obtained by adding an error correction mechanism

$$\Delta x_{Nt} = x_{Nt} - x_{N(t-1)} = \sum_{i=1}^{p-1} \mathbf{\Gamma}_i \Delta x_{N(t-i)} + \alpha\beta^{\mathrm{T}} x_{N(t-1)} + \delta_t \tag{11}$$

where $\mathbf{\Gamma}_i = -\sum_{j=i+1}^{p} \mathbf{C}_j$, $\alpha \in \mathbb{R}^{M_N\times r}$ and $\beta \in \mathbb{R}^{M_N\times r}$ are loading and cointegration matrixes, respectively (Johansen & Juselius, 1990).

Further, the nonstationary time series x_{Nt} can be divided into a combination of common trends and a set of stationary factors by means of Gonzalo–Granger decomposition

$$x_{Nt} = \beta_\perp(\alpha_\perp^{\mathrm{T}}\beta_\perp)^{-1}\alpha_\perp^{\mathrm{T}} x_{Nt} + \alpha(\beta^{\mathrm{T}}\alpha)^{-1}\beta^{\mathrm{T}} x_{Nt} \tag{12}$$

where $\alpha_\perp \in \mathbb{R}^{M_N\times(M_N-r)}$ are $\beta_\perp \in \mathbb{R}^{M_N\times(M_N-r)}$ are the orthogonal complements of α and β.

It can be seen that the dynamic equilibrium relationships among nonstationary variables can be reflected by the stationary factor $\alpha(\beta^{\mathrm{T}}\alpha)^{-1}\beta^{\mathrm{T}}\mathbf{x}_{Nt}$ (Escribano & Peña, 2010; Zhou et al., 2021). After that, the previous stationary variables and the stationary factors are combined for obtaining a new stationary variable set $\mathbf{X} = [x_1, x_2, \dots, x_M]^{\mathrm{T}} \in \mathbb{R}^{T\times M}$ with the purpose of the subsequent nonlinear and dynamic causality analysis. The flowchart of the Gonzalo–Granger decomposition based stationary factor extraction is presented in Fig. 3.

3.2. The AGRU based nonlinear and dynamic causality analysis

Suppose that $x_t \in \mathbf{X}$ is a stationary time series, then the classical full and reduced GC models can be built by a single VAR model

$$x_t = \sum_{l=1}^{q} \mathbf{A} x_{t-l} + \varepsilon_t \tag{13}$$

where $\mathbf{A} \in \mathbb{R}^{M\times M}$ denotes how the time lag q affects x_t, and ε_t is the noise with a mean of zero.

In above model, x_j does not Granger-cause x_i iff $\forall q$, $\mathbf{A}^q_{ij} = 0$. As a result, a GC analysis in that model reduces to determine which values in $\mathbf{A}$ are 0 over all lags. Nevertheless, it is hard to ensure partial values of $\mathbf{A}$ are 0 for all lags. Alternatively, whether $\mathbf{A}_{ij}$ is relatively small is a feasible way. Thus, an AGRU is designed for transforming the VAR model into a nonlinear form.

Assume that $\mathbf{A} = [a_1^{\mathrm{T}}, a_2^{\mathrm{T}}, \dots, a_M^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{M\times M}$, of which $a_i = [a_{i1}, a_{i2}, \dots, a_{iM}]$, and for obtaining smooth causalities, the attention score a_{ij} can be given by

$$a_{ij} = \frac{\exp(a_{ij})}{\sum_{j=1}^{M}\exp(a_{ij})} \tag{14}$$

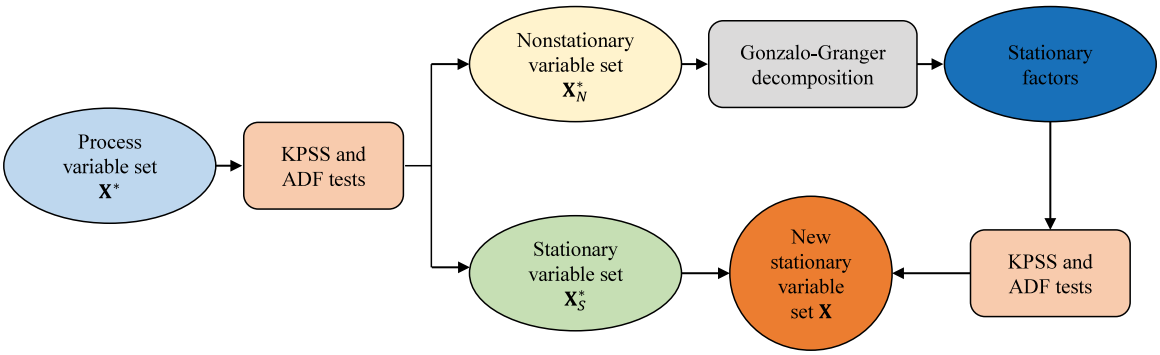


Fig. 3. Flowchart of the Gonzalo–Granger decomposition based stationary factor extraction.

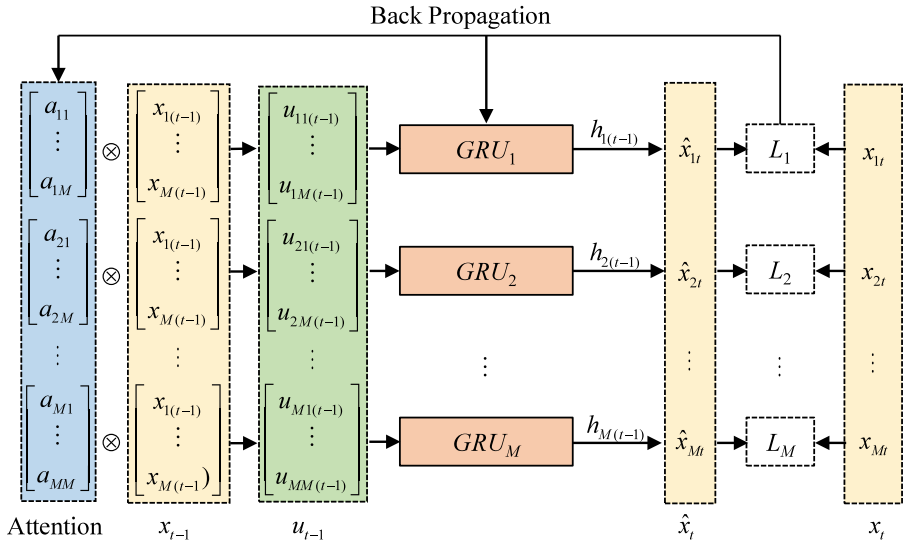


Fig. 4. The framework of AGRU for nonlinear and dynamic causality analysis.

If $a_{ij} > 0$, then x_i is the Granger causality of x_j. Thus, the nonlinear form is

$$x_{i(t+1)} = \mathrm{AGRU}(a_i \odot x_t) + \epsilon_{i(t+1)} \tag{15}$$

which can be generalized to multidimensional forms by means of M parallel and independent regression estimators (Bahdanau et al., 2014; He et al., 2020).

Fig. 4 is the developed AGRU framework of nonlinear and dynamic causality analysis, where $u_{it} = a_i x_t$ is an input hidden layer in view of causalities. It can be seen that there is no need for retraining AGRU for the restricted model. Thus, the limitation of nested loops can be avoided.

Then, the mean squared error (MSE) can be used for assessing the estimation loss

$$E = \frac{1}{2}\sum_{i=1}^{M}\sum_{t=1}^{T}[x_{i(t+1)} - \hat{x}_{i(t+1)}]^2 \tag{16}$$

where $\hat{x}_{i(t+1)}$ is the estimated value, and the following is the associated backpropagation learning algorithm.

The derivatives of the MSE with respect to the trained weights and bias term for output units are

$$\frac{\partial E}{\partial W_o} = \frac{\partial E}{\partial o_t}\frac{\partial o_t}{\partial W_o} = \frac{\partial E}{\partial o_t} \odot o_t \odot (1 - o_t) \cdot h_t^{\mathrm{T}} \tag{17}$$

and

$$\frac{\partial E}{\partial b_o} = \frac{\partial E}{\partial o_t}\frac{\partial o_t}{\partial b_o} = \frac{\partial E}{\partial o_t} \odot o_t \odot (1 - o_t). \tag{18}$$

The derivative of the MSE with respect to the activation is

$$\frac{\partial E}{\partial h_t} = \frac{\partial E}{\partial o_t}\frac{\partial o_t}{\partial h_t} + \frac{\partial E}{\partial h_{t+1}}\frac{\partial h_{t+1}}{\partial h_t} + \frac{\partial E}{\partial z_{t+1}}\frac{\partial z_{t+1}}{\partial h_t} + \frac{\partial E}{\partial r_{t+1}}\frac{\partial r_{t+1}}{\partial h_t} + \frac{\partial E}{\partial \tilde{h}_{t+1}}\frac{\partial \tilde{h}_{t+1}}{\partial h_t} \tag{19}$$

where

$$\frac{\partial E}{\partial o_t}\frac{\partial o_t}{\partial h_t} = \frac{\partial o_t^{\mathrm{T}}}{\partial h_t} \cdot \frac{\partial E}{\partial o_t}, \tag{20}$$

$$\frac{\partial E}{\partial h_{t+1}}\frac{\partial h_{t+1}}{\partial h_t} = \frac{\partial E}{\partial h_{t+1}} \odot (1 - z_{t+1}), \tag{21}$$

$$\begin{aligned}\frac{\partial E}{\partial z_{t+1}}\frac{\partial z_{t+1}}{\partial h_t} &= \frac{\partial E}{\partial h_{t+1}}\frac{\partial h_{t+1}}{\partial z_{t+1}}\frac{\partial z_{t+1}}{\partial h_t} \\ &= \frac{\partial E}{\partial h_{t+1}} \odot (\tilde{h}_{t+1} - h_t) \odot [U_z^{\mathrm{T}} \cdot (z_{t+1} \odot (1 - z_{t+1}))],\end{aligned} \tag{22}$$

$$\begin{aligned}\frac{\partial E}{\partial r_{t+1}}\frac{\partial r_{t+1}}{\partial h_t} &= \frac{\partial E}{\partial h_{t+1}}\frac{\partial h_{t+1}}{\partial \tilde{h}_{t+1}}\frac{\partial \tilde{h}_{t+1}}{\partial r_{t+1}}\frac{\partial r_{t+1}}{\partial h_t} \\ &= \frac{\partial E}{\partial h_{t+1}} \odot [U_h^{\mathrm{T}} \cdot (z_{t+1} \odot (1 - \tilde{h}_{t+1}^2))] \\ &\quad \odot h_t \odot [U_r^{\mathrm{T}} \cdot (r_{t+1} \odot (1 - r_{t+1}))],\end{aligned} \tag{23}$$

and

$$\frac{\partial E}{\partial \tilde{h}_{t+1}}\frac{\partial \tilde{h}_{t+1}}{\partial h_t} = \frac{\partial E}{\partial h_{t+1}}\frac{\partial h_{t+1}}{\partial \tilde{h}_{t+1}}\frac{\partial \tilde{h}_{t+1}}{\partial h_t} = \frac{\partial E}{\partial h_{t+1}} \odot z_{t+1} \odot [U_h^{\mathrm{T}} \cdot (1 - \tilde{h}_{t+1}^2)] \odot r_{t+1}. \tag{24}$$

The derivatives of the MSE with respect to the trained weights for the update gate are

$$\frac{\partial E}{\partial W_z} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial z_t}\frac{\partial z_t}{\partial W_z} = \frac{\partial E}{\partial h_t} \odot (\tilde{h}_t - h_{t-1}) \odot (z_t \odot (1 - z_t)) \cdot x_t^{\mathrm{T}} \tag{25}$$

and

$$\frac{\partial E}{\partial U_z} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial z_t}\frac{\partial z_t}{\partial U_z} = \frac{\partial E}{\partial h_t} \odot (\tilde{h}_t - h_{t-1}) \odot (z_t \odot (1 - z_t)) \cdot h_{t-1}^{\mathrm{T}}. \tag{26}$$

The derivatives of the MSE with respect to the reset gate and the associated trained weights are

$$\frac{\partial h_t}{\partial r_t} = \frac{\partial h_t}{\partial \tilde{h}_t}\frac{\partial \tilde{h}_t}{\partial r_t} = [U_h^{\mathrm{T}} \cdot (z_t \odot (1 - \tilde{h}_t^2))] \odot h_{t-1}, \tag{27}$$

$$\frac{\partial E}{\partial W_r} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial r_t}\frac{\partial r_t}{\partial W_r} = \frac{\partial E}{\partial h_t} \odot \frac{\partial h_t}{\partial r_t} \odot (r_t \odot (1 - r_t)) \cdot x_t^{\mathrm{T}}, \tag{28}$$

and

$$\frac{\partial E}{\partial U_r} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial r_t}\frac{\partial r_t}{\partial U_r} = \frac{\partial E}{\partial h_t} \odot \frac{\partial h_t}{\partial r_t} \odot (r_t \odot (1 - r_t)) \cdot h_{t-1}^{\mathrm{T}}. \tag{29}$$

The derivatives of the MSE with respect to the trained weights for the candidate activation are

$$\frac{\partial E}{\partial W_h} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial \tilde{h}_t}\frac{\partial \tilde{h}_t}{\partial W_h} = \frac{\partial E}{\partial h_t} \odot z_t \odot (1 - \tilde{h}_t^2) \cdot x_t^{\mathrm{T}} \tag{30}$$

and

$$\frac{\partial E}{\partial U_h} = \frac{\partial E}{\partial h_t}\frac{\partial h_t}{\partial \tilde{h}_t}\frac{\partial \tilde{h}_t}{\partial U_h} = \frac{\partial E}{\partial h_t} \odot z_t \odot (1 - \tilde{h}_t^2) \cdot (r_t \odot h_{t-1})^{\mathrm{T}}. \tag{31}$$

Then, based on above gradient calculations, the updated weights and bias terms can be got by

$$\begin{cases} \widehat{W}_n = W_n - \eta \dfrac{\partial E}{\partial W_n} \\ \widehat{U}_n = U_n - \eta \dfrac{\partial E}{\partial U_n} \\ \hat{b}_n = b_n - \eta \dfrac{\partial E}{\partial b_n} \end{cases} \tag{32}$$

where η is the learning rate.

Based on the chain rule, the derivatives of the MSE with respect to the attention score a_i can be obtained, and the associated parameters can be updated by means of training AGRU-GC at each epoch.

4. Simulations and applications to HRP

In this section, the feasibility of the systematic framework is verified by a representative manufacturing process, i.e. HRP. The classical linear GC (Bressler & Seth, 2010), LSTM-GC (Wang et al., 2018), GRU-GC methods are for competitive ones.

4.1. Process description

The HRP consists of numerous control loops and process variables, which are coupled with each other. Under the drive and action of energy flow, the material flow runs dynamically and orderly along a specific process network according to the preset program. Due to the changes in operating states, customized product varieties and specifications, frequent switching of operating conditions in the production process make the production process exhibit strong nonstationary characteristics. Once a fault occurs, it will propagate and evolve along different paths with the flow of material, energy and information, which may be covered by the nonstationary trends of variables. Thus, how to analyze the nonstationary causalities between time series before and after faults aiming at root cause diagnosis of faults is a meaningful topic.

The HRP is mainly composed of many production processes, such as furnaces, rolling mill process (RMP), finishing mill process (FMP), laminar cooling, coiler, and coiling, which is a continuous irreversible production process from raw materials to final products, as presented in Fig. 5. As the core part of HRP, the quality level of rolled products mainly depends on the technical equipment and control levels of FMP, which will be used for the background process in this paper. The FMP is composed of seven stands, and each stand is mainly composed of a pair of supporting rolls, a pair of working rolls and the corresponding hydraulic pressing devices. A total force detection device is generally installed at the lower part of the backup roll of rolling mill for measuring the rolling force, and the average gap is mainly controlled by a high-precision hydraulic servo control system for ensuring the thickness. Various instruments and sensors, such as thickness gauges and thermometers, are generally installed at the exit of the finishing mill for recording and storing a large amount of data.

Table 1
Description of the process variables in the FMP.

Variables	Descriptions
G_1 to G_7	Average gap of the eth stand, $e = 1, \ldots, 7$
F_1 to F_7	Total force of the eth stand, $e = 1, \ldots, 7$
B_2 to B_7	Work roll bending force of the eth stand, $e = 2, \ldots, 7$

According to the above analysis, normal and faulty data sets collected from debugging operation and returning product information were used. Two frequent fault cases were applied for validation of the proposed systematic nonstationary causality analysis method, and the relevant process variables are listed in Table 1 (the work roll bending force of F_1 is not measured). Three thousand samples were collected before and after the faults occurred for evaluating the performance.

4.2. Results and discussions

To evaluate the root cause diagnosis performance of the proposed systematic framework, two typical fault cases that frequently occurred in the HRP are taken into consideration.

(1) The first case

Fault 1 is induced by the failure of hydraulic roll-gap control system in the 4th stand. The occurrence of this fault makes the relevant stand unable to meet the reduction rate. As a result, other control units are affected by the abnormal behaviors, and the exit thickness or flatness is ultimately influenced.

Fig. 6 shows the plots of the process variable trajectories in FMP after this fault. It can be seen that some variables are nonstationary, while the other ones are stationary after the 2000th sample. Then, to further distinguish nonstationary and stationary variables, the combination of KPSS and ADF tests are carried out. Based on the hypothesis test results, if the KPSS statistic (Kwiatkowski et al., 1992) is larger than −3.000, and the ADF statistic (Dickey & Fuller, 1981) is larger than −4.000, then the variable trajectories are nonstationary. As can be seen from Table 2, the G_4, G_5, G_6, G_7, F_4, F_5, F_6, F_7, and B_7 are nonstationary. After that, stationary factors are obtained by means of Gonzalo–Granger decomposition, and the combination of KPSS and ADF tests are reapplied for testing the stationary of the stationary factors, where the critical values are −20.000 and −24.000, respectively. The associated test results are shown in Table 3.

After merging the stationary variables and stationary factors into the new stationary variable set, the proposed AGRU is used for nonlinear and dynamic causality analysis aiming at root cause diagnosis of this fault. In the simulations, to explore the effect of the number of hidden layers, the experiments of LSTM-GC, GRU-GC, and AGRU were carried out. The results are presented in Table 4, where the root MSE (RMSE) and mean absolute error (MAE) are used for evaluating the performance. It can be seen that the LSTM-GC, GRU-GC, and AGRU obtain the best performance with 2, 2, and 1 hidden layer, respectively.

Fig. 7 shows the causality analysis results of case 1, of which the ith row and the jth column denotes the directed causality from x_i to x_j, and the shade of color indicates the strength of causalities. It can be shown

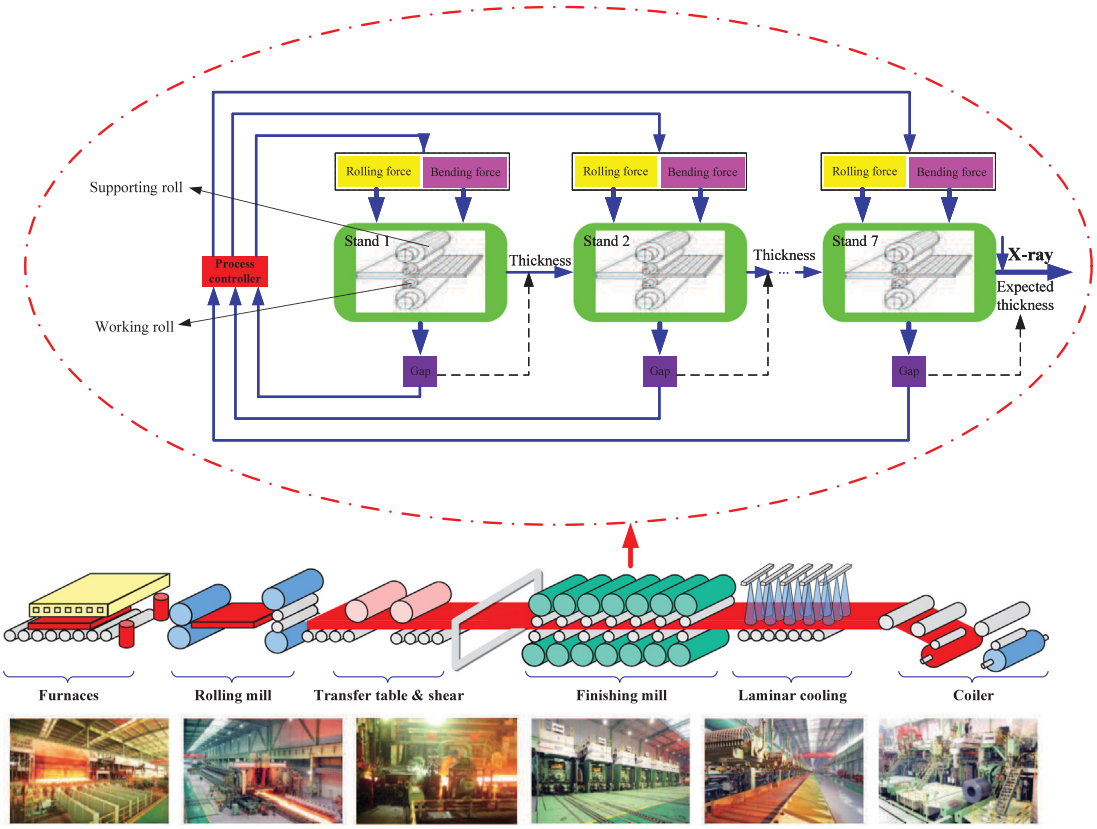


Fig. 5. Schematic layout of the HRP.

Table 2
Nonstationary test results for the first case.

Variables	KPSS statistic	ADF statistic	Test results	Variables	KPSS statistic	ADF statistic	Test results
G_1	−4.577	−17.038	Stationary	F_4	−2.124	−1.770	Nonstationary
G_2	−6.288	−11.066	Stationary	F_5	−2.077	−2.094	Nonstationary
G_3	−5.369	−8.314	Stationary	F_6	−2.990	−2.336	Nonstationary
G_4	−2.644	−3.224	Nonstationary	F_7	−1.998	−1.996	Nonstationary
G_5	−2.799	−2.766	Nonstationary	B_2	−6.003	−15.281	Stationary
G_6	−2.918	−1.990	Nonstationary	B_3	−4.906	−12.335	Stationary
G_7	−2.901	−2.033	Nonstationary	B_4	−3.962	−7.133	Stationary
F_1	−9.124	−7.652	Stationary	B_5	−3.186	−22.051	Stationary
F_2	−10.268	−6.399	Stationary	B_6	−5.299	−17.029	Stationary
F_3	−7.775	−13.260	Stationary	B_7	−2.118	−1.176	Nonstationary

Table 3
Nonstationary test results for stationary factors of the first case.

Variables	KPSS statistic	ADF statistic	Test results
G_4	−6.266	−7.436	Stationary
G_5	−5.299	−8.762	Stationary
G_6	−15.361	−8.305	Stationary
G_7	−12.133	−23.005	Stationary
F_4	−19.121	−7.620	Stationary
F_5	−11.445	−16.320	Stationary
F_6	−7.014	−10.269	Stationary
F_7	−8.471	−11.003	Stationary
B_7	−14.531	−12.129	Stationary

that the classical GC method cannot effectively capture the nonlinear causalities like $B_7 \longrightarrow F_7$ and $G_4 \longrightarrow G_5$, and the LSTM-GC method cannot capture the nonlinear causality like $G_6 \longrightarrow F_6$. However, the proposed AGRU-GC performs well for the nonlinear and dynamic time series. Furthermore, the propagation path identification results got by the expert experience and the proposed AGRU-GC method are shown in Figs. 8 and 9, of which the red solid line is the fault propagation path, while the blue dotted line is the bidirectional causality. It can be indicated that variable G_4 is diagnosed as the root cause. Subsequently, the work roll bending forces and average gaps of the posterior stands are affected by this abnormal behavior.

Causality analysis results of the proposed and competition methods are presented in Table 5, where E is the ratio of the quantity of correct causalities to the total of causalities, F is the ratio of the quantity of missing causalities to the total of causalities, and G is the ratio of the quantity of false causalities to the total of causalities. It can be shown that the proposed AGRU-GC method is superior to other three ones from the perspective of correct rate and missing causalities.

(2) The second case

To further verify the root cause diagnosis performance of the proposed method, the new approach and competitors are applied to the second fault case that is caused by the actuator failure of the cooling water control valve between the $2nd$ and the $3rd$ stands. This fault affects the bending force and rolling force of the subsequent stands. Consequently, the exit thickness or flatness of the final stand is affected.

Fig. 10 shows the plots of the process variable trajectories for the second case. It can be seen that many variables are nonstationary after the $1300th$ sample. Then, the combination of KPSS and ADF tests are used for estimating the number of nonstationary variables, where the

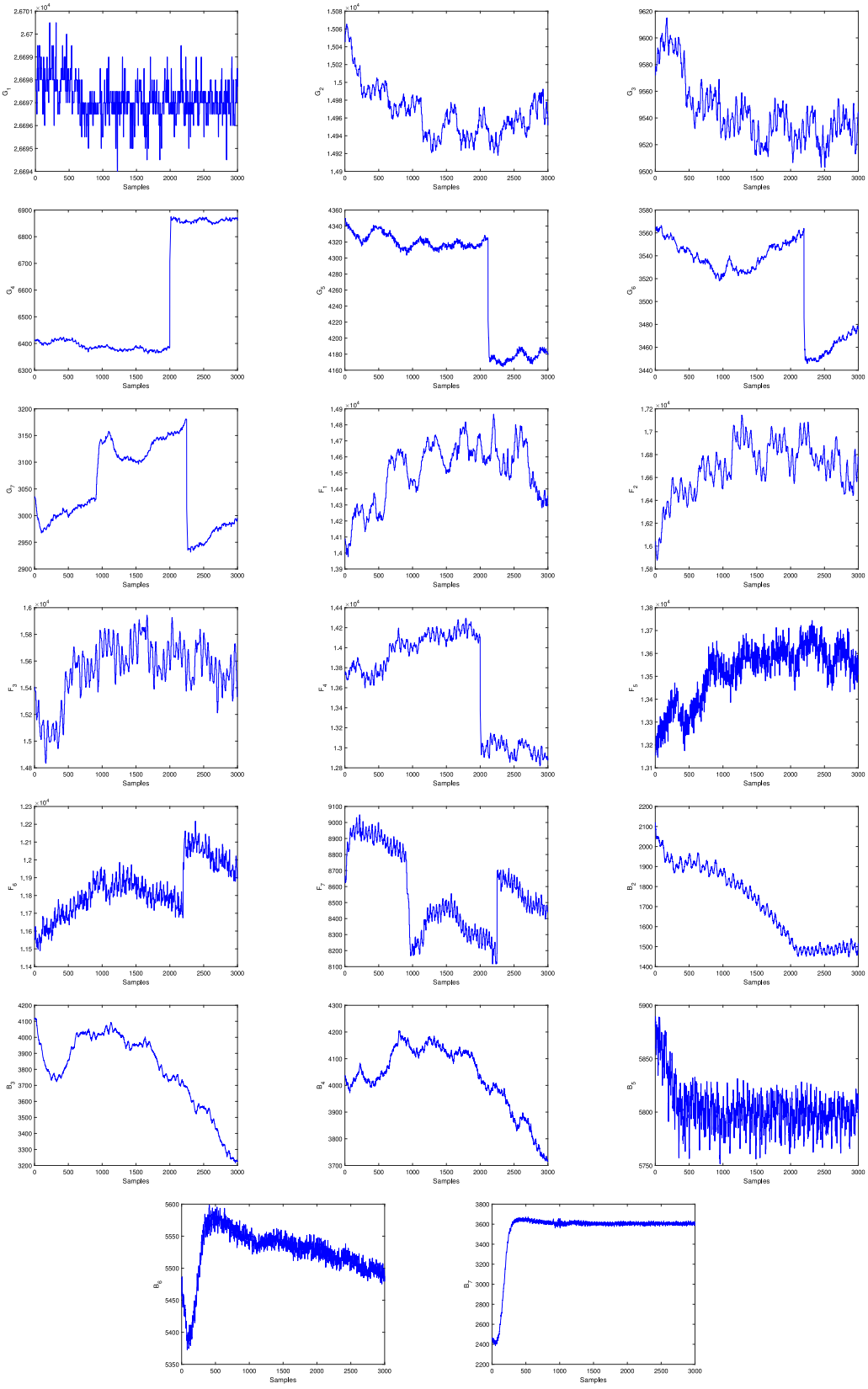


Fig. 6. Variable trajectories for the first case.

critical values are −4.000 and −7.000, respectively. From Table 6, it can be shown that G_3, G_4, G_5, G_6, F_3, F_4, F_5, F_6, F_7, and B_7 are nonstationary. Then, stationary factors are obtained by the Gonzalo–Granger decomposition, and the combination of KPSS and ADF tests are used for testing the stationary of the stationary factors, where the critical values are −5.000 and −6.000, respectively. The associated test results are shown in Table 7. Similar to the first case, the experiments of LSTM-GC, GRU-GC, and AGRU with different hidden layers were also

Table 4
Comparison of RMSE and MAE for different hidden layers in the first case.

Model	1		2		3		4		5		6	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
LSTM-GC	0.277	0.990	**0.057**	**0.090**	0.207	0.229	0.407	0.405	0.568	0.599	0.650	0.880
GRU-GC	0.415	0.718	**0.064**	**0.088**	0.225	0.346	0.487	0.422	0.499	0.475	0.763	0.790
AGRU-GC	**0.085**	**0.092**	0.177	0.195	0.218	0.496	0.407	0.438	0.533	0.516	0.660	0.677

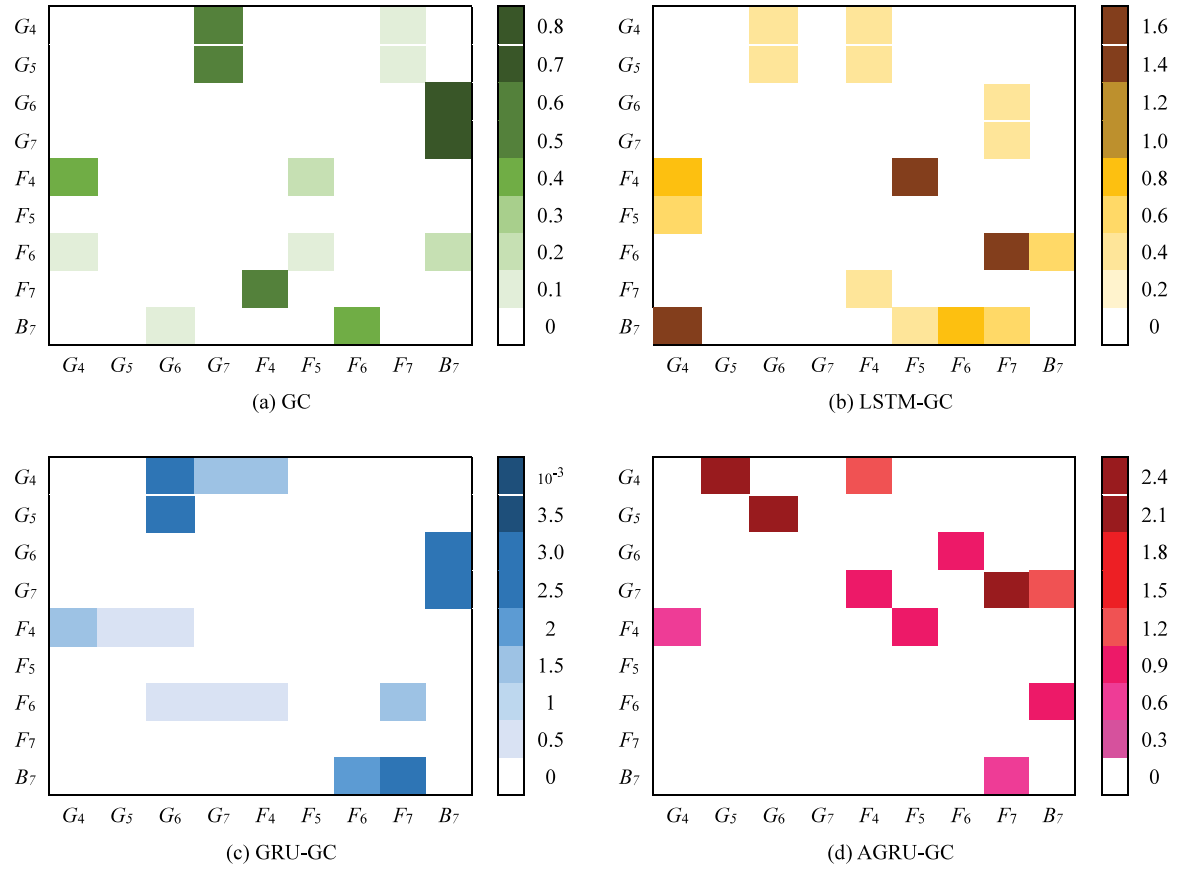


Fig. 7. Causality analysis results for the first case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

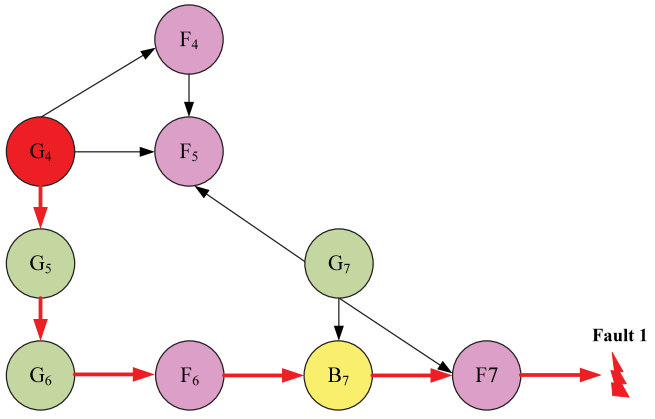


Fig. 8. Expert experience based propagation pathway for the first case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

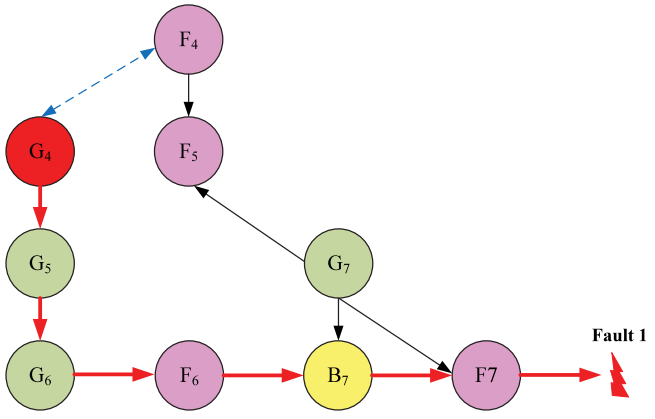


Fig. 9. AGRU-GC based propagation pathway for the first case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performed, and the results are presented in Table 8. It can be shown that the best performance can be obtained by 3, 2, and 1 hidden layer, respectively.

Fig. 11 presents the causality analysis results. It can be found that the nonlinear causalities can be well captured by the GRU-GC and AGRU-GC methods. Furthermore, as shown in Figs. 12 and 13, variable

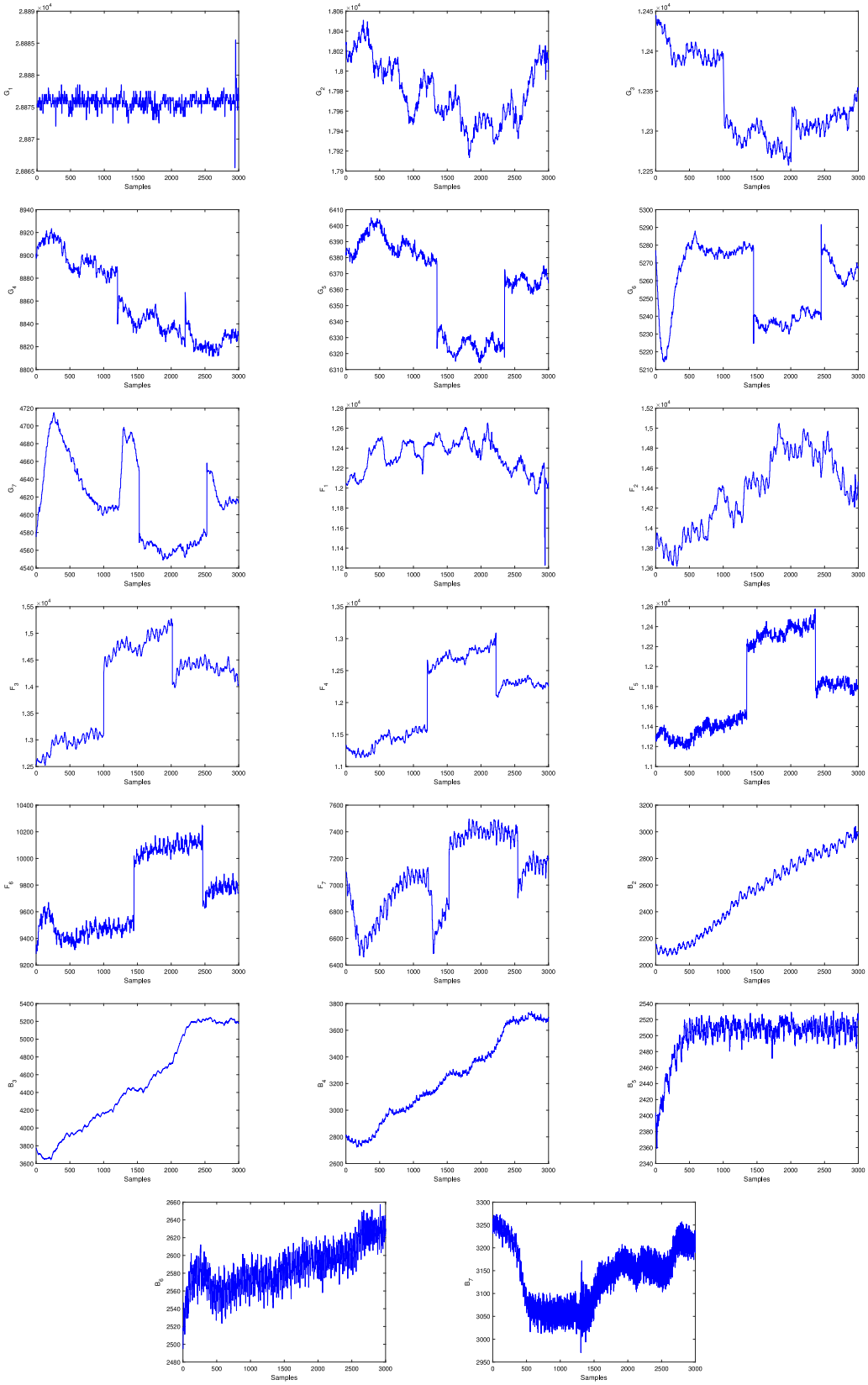


Fig. 10. Variable trajectories for the second case.

G_3 can be accurately diagnosed as the root cause, and the total forces F_6, F_7, and average gaps G_4, G_5, and G_6 of the subsequent stands are affected, which coincides with the actual industrial practice. Moreover, different methods based causality analysis results are presented in Table 9. It can be shown that the AGRU-GC method obtains satisfactory correct rate and missing causalities.

Based on the above results and discussions, it can be indicated that the new systematic nonstationary causality analysis framework works

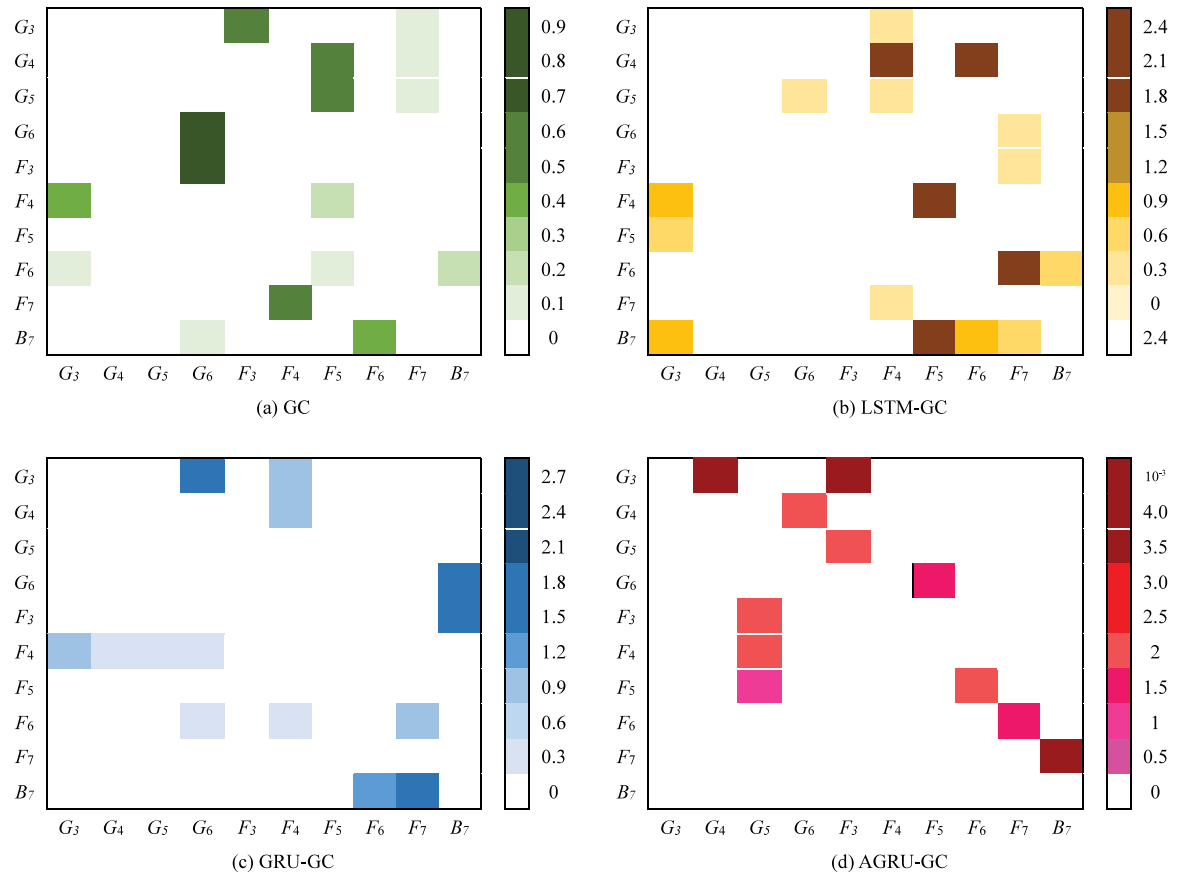


Fig. 11. Causality analysis results for the second case.

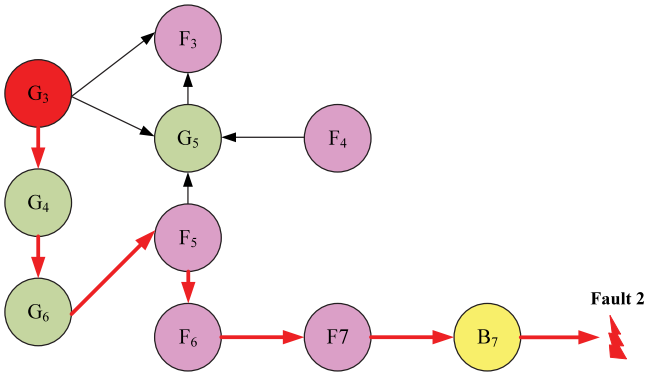


Fig. 12. Expert experience based propagation pathway for the second case.

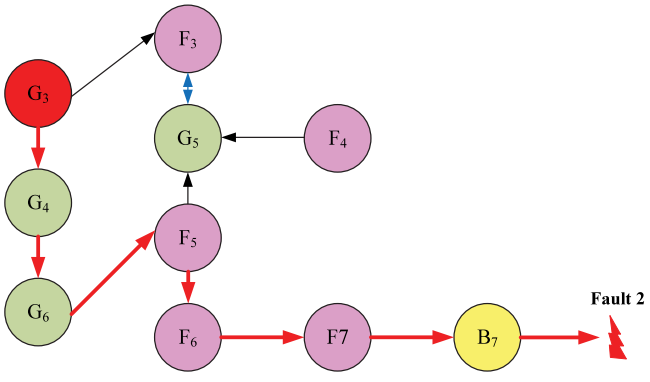


Fig. 13. AGRU-GC based propagation pathway for the second case.

Table 5
Causality analysis results for different algorithms in the first case.

Models	E	F	G
GC	0.599	0.310	0.401
LSTM-GC	0.677	0.291	0.323
GRU-GC	0.899	0.137	0.101
AGRU-GC	**0.946**	**0.083**	**0.054**

well for root cause diagnosis and propagation pathway identification of the faults in the FMP, which will offer timely operational information for field staffs.

5. Conclusions

In this work, a systematic nonstationary causality analysis scheme has been designed for root cause diagnosis of faults in manufacturing processes. Based on this new framework, the combination of KPSS and ADF tests is used for stationarity assessment. Furthermore, the Gonzalo–Granger decomposition is applied in extracting stationary factors from the nonstationary time series, and then the proposed nonlinear and dynamic causality analysis method, termed AGRU, is applied in root cause diagnosis and propagation pathway identification of faults. Subsequently, a typical manufacturing process, i.e. the HRP, is used for verification. The simulation and comparison results show that the nonlinear and dynamic causalities can be more accurately captured by the AGRU method than other competitive algorithms.

Future work will be dedicated to a systematic root cause diagnosis framework with the coexistence of continuous and binary variables in nonstationary manufacturing processes.

Table 6
Nonstationary test results for the second case.

Variables	KPSS statistic	ADF statistic	Test results	Variables	KPSS statistic	ADF statistic	Test results
G_1	−6.617	−19.021	Stationary	F_4	−2.098	−2.771	Nonstationary
G_2	−5.348	−10.065	Stationary	F_5	−1.470	−1.033	Nonstationary
G_3	−3.110	−6.316	Nonstationary	F_6	−2.336	−2.306	Nonstationary
G_4	−2.607	−5.424	Nonstationary	F_7	−1.921	−1.991	Nonstationary
G_5	−3.002	−3.169	Nonstationary	B_2	−16.012	−16.281	Stationary
G_6	−2.776	−2.066	Nonstationary	B_3	−14.802	−13.335	Stationary
G_7	−12.902	−12.097	Stationary	B_4	−6.268	−17.004	Stationary
F_1	−9.780	−9.150	Stationary	B_5	−13.077	−12.053	Stationary
F_2	−10.125	−16.322	Stationary	B_6	−15.190	−7.020	Stationary
F_3	−1.772	−3.268	Nonstationary	B_7	−1.468	−1.179	Nonstationary

Table 7
Nonstationary test results for stationary factors of the second case.

Variables	KPSS statistic	ADF statistic	Test results
G_3	−9.331	−8.432	Stationary
G_4	−16.120	−6.152	Stationary
G_5	−5.236	−9.330	Stationary
G_6	−7.423	−8.517	Stationary
F_3	−9.100	−12.001	Stationary
F_4	−12.129	−17.028	Stationary
F_5	−11.447	−11.347	Stationary
F_6	−8.012	−10.260	Stationary
F_7	−7.470	−9.001	Stationary
B_7	−9.131	−11.035	Stationary

Table 8
Comparison of RMSE and MAE for different hidden layers in the second case.

Model	1		2		3		4		5		6	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
LSTM-GC	0.265	0.990	0.167	0.392	**0.097**	**0.102**	0.517	0.425	0.569	0.611	0.623	0.805
GRU-GC	0.320	0.512	**0.079**	**0.080**	0.323	0.349	0.407	0.451	0.519	0.498	0.723	0.710
AGRU-GC	**0.077**	**0.085**	0.309	0.292	0.298	0.402	0.400	0.531	0.506	0.533	0.030	0.527

Table 9
Causality analysis results for different algorithms in the second case.

Models	E	F	G
GC	0.408	0.366	0.592
LSTM-GC	0.722	0.334	0.278
GRU-GC	0.909	0.127	0.091
AGRU-GC	**0.927**	**0.083**	**0.073**

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) under Grants (62003030, U21A20483, 72171171), the National Key R&D Program of China (No. 2021YFB3301200), and the Interdisciplinary Research Project for Young Teachers of USTB (Fundamental Research Funds for the Central Universities) (FRFIDRY-20-017).

References

Amin, M., Khan, F., & Imtiaz, S. (2019). Fault detection and pathway analysis using a dynamic Bayesian network. *Chemical Engineering Science*, *195*, 777–790.

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv Preprint: 1409.0473.

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

Bressler, S., & Seth, A. (2010). Wiener-granger causality: A well established methodology. *NeuroImage*, *58*, 323–329.

Chen, G. J., & Ge, Z. Q. (2020). Hierarchical Bayesian network modeling framework for large-scale process monitoring and decision making. *IEEE Transactions on Control Systems Technology*, *28*(2), 671–679.

Chen, H. T., Liu, Z. G., Alippi, C., Huang, B., & Liu, D. L. (2022). Explainable intelligent fault diagnosis for nonlinear dynamic systems: from unsupervised to supervised learning. http://dx.doi.org/10.36227/techrxiv.19101512.v1, TechRxiv.

Chen, X. T., Wang, T., Ying, R. X., & Cao, Z. B. (2021). A fault diagnosis method considering meteorological factors for transmission networks based on p systems. *Entropy*, *23*(1008).

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

Cho, K., Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. arXiv:1409.1259.

Dickey, D., & Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrics*, *49*, 1057–1072.

Duan, P., Chen, T. W., Shah, S. L., & Yang, F. (2012). Methods for root cause diagnosis of plant-wide oscillations. *AIChE Journal*, *60*, 2019–2034.

Escribano, A., & Peña, D. (2010). Cointegration and common factors. *Journal of Time Series Analysis*, *15*, 577–586.

He, R., Chen, G. M., Sun, S. F., Dong, C., & Jiang, S. Y. (2020). Attention-based long short-term memory method for alarm root-cause diagnosis in chemical processes. *Industrial and Engineering Chemistry Research*, *59*, 11559–11569.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780.

Huang, B., & Kadali, R. (2008). *Dynamic modelling, predictive control and performance monitoring, a data-driven subspace approach*. London: Springer-Verlag.

Jiang, H. L., Patwardhan, R., & Shah, S. L. (2009). Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix. *Journal of Process Control*, *19*(8), 1347–1354.

Jiang, Y. C., Yin, S., Dong, J. W., & Kaynak, O. (2021). A review on soft sensors for monitoring, control, and optimization of industrial processes. *IEEE Sensors Journal*, *21*(11), 12868–12881.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration-with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, *52*, 169–210.

Kabir, S. (2017). An overview of fault tree analysis and its application in model based dependability analysis. *Expert Systems with Applications*, *77*(1), 114–135.

Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*, 159–178.

Landman, R., Kortela, J., Sun, Q., & Jömsö-Jounela, S. (2014). Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity. *Computers & Chemical Engineering*, *71*, 446–456.

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

Lindner, B., Auret, L., & Bauer, M. (2020). A systematic workflow for oscillation diagnosis using transfer entropy. *IEEE Transactions on Control Systems Technology*, *28*, 908–919.

Maurya, M., Rengaswamy, R., & Venkatasubramanian, V. (2004). Application of signed digraphs-based analysis for fault diagnosis of chemical process flowsheets. *Engineering Applications of Artificial Intelligence*, *17*(5), 501–518.

Naghoosi, E., Huang, B., Domlan, E., & Kadali, R. (2013). Information transfer methods in causality analysis of process variables with an industrial application. *Journal of Process Control*, *23*, 1296–1305.

Qin, S. J., Dong, Y. N., Zhu, Q. Q., Wang, J., & Liu, Q. (2020). Bridging systems theory and data science: A unifying review of dynamic latent variable analytics and process monitoring. *Annual Reviews in Control*, *50*, 29–48.

Sun, Q. Q., & Ge, Z. Q. (2021). A survey on deep learning for data-driven soft sensors. *IEEE Transactions on Industrial Informatics*, *17*(9), 5853–5866.

Wang, Y. M., Lin, K., Qi, Y., Lian, Q., Feng, S. Z., Wu, Z. H., & Pan, G. (2018). Estimating brain connectivity with varying length time lags using recurrent neural network. *IEEE Transactions on Biomedical Engineering*, *65*(9), 1953–1963.

Wang, T., Liu, W., Cabrera, L. V., Wang, P., Wei, X. G., & Zang, T. L. (2022). A novel fault diagnosis method of smart grids based on memory spiking neural p systems considering measurement tampering attacks. *Information Sciences*, *596*, 520–536.

Wang, T., Liu, W., Zhao, J. B., Guo, X. K., & Terzija, V. (2020). A rough set-based bio-inspired fault diagnosis method for electrical substations. *International Journal of Electrical Power & Energy Systems*, *119*, Article 105961.

Wang, J. D., Yang, F., Chen, T. W., & Shah, S. L. (2016). An overview of industrial alarm systems: Main causes for alarm overloading, research status, and open problems. *IEEE Transactions on Automation Science and Engineering*, *13*(2), 1045–1061.

Yang, F., Duan, P., Shah, S. L., & Chen, T. W. (2014). *Capturing connectivity and causality in complex industrial processes*. USA, NY, New York: Springer.

Yu, F., Xiong, Q. L., Cao, L., & Yang, F. (2022). Stable soft sensor modeling based on causality analysis. *Control Engineering Practice*, *122*, Article 105109.

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

Yuan, X. H., Shi, J. F., & Gu, L. C. (2021). A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications*, *169*, Article 114417.

Zhang, C., Hu, D., & Yang, T. (2022). Anomaly detection and diagnosis for wind turbines using long short-term memory-based stacked denoising autoencoders and xgboost. *Reliability Engineering & System Safety*, *222*, Article 108445.

Zhang, H. J., Zhang, C., Dong, J., & Peng, K. X. (2022). A new key performance indicator oriented industrial process monitoring and operating performance assessment method based on improved hessian locally linear embedding. *International Journal of Systems Science*, http://dx.doi.org/10.1080/00207721.2022.2093420.

Zhao, C. H., Chen, J. H., & Jing, H. (2021). Condition-driven data analytics and monitoring for wide-range nonstationary and transient continuous processes. *IEEE Transactions on Automation Science and Engineering*, *18*(4), 1563–1574.

Zhao, Y., & Zhao, C. H. (2022). Dynamic multivariate threshold optimization and alarming for nonstationary processes subject to varying conditions. *Control Engineering Practice*, *124*, Article 105180.

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. D. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3212–3232.

Zhou, C., Liu, T. H., Zhu, H. Q., Li, Y. G., & Li, F. B. (2021). Nonstationary and multirate process monitoring by using common trends and multiple probability principal component analysis. *Industrial and Engineering Chemistry Research*, *60*, 18031–18044.