IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 17, 2020, accepted November 6, 2020, date of publication November 10, 2020,
date of current version November 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037117

Sensor Drift Detection Based on Discrete Wavelet Transform and Grey Models

XIAOJIA HAN[1,2,3], JING JIANG[4], (Fellow, IEEE), AIDONG XU[1,2], ATAUL BARI[4], CHAO PEI[1,2,3], AND YUE SUN[1,2,3]

[1]Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China
[2]Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China
[4]Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada

Corresponding author: Aidong Xu (xad@sia.cn)

This work was supported by the Research and Application of Key Technologies for High-Level Safety Integrity Transmitter under Grant 2018YFB2004101. The work of Xiaojia Han was supported by the UCAS Joint Ph.D. Training Program. The work of Jing Jiang and Ataul Bari was supported by a research grant from the Ontario Research Fund–Research Excellence 8 under Grant ORF-RE8.

ABSTRACT Drift detection has been a difficult problem in the field of sensor fault diagnosis. In this article, a sensor drift detection method using discrete wavelet transform (DWT) and a grey model GM(1,1) is proposed. DWT is used to separate the noise part from the trend part of the sensor data. Then, the GM(1,1) model is used for time series prediction in the trend part. Finally, residuals generated by predicted and current denoised sensor data are calculated and compared with a pre-selected threshold for drift detection. The residuals may not necessarily be Gaussian distribution. Therefore, the pre-selected threshold is chosen by using the kernel density estimation (KDE) method without Gaussian assumption. The effectiveness of the proposed method has been demonstrated using a simulated temperature sensor output from a sensor model on a continuous stirred-tank reactor (CSTR), as well as measurements from a physical temperature sensor in the nuclear power control test facility (NPCTF).

INDEX TERMS Discrete wavelet transform, fault detection, grey models, kernel density estimation, sensor drift.

I. INTRODUCTION

Sensors are essential parts of technical processes that measure some physical variables in real time. Sensor faults can lead to catastrophic consequences. For example, in Three Mile Island accident in 1979 [1], measurements from a faulty sensor that the operators were relying on played a major role in the accident sequence. Reliability of sensors is extremely important for safe and reliable operation of technical processes. In practice, many sensors can be exposed to harsh environments over a long period of time. This may lead to deterioration in some sensing elements and cause the entire sensor to malfunction. A fault is defined as an unpermitted deviation of at least one characteristic property of a variable from an acceptable behavior [2]. Sensor faults can be categorized into abrupt faults and incipient faults. Abrupt faults can typically be modeled as a step-like deviation, whereas incipient faults (slowly developing) are represented by a drift [3]. Sensor drift refers to the case where the difference between the sensor outputs and the actual value of the process variables diverge linearly with time [4]. Traditional solutions to deal with sensor drift is through periodic calibration (i.e. time-based maintenance). There are some shortcomings associated with periodic calibration: first, an industrial plant may have a large quantity of sensors, and periodic calibration for every sensor can be time-consuming and expensive. Second, in the process of calibration, mistakes can be introduced inadvertently. Furthermore, periodic calibration cannot ensure that a drift that has occurred between calibration intervals can be detected in time. Sensor fault detection refers to techniques to locate faulty sensors in a system. Using fault detection techniques, faulty sensors can be identified so that targeted calibration can be performed. This is also known as condition-based maintenance [5].

Drift detection has been a difficult problem in the field of sensor fault diagnosis. For the sensor drift detection problem, sensor drift detection approaches are typically based on hardware redundancy or analytical redundancy.

The associate editor coordinating the review of this manuscript and approving it for publication was Moussa Boukhnifer.

Hardware redundancy requires multiple physical sensors, which may be associated with higher cost, increased maintenance and extra space for installation [6]. Cross-calibration technique [7] is a method of hardware redundancy, but it is difficult for cross-calibration technique to detect faulty sensors with similar drift directions [8]. Analytical redundancy estimates the output of a sensor analytically from other correlated measurements in the system [8]. Fault detection methods based on analytical redundancy can be broadly divided into model-based and data-based methods [9]. The former [10], [11] requires accurate mathematical models, which may be difficult to obtain. The data-based fault detection methods do not require accurate models. Instead, data-based fault detection methods detect faults through analysis of fault-free training data obtained during normal (fault-free) operations [8].

Many data-driven methods have been proposed to solve the drift detection problem of sensors. Principal component analysis (PCA), which is a data-based method, is used in [12] to detect fixed and drifting biases of temperature and pressure sensors in a refrigeration and air conditioning (R&AC) system. PCA can transform correlated variables into new sets of variables that are uncorrelated and retains the key information of the original data set [13]. Nevertheless, one assumption of using PCA is that variables should obey Gaussian distribution. Furthermore, nonlinearities among different variables increases the difficulty of detecting faults [14]. A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based D-statistic and DV-contribution plots is proposed in [14]. The proposed DV-contribution plot showed more accurate fault diagnosis results compared with the PCA-based Q-contribution plot [14]. Furthermore, a multivariate chart using partial least squares (PLS) with a continuous ranked probability score (CRPS) is proposed in [15]. The proposed approach uses PLS to generate residuals, and then the CRPS-based chart is applied to reveal any abnormality [15]. The experiment results show that the proposed method can detect the drift. A canonical variate dissimilarity analysis (CVDA) method for process incipient fault detection is proposed in [16]. Both PLS and CVDA methods use correlations between variables in a system to detect faults. However, only abnormal behaviors can be detected by these methods. For locating the faulty sensor, intervention of maintenance personnel or other fault isolation methods are needed in those work [12], [14]–[16]. To detect the faulty sensor directly without fault isolation methods, a sensor fault detection method based on continuous wavelet transform and image analysis techniques is proposed in [17] for spike, noise, freezing and quantization faults. It can detect which faulty sensor directly without fault isolation methods, but it cannot identify sensor drift. This article solves the sensor drift fault detection problem without fault isolation methods.

A sensor drift detection method using discrete wavelet transform (DWT) and a grey model GM(1,1) is proposed in this article. An incipient fault detection approach via detrending and denoising has been proposed in [18], which is applicable particularly for detecting short term incipient faults with recovery features. A polynomial model is used in [18] to remove the trend to obtain the residual, and then perform denoising operation. In this article, sensor data is firstly decomposed into high frequency (noise portion) and low frequency (trend portion) components through DWT. The reason is that different frequency components of sensor output may contain different information about the process being measured. Therefore, different frequency components of sensor output can be used for a wide range of monitoring applications [1]. For example, high frequency components can be used for sensor abrupt fault detection [19], whereas low frequency components can be used for incipient fault detection. The drift signal is of low frequency. Therefore, low frequency components can be used for drift detection. In order to detect fault, sensor drift detection problem can be transformed into a time series prediction problem in time series data analysis [20], using grey models [21] or neural networks [22]–[24].

In the grey model, grey system theory [25], [26] is used to extract governing laws between sensor sequence data. The basic idea of the grey model is to use the original data to form the original sequence, and the new sequence is generated by the cumulative generation method. It can weaken the randomness of the original data and make it show more obvious characteristics. It is found that the cumulative generation curve is an approximate exponential growth curve, and exponential growth is in the form of differential equations. Grey model is established as the differential equation model for the generated new sequence. The GM(1,1) model represents a 1st order, 1 variable differential equation model. Compared to neural networks, the computational burden of the grey model is much lower and prediction accuracy is much higher [27]. Therefore, in the proposed method, the grey model is chosen for time series prediction about the trend part of the sensor signal. Finally, residuals generated by predicted and current denoised sensor data are computed and compared with a pre-selected threshold for drift detection. In summary, the process of obtaining current trend data by the DWT is referred to as ''denoising''. The process of obtaining predicted trend data by the grey model and calculating residuals according to predicted values is referred to as ''detrending''. The overall sensor drift detection process is shown in Fig. 1.

The effectiveness of the proposed approach has been investigated using a temperature sensor model in a continuous stirred-tank reactor (CSTR) [16] and using actual measurements from a temperature sensor of a nuclear power plant physical simulator, known as nuclear power control test facility (NPCTF) [28]. In order to describe the detection process, the following assumptions are made. (1) Fault-free sensor data can be collected while the sensor is operating normally, (2) Fault (drift) occurs while the system is operating in a steady state condition, and (3) A single fault in the form of drift is considered. The contributions of this article can be summarized as follows: 1) A sensor drift detection method

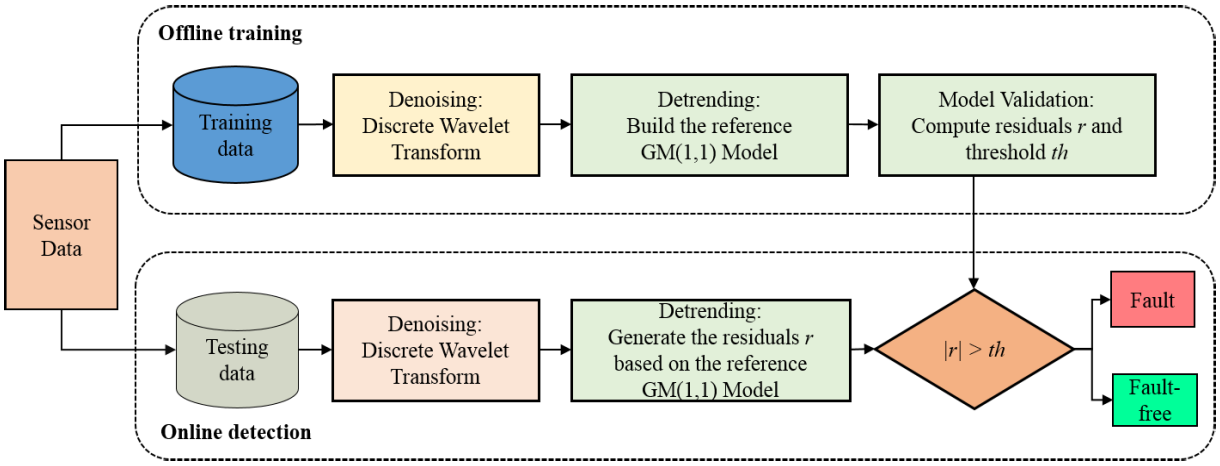


FIGURE 1. The overall process for sensor drift detection.

is developed. To the best of our knowledge, this is the first approach that has used this technique to detect the drift. 2) The residuals do not have to follow a Gaussian distribution, because the kernel density estimation (KDE) method is applied to calculate the threshold. Most of the existing approaches assume that the residuals follow a Gaussian distribution. This may be too idealistic due to process nonlinearities in actual systems and model error in residuals. The rest of the paper is organized as follows. Sensor drift detection using DWT and grey models is described in Section 2. Case studies are presented in Section 3, and conclusions are drawn in Section 4.

II. SENSOR DRIFT DETECTION USING DWT AND GREY MODELS

A. SENSOR DATA ANALYSIS

Sensor measurement data is represented by $y_s(k)$, $k = 1, 2, 3, \ldots, N$, where N is a positive integer indicating the total number of sensor readings collected in a sequence. $y_s(k)$ is a general expression for sensor data. For a fault-free sensor, $y_s(k)$ is represented as $y_{fault-free}(k)$ and for a faulty sensor, $y_s(k)$ is represented as $y_{faulty}(k)$. If the actual value of the variable being measured (referred to as the trend part) is expressed as $y(k)$, then $y_{fault-free}(k)$ can be expressed as:

$$y_{fault-free}(k) = y(k) + e(k) \tag{1}$$

where $e(k)$ is sensor noise (due to the influence of external environment and internal factors of the sensor). Under a drift fault, $y_{faulty}(k)$ deviates from the real value in a ramp manner [29]. If the sensor output drift rate is β, then the $y_{faulty}(k)$ can be expressed as:

$$y_{faulty}(k) = y(k) + \beta * k + e(k) \tag{2}$$

In brief, sensor data from a fault-free sensor includes a trend part and a noise part, whereas sensor data from a faulty sensor include a trend part mixed with a drift part and a noise part. For sensor drift detection, the sensor data is denoised in the initial step of the proposed method. Furthermore, the method first learns a normal grey model from the fault-free sensor data. This grey model is then used to predict the trend data of the faulty sensor. In order to detect the drift, residuals generated by predicted and current denoised sensor data are calculated and compared with a pre-selected threshold. The residuals from a faulty sensor should include a drift part, which is assessed to detect the fault. The specific algorithm is discussed in the following section.

B. DENOISING PROCESS USING DWT

Wavelet transform is a time-frequency analysis method. It has been applied in several fault detection scenarios, such as in heating, ventilating, and air-conditioning (HVAC) systems [30], in feed water treatment process [31] and in distillation columns [32]. Discrete wavelet transform (DWT) [33] can decompose signals into different frequency components which can be used for denoising the data or for other analysis. The purpose of DWT is to separate signals of different frequencies in sensor data. High frequency signals represent noise signals and low frequency signals represent trend signals. Since drift of the sensor is a low-frequency signal, the goal of this article is to use DWT to extract low-frequency signals in sensor measurement values, and then combine the grey model algorithm for subsequent drift fault detection. In the detection algorithm, the analysis and processing of high-frequency signals are not involved. On the other hand, the wavelet packet transform [34] will continue to decompose high-frequency signal components, which indirectly increases the computational complexity. Therefore, DWT is used to decompose sensor data $y_s(k)$ and extract the trend part $y(k)$ as follows:

The theoretical essence of wavelet transform is space partition. Given $y_s(k) \in L^2(R)$, $L^2(R)$ is decomposed into the set of subspaces W and V with J being any predetermined scale. The number of wavelet decomposition layers J has a great influence on the effect of DWT in extracting trend

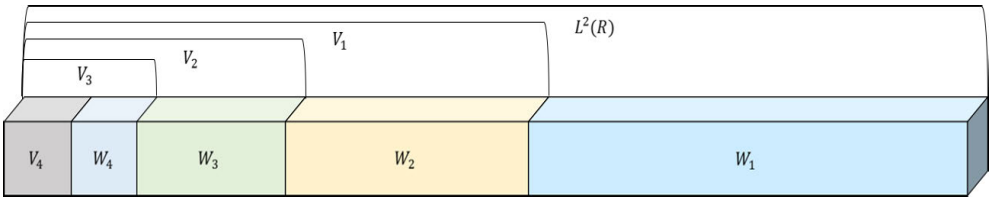


FIGURE 2. The space partition diagram of wavelet transform when $J = 4$.

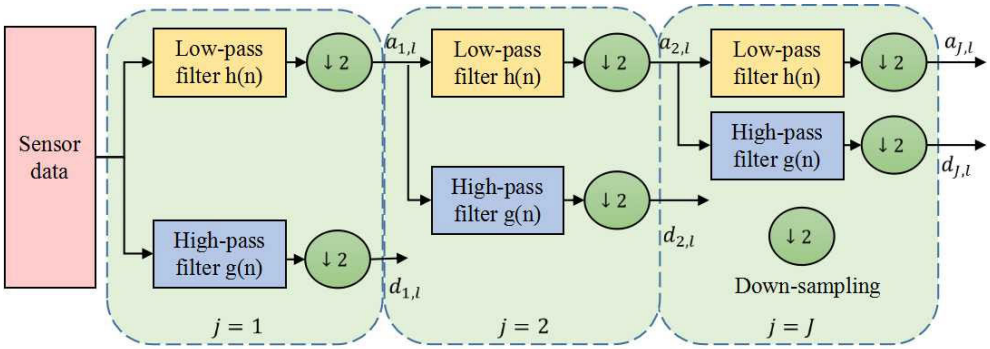


FIGURE 3. DWT process diagram.

signals. Too many decomposition layers J will cause the loss of low-frequency signal information and increase the amount of calculation. If the number of decomposition layers J is too small, the denoising effect will be unsatisfactory. Many tests show that $J = 4$ can meet the detection requirements. The space partition diagram of wavelet transform when $J = 4$ is shown in Fig. 2.

$$L^2(R) = \sum_{j=1}^{J} W_j \oplus V_J \tag{3}$$

$$V_{j-1} = V_j \oplus W_j \tag{4}$$

$$W_j \perp W_i,\ j \neq i \tag{5}$$

where the symbol $\oplus$ denotes direct summation operator and $\perp$ denotes an orthogonal operator.

Correspondingly, the signal $y_s(k)$ can be decomposed into detailed and approximate parts [33],

$$y_s(k) = \sum_{j=1}^{J} \sum_{l=-\infty}^{\infty} d_{j,l} \psi_{j,l}(k) + \sum_{l=-\infty}^{\infty} a_{J,l} \phi_{j,l}(k) \tag{6}$$

$$\psi(k) = \sqrt{2} \sum_{n} g(n)\, \phi(2k - n) \tag{7}$$

$$\phi(k) = \sqrt{2} \sum_{n} h(n)\, \phi(2k - n) \tag{8}$$

where $j, l \in Z$, $d_{j,l}$ is detailed coefficient and represents the high-frequency information, $a_{J,l}$ is approximate coefficient and represents the low frequency information of the signal. $\psi_{j,l}(k)$ is the wavelet function, $\phi_{j,l}(k)$ is the scale function.

Dual-scale equations (7)(8) [33] show the relationship of wavelet function $\psi(k)$ and scale function $\phi(k)$. The physical realization of DWT can be transformed into the process of constructing filters to realize signal decomposition. The coefficients $g(n)$ and $h(n)$ are called a pair of high-pass and low-pass wavelet filters. The coefficients $g(n)$ and $h(n)$ are determined by the wavelet function $\psi(k)$ and scale function $\phi(k)$. Through $g(n)$ and $h(n)$, the signal $y_s(k)$ is decomposed into low frequency component and high-frequency component, respectively [33]. This process is shown in Fig. 3. In order to maintain the consistency of the number of data points N before and after the decomposition, low frequency information $a_{j,l}$ and high frequency information $d_{j,l}$ in the decomposition structure are reconstructed to obtain the low frequency signal aj and the high frequency signal dj. Therefore, trend data $y(k)$ is the low frequency signal aj.

C. DETRENDING PROCESS USING GREY MODEL

Grey model can predict future changes according to the grey system theory [25]. It can be used for status self-validation of sensor arrays [27]. Trend data $y(k)$ of the sensor can be achieved in the DWT process. $Y^{(0)}$ is expressed as sensor output trend series as follows,

$$Y^{(0)} = [y(1),\ y(2),\ \ldots,\ y(M)] \tag{9}$$

where M is a positive integer and $y(k) \geq 0$. If $y(k)$ is negative, some processing is needed to be done to make them positive, such as adding a positive number. The first order accumulate generating operation (AGO) [35] of $Y^{(0)}$ is defined as

$$y^{(1)}(k) = \sum_{i=1}^{k} y(i) \tag{10}$$

The AGO is aimed at transforming an irregular, scattered series of data into a smooth and monotonically increasing

series to lessen the effects of random characteristics and noise content in the data [36]. Then new sequence $Y^{(1)}$ can be obtained. $z^{(1)}(k)$ is called the background value [27], which means it is a weighted neighbor generation data.

$$Y^{(1)} = \left[y^{(1)}(1), y^{(1)}(2), \ldots, y^{(1)}(M)\right] \tag{11}$$

$$z^{(1)}(k) = \lambda y^{(1)}(k) + (1-\lambda) y^{(1)}(k-1), \ \lambda \in [0,1] \tag{12}$$

New sequences are used to construct differential equations. Based on this differential equation, the predicted value is calculated by solving the equation. Differential equation is defined as,

$$\frac{dy^{(1)}(t)}{dt} + py^{(1)}(t) = b \tag{13}$$

where p is a developing coefficient and b is a control parameter. By discretizing (13), the grey difference equation, also known as GM(1,1) [37], can be expressed as:

$$y(k) + p\left[\lambda y^{(1)}(k) + (1-\lambda) y^{(1)}(k-1)\right] = b \tag{14}$$

The parameters p and b are estimated by using least square estimation on (14) from the fault-free sensor data.

$$[p, b]^T = \left(B^T B\right)^{-1} B^T Y \tag{15}$$

where,

$$Y = \begin{pmatrix} y(2) \\ y(3) \\ \vdots \\ y(M) \end{pmatrix} \tag{16}$$

$$B = \begin{pmatrix} -\left[\lambda y^{(1)}(2) + (1-\lambda) y^{(1)}(1)\right] & 1 \\ -\left[\lambda y^{(1)}(3) + (1-\lambda) y^{(1)}(2)\right] & 1 \\ \vdots & \vdots \\ -\left[\lambda y^{(1)}(M) + (1-\lambda) y^{(1)}(M-1)\right] & 1 \end{pmatrix} \tag{17}$$

In the case of known parameters p and b, the solution of GM(1,1) is as follows. Note that $y_{pre}^{(1)}(k+1)$ is also the predicted value of the new sequence $Y^{(1)}$.

$$y_{pre}^{(1)}(k+1) = \left[y(1) - \frac{b}{p}\right] e^{-pk} + \frac{b}{p} \tag{18}$$

Thus, the predicted value $y_{pre}(k+1)$ of the actual sequence $Y^{(0)}$ can be given as follows,

$$y_{pre}(k+1) = y_{pre}^{(1)}(k+1) - y_{pre}^{(1)}(k) \tag{19}$$

D. DETECTION PROCESS USING RESIDUALS

For drift detection, residuals generated by predicted trend data $y_{pre}(k+1)$ and the actual trend data $y(k+1)$ are calculated and compared with the threshold for fault detection. A threshold can be set based on the residuals of fault-free sensor data. However, since residuals may not satisfy Gaussian distribution, therefore, the kernel density estimation (KDE) method [16] is used to select the threshold. The chosen kernel is given in (21). The probability is calculated by estimating the probability density function of residuals from fault-free sensor data. Then the threshold th is obtained by a set confidence level α.

$$r(k+1) = y(k+1) - y_{pre}(k+1) \tag{20}$$

$$K(g) = \frac{1}{\sqrt{2}} \exp\left(-\frac{g^2}{2}\right) \tag{21}$$

$$P(-th < r < th) = \int_{-th}^{th} \frac{1}{Mh} \sum_{k=1}^{M} K\left(\frac{r - r(h)}{h}\right) = \alpha \tag{22}$$

where $r(k)$ is the kth sample of r, h is a smoothing parameter. If the absolute value of residual $r(k)$ is greater than the threshold value th, then the sensor is considered to be faulty. The flow chart of the sensor drift detection process is shown in Fig. 4.

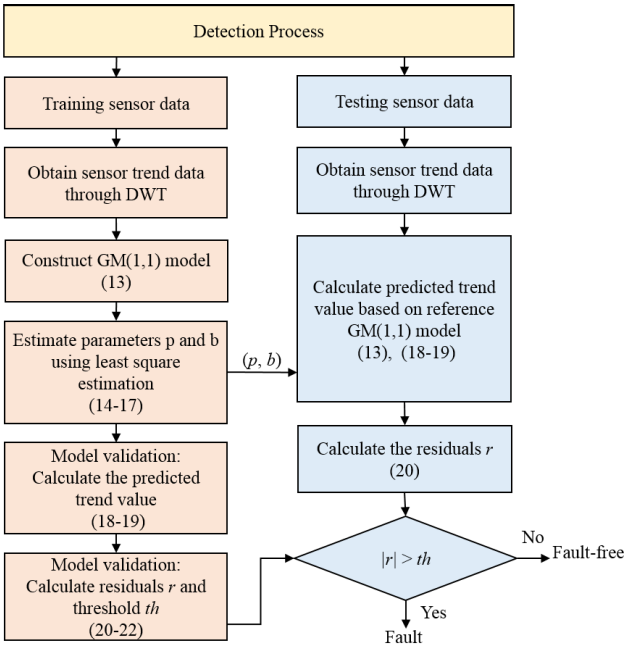


FIGURE 4. Flow chart of the sensor drift detection process.

III. CASE STUDIES

A. CSTR SIMULATION EXPERIMENT

CSTR (continuous stirred-tank reactor), shown in Fig. 5, is a Simulink simulation model [16] of a chemical reactor that has been used to investigate the effectiveness of the proposed sensor drift detection method. CSTR is a tank reactor with stirring and is a typical industrial chemical production process. The purpose of stirring is to make the reaction of materials in the container more uniform. The reaction process includes physical and chemical changes of the materials in the system, and parameters that characterize the system characteristics include temperature, pressure, liquid level, concentration and so on. As shown in Fig. 5, T is temperature, C is concentration and Q is flow rate. A temperature sensor T_{ci} in the CSTR Simulink model is selected for investigation. The temperature

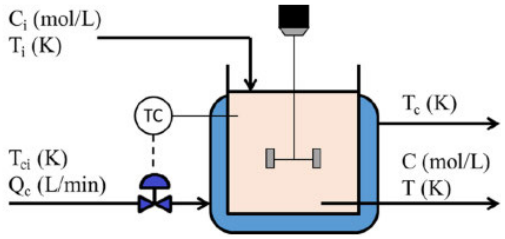


FIGURE 5. Schematic of the CSTR [16].

measurement duration is 1200 minutes. The variance of the Gaussian noise is 0.05 K^2. Simulation experiments have been performed in two scenarios, first on the training sensor data with fault free sensor and then on the testing sensor data.

For the first scenario, training T_{ci} sensor data is collected with a sampling interval of 1 min. Due to noise and other interference, T_{ci} sensor data fluctuates. Then, DWT is used to decompose sensor data $y_s(k)$ and extract the trend part $y(k)$ as discussed in Section 2. Level 4 of approximate information and db4 wavelet are chosen as level 4 contains the least noise signal. Therefore, trend data $y(k)$ is obtained by reconstructing the approximate information of level 4. The next step is to construct GM(1,1) model based on trend data $y(k)$, which essentially transforms the drift detection problem into a time series prediction problem. GM(1,1) model is built in this part. In order to validate the GM(1,1) model, another fault-free T_{ci} sensor data is collected and processed (denoised and detrended), which is shown in Fig. 6. Actual value $y(k)$ and predicted value $y_{pre}(k)$ of T_{ci} sensor trend data are obtained by the GM(1,1) model. Then residuals are calculated through actual value $y(k)$ and predicted value $y_{pre}(k)$ of the T_{ci} sensor trend data. The absolute values of residuals in the fault-free state are shown in Fig. 7 while the probability density function (PDF) curve of residuals is shown in Fig. 8. With the confidence level set to 99.9%, the threshold th is found to be 4.6468 K. It can be seen from Fig. 7 that the absolute values of the residuals are all smaller than the threshold, so the CSTR temperature sensor is in a normal state. Fig. 8 is the PDF curve of the residuals. The PDF curve has three peaks and is not a Gaussian distribution. Therefore, it cannot be

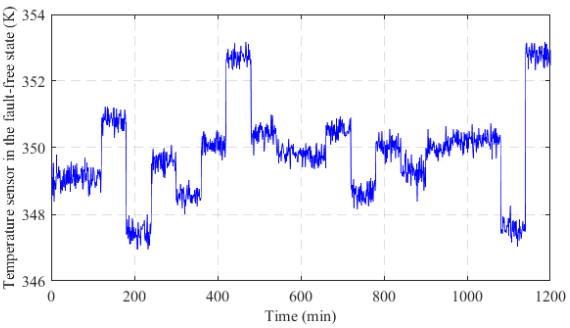


FIGURE 6. CSTR temperature sensor data in the fault-free state.

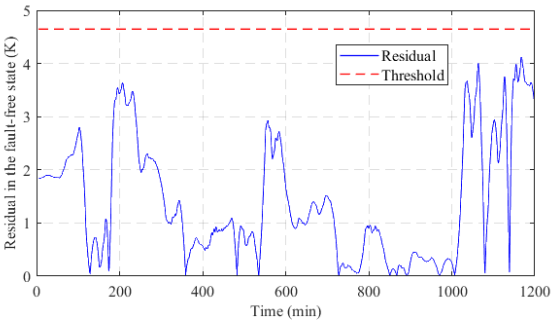


FIGURE 7. CSTR temperature sensor absolute values of residuals in the fault-free state.

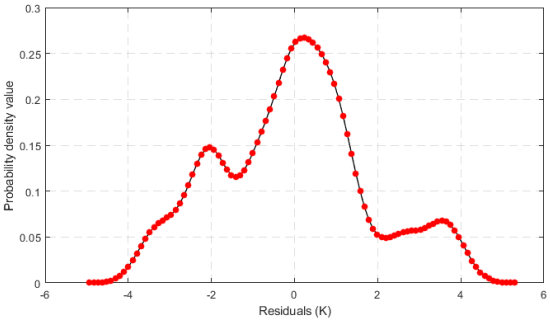


FIGURE 8. PDF of residuals in the fault-free state.

directly assumed that residuals follow a Gaussian distribution. This article uses the kernel density estimation method to calculate the threshold. This method does not have the assumption that the residuals follow a Gaussian distribution and is more realistic in practical applications.

For the second scenario, drift is introduced in the T_{ci} sensor at 200th minute while the other components in the CSTR system are working normally. The drift rate is set to be at 0.02 K/min. DWT is used to denoise, whose result is shown in Fig. 9. In the first layer of DWT, the sensor data is decomposed into $a1$ and $d1$ through wavelet analysis. $a1$ is the sensor trend signal, and $d1$ is the sensor noise signal. In order to further denoise, wavelet decomposition is performed on the sensor trend signal. $a1$ is decomposed into a sensor trend

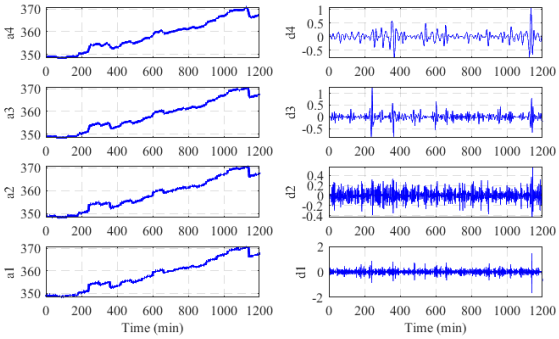


FIGURE 9. DWT result of the testing data.

signal $a2$ and a sensor noise signal $d2$. $a2$ is decomposed into a sensor trend signal $a3$ and a sensor noise signal $d3$. $a3$ is decomposed into a sensor trend signal $a4$ and a sensor noise signal $d4$. $a4$ is selected as the sensor trend data because $a4$ contains less noise. For detrending, actual value $y(k)$ and the predicted value $y_{pre}(k)$ of T_{ci} sensor data are calculated. The denoising and detrending processes use the same parameters as in the training T_{ci} sensor detection procedure. Finally, residuals are calculated through the actual value $y(k)$ and the predicted value $y_{pre}(k)$ of T_{ci} sensor trend data. The absolute values of residuals are shown in Fig. 10. The threshold th is set at 4.6468 K. As can be seen from Fig. 10, the absolute values of the residuals exceed the threshold th when $k =$ 333rd∼353rd min and $k >$ 524th min. Therefore, it can be detected that the T_{ci} sensor has a drift fault.

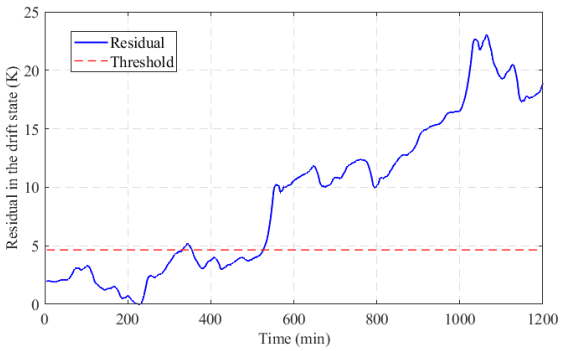


FIGURE 10. Absolute values of residuals of the testing data.

Compared to neural networks, the computational requirement (measured in terms of execution time) of the grey model is lower and the prediction accuracy is higher [27]. Here, the back propagation neural network (BPNN) [23] is used as a neural network method to compare with the grey model GM(1,1). For comparison, both were implemented on a PC platform (2.2GHz frequency and 8G memory). The programming language used in this article is MATLAB. Performance comparison of neural network and grey model is shown in the Table 1, Fig. 11 and Fig. 12. From Table 1, the execution time of BPNN is much longer than that of GM(1,1). And the mean absolute error of BPNN is also larger than that of GM(1,1). The accuracy of the prediction model directly affects the mean absolute error. The more accurate the model, the smaller the mean absolute error, and vice versa, the greater the mean absolute error. Through 5 cycles of simulation experiment comparison, the performance data of BPNN has a certain degree of instability (fluctuation) compared with GM(1,1). Therefore, grey model is chosen for time series prediction in the trend part of the sensor signal in the proposed method.

B. NPCTF EXPERIMENT

NPCTF (Nuclear Power Control Test Facility) [28] is a nuclear power plant physical simulator at Western University, which physically simulates the major components, process

TABLE 1. Performance comparison of neural network and grey model (CSTR).

Methods		Execution time(s)	Mean absolute error
BPNN	1	77.214	0.062
	2	73.801	0.051
	3	84.919	0.085
	4	71.217	0.069
	5	77.550	0.088
	Average value	76.940	0.071
GM(1,1)	1	1.842	0.011
	2	1.759	0.011
	3	1.755	0.011
	4	1.829	0.011
	5	1.721	0.011
	Average value	1.781	0.011

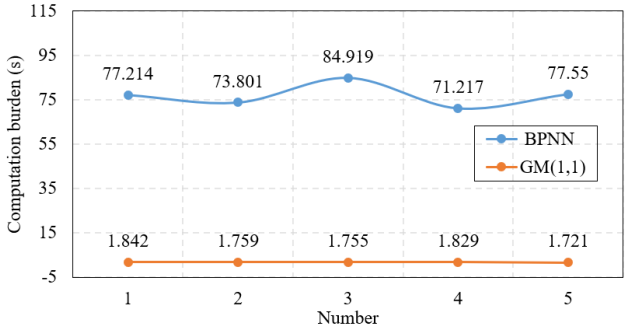


FIGURE 11. Execution time comparison NN vs GM(1,1) in CSTR.

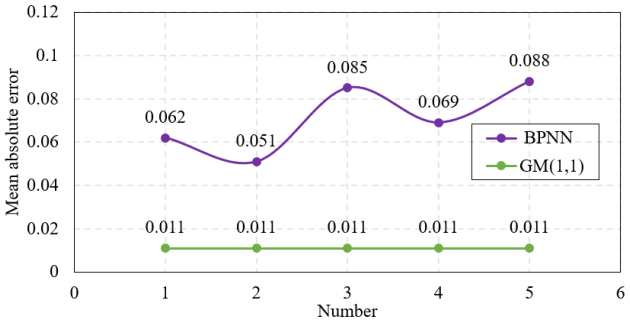


FIGURE 12. Mean absolute error comparison NN vs GM(1,1) in CSTR.

and instruments and control system of a nuclear power plant. NPCTF is a physical experiment platform for simulating the power generation process of nuclear power plants. The platform uses a heater to simulate the heat released by nuclear reactions. NPCTF transfers the heat released by the reaction by the water cycle in the whole control system. Finally, the NPCTF simulates the entire power generation process in the nuclear power plant. The Part of the control loop is shown in Fig. 13. The partial P&ID diagram of the NPCTF is shown in Fig. 13. In Fig. 13, prefix CV stands for control valve. Similarly, prefix P is pressure sensor, T is temperature sensor, F is flow sensor, L is level sensor, and Po is position of the control valve. This system has been used to investigate the effectiveness of the proposed fault detection method. A temperature sensor, $T2$, has been used. As in CSTR, Training

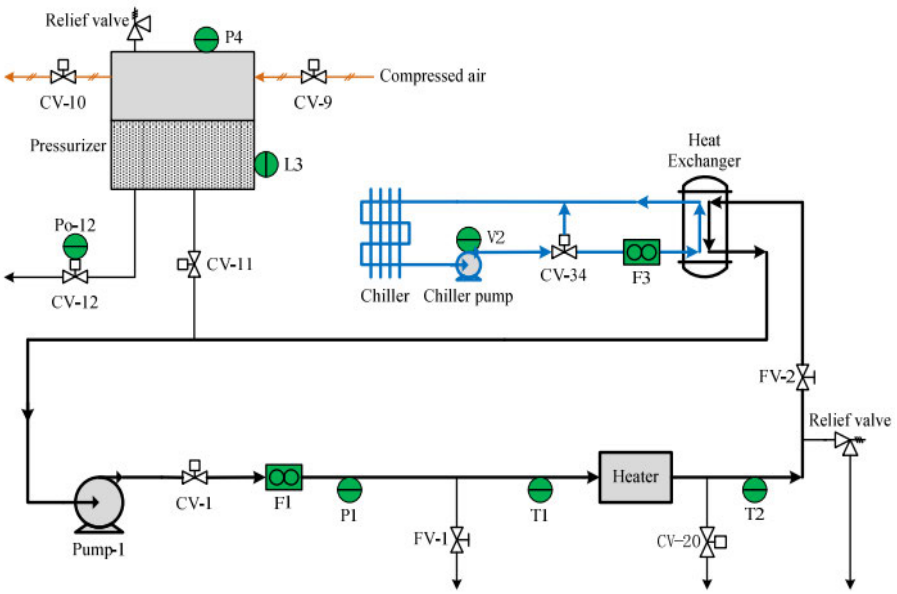


FIGURE 13. Partial P&ID diagram of NPCTF [28].

$T2$ sensor data and testing $T2$ sensor data are denoised and detrended, and the residuals are calculated. Details of the fault detection experiment process is as follows.

For the first scenario, training $T2$ sensor data is collected with sampling interval 1s, while the sensor is in a fault-free state. Then, DWT is used to obtain the trend part $y(k)$. In the next step, GM(1,1) model is constructed based on the trend data $y(k)$. In order to validate the GM(1,1) model, another set of fault-free $T2$ sensor data is collected and processed, which is shown in Fig. 14. Actual value $y(k)$ and predicted value $y_{pre}(k)$ of $T2$ sensor trend data are obtained by the GM(1,1) model. Then residuals are calculated through actual value $y(k)$ and predicted value $y_{pre}(k)$. The absolute values of residuals in the fault-free state are shown in Fig. 15 while the probability density function (PDF) curve of residuals is shown in Fig. 16. With a confidence level 99.9%, the threshold th is found to be 0.1488 °C. In the Fig. 15, the absolute values of the residuals are all smaller than the threshold, so the

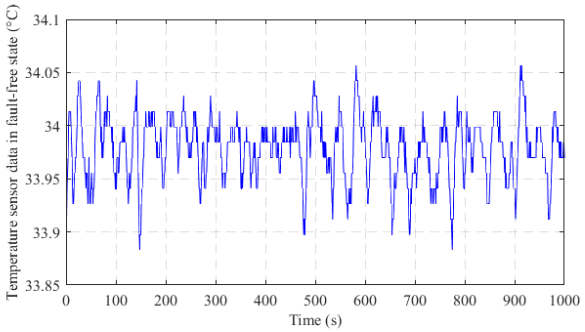


FIGURE 14. Temperature sensor data in the fault-free state.

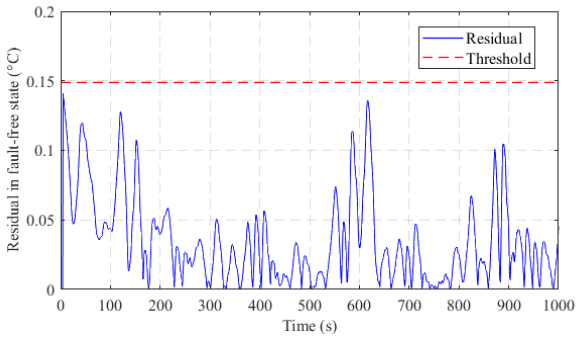


FIGURE 15. Temperature sensor absolute values of residuals in the fault-free state.

NPCTF temperature sensor is in a normal state. Fig. 16 is the PDF curve of the residuals. The PDF curve has two peaks and is not fully a Gaussian distribution. Therefore, it cannot be directly assumed that the residuals follow a Gaussian distribution. This article uses the kernel density estimation method to calculate the threshold. This method does not have the assumption that the residuals follow a Gaussian distribution, and is more realistic in practical applications.

In order to simulate the drift fault based on the real sensor data, a drift is artificially introduced at the 100th second in the temperature sensor $T2$, while other components of the NPCFT system are kept unchanged. The drift rate is set to 0.0006 °C/s. The result of DWT step is shown in Fig. 17. In the first layer of DWT, the sensor data is decomposed into $a1$ and $d1$ through wavelet analysis. $a1$ is the sensor trend signal, and $d1$ is the sensor noise signal. In order to further denoise, wavelet decomposition is performed on the sensor trend signal. $a1$ is decomposed into a sensor trend signal

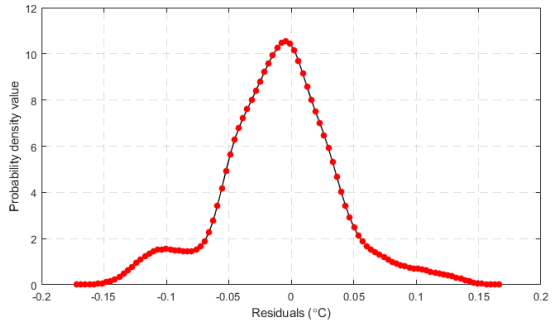


FIGURE 16. PDF of residuals in the fault-free state.

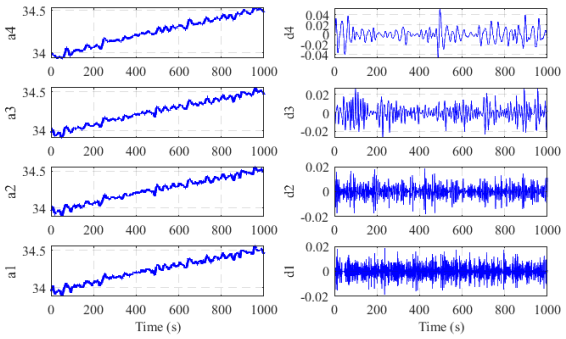


FIGURE 17. DWT result for the measurements.

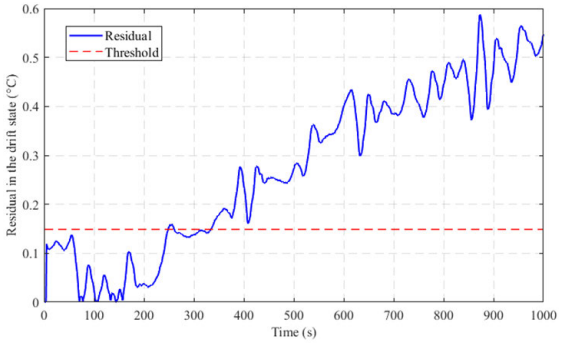


FIGURE 18. Absolute values of residuals for the measurements.

$a2$ and a sensor noise signal $d2$. $a2$ is decomposed into a sensor trend signal $a3$ and a sensor noise signal $d3$. $a3$ is decomposed into a sensor trend signal $a4$ and a sensor noise signal $d4$. $a4$ is selected as the sensor trend data because $a4$ contains less noise. Therefore, trend data $y(k)$ is obtained by reconstructing the approximate information of level 4. For detrending, predicted value $y_{pre}(k)$ is computed by GM(1,1) model. Both DWT and GM(1,1) processes use the same parameters as in the training $T2$ sensor detection procedure. Finally, residuals are calculated through actual value $y(k)$ and predicted value $y_{pre}(k)$ of $T2$. Detection result in the temperature sensor is shown in Fig. 18, which indicates that drift can be successfully detected when $k = 252$nd$\sim$262nd s and $k > 340$th s.

The back propagation neural network (BPNN) [23] is used as a neural network method to compare with the grey model GM(1,1). For comparison, both were implemented on a PC platform (2.2GHz frequency and 8G memory). The programming language used in this article is MATLAB. Performance comparison of the neural network and the grey model approaches is shown in Table 2, Fig. 19 and Fig. 20 From Table 2, The execution time of BPNN is much longer than that of GM(1,1). And the mean absolute error of BPNN is also larger than that of GM(1,1). The accuracy of the prediction model directly affects the mean absolute error. The more accurate the model, the smaller the mean absolute error, and vice versa, the greater the mean absolute error. Through 5 cycles of simulation experiment comparison, the performance data of BPNN has a certain degree of instability

TABLE 2. Performance comparison of neural network and grey model (NPCTF).

Methods		Execution time(s)	Mean absolute error
BPNN	1	77.559	0.0025
	2	71.745	0.0022
	3	70.670	0.0027
	4	73.382	0.0030
	5	74.141	0.0038
	Average value	73.499	0.0028
GM(1,1)	1	1.626	0.0009
	2	1.704	0.0009
	3	1.618	0.0009
	4	1.566	0.0009
	5	1.670	0.0009
	Average value	1.637	0.0009

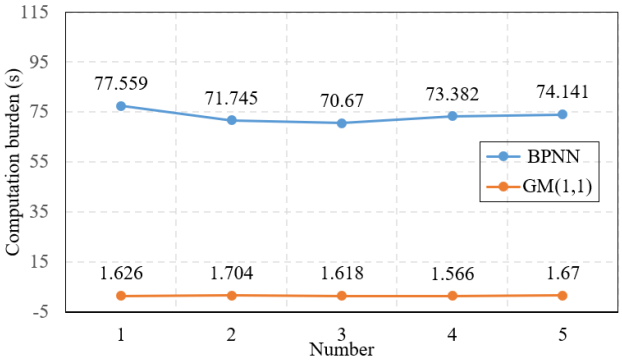


FIGURE 19. Execution time comparison NN vs GM(1,1) in NPCTF.

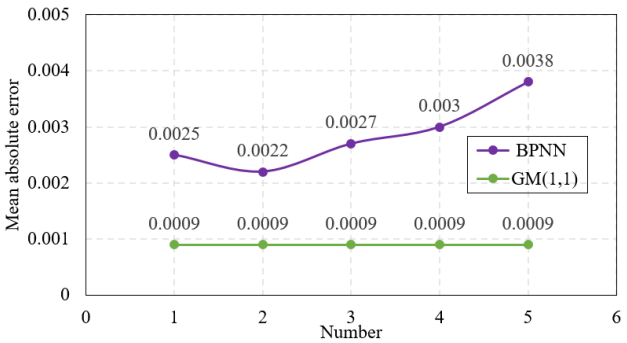


FIGURE 20. Mean absolute error comparison NN vs GM(1,1) in NPCTF.

compared with GM(1,1). Therefore, grey model is chosen for time series prediction in the trend part of the sensor signal in the proposed method.

IV. CONCLUSION

Sensor drift detection has been a difficult problem in the field of sensor fault diagnosis. Sensor drift detection on the field device side requires a non-invasive, low computational burden method. This article proposes a sensor drift detection method using discrete wavelet transform (DWT) and grey models. DWT is used to decompose the signal. Grey models are done for detrending. In the detection phase, the kernel density estimation (KDE) method is used to set the threshold in order to relax the assumption of residual Gaussian distribution. Compared with neural networks, this method has lower computational burden and higher prediction accuracy. To the best of our knowledge, this is the first approach that has used this technique to detect a sensor drift. Through CSTR and NPCTF experiments, it has been demonstrated that the proposed method can effectively detect a sensor drift.

REFERENCES

[1] H. M. Hashemian, "On-line monitoring applications in nuclear power plants," *Prog. Nucl. Energy*, vol. 53, no. 2, pp. 167–181, Mar. 2011.
[2] R. Isermann, "Model-based fault-detection and diagnosis-status and applications," *Annu. Rev. Control*, vol. 29, no. 1, pp. 71–85, 2005.
[3] H. Qiu Zhang and Y. Yan, "A wavelet-based approach to abrupt fault detection and diagnosis of sensors," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 5, pp. 1389–1396, Oct. 2001.
[4] T.-H. Yi, H.-B. Huang, and H.-N. Li, "Development of sensor validation methodologies for structural health monitoring: A comprehensive review," *Measurement*, vol. 109, pp. 200–214, Oct. 2017.
[5] R. M. Ayo-Imoru and A. C. Cilliers, "A survey of the state of condition-based maintenance (CBM) in the nuclear power industry," *Ann. Nucl. Energy*, vol. 112, pp. 177–188, Feb. 2018.
[6] X. Han, J. Jiang, A. Xu, B. Yan, and C. Pei, "A temperature sensor failure diagnosis approach in frequency," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*. New York, NY, USA: IEEE, Oct. 2016, pp. 1–4.
[7] H. M. Hashemian, *Maintenance of Process Instrumentation in Nuclear Power Plant*. Berlin, Germany: Springer-Verlag, 2006.
[8] J. Ma and J. Jiang, "Applications of fault detection and diagnosis methods in nuclear power plants: A review," *Prog. Nucl. Energy*, vol. 53, no. 3, pp. 255–266, Apr. 2011.
[9] Y. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annu. Rev. Control*, vol. 32, no. 2, pp. 229–252, Dec. 2008.
[10] P. Tadić and Ž. Đurović, "Particle filtering for sensor fault diagnosis and identification in nonlinear plants," *J. Process Control*, vol. 24, no. 4, pp. 401–409, Apr. 2014.
[11] S. Gautam, P. K. Tamboli, V. H. Patankar, S. P. Duttagupta, and K. Roy, "Real-time statistical detection and identification of sensor incipient fault using Kalman filter," in *Proc. Indian Control Conf. (ICC)*, Kanpur, India, Jan. 2018, pp. 65–70.
[12] Z. Du, L. Chen, and X. Jin, "Data-driven based reliability evaluation for measurements of sensors in a vapor compression system," *Energy*, vol. 122, pp. 237–248, Mar. 2017.
[13] W. Li, M. Peng, and Q. Wang, "Improved PCA method for sensor fault detection and isolation in a nuclear power plant," *Nucl. Eng. Technol.*, vol. 51, no. 1, pp. 146–154, Feb. 2019.
[14] G. Li, Y. Hu, H. Chen, H. Li, M. Hu, Y. Guo, S. Shi, and W. Hu, "A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based D-statistic and DV-contribution plots," *Energy Buildings*, vol. 133, pp. 230–245, Dec. 2016.
[15] F. Harrou, Y. Sun, M. Madakyaru, and B. Bouyedou, "An improved multivariate chart using partial least squares with continuous ranked probability score," *IEEE Sensors J.*, vol. 18, no. 16, pp. 6715–6726, Aug. 2018.
[16] K. E. S. Pilario and Y. Cao, "Canonical variate dissimilarity analysis for process incipient fault detection," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5308–5315, Dec. 2018.
[17] F. Cannarile, P. Baraldi, P. Colombo, and E. Zio, "A novel method for sensor data validation based on the analysis of wavelet transform scalograms," *Int. J. Prognostics Health Manage.*, vol. 9, no. 1, pp. 1–17, 2018.
[18] Z. He, Y. A. W. Shardt, D. Wang, B. Hou, H. Zhou, and J. Wang, "An incipient fault detection approach via detrending and denoising," *Control Eng. Pract.*, vol. 74, pp. 1–12, May 2018.
[19] X. Han, A. Xu, K. Wang, H. Guo, N. Zhang, Y. Liu, and S. H. Hong, "Quadratic-wavelet-transform-based fault detection approach for temperature sensor," *IEEJ Trans. Electr. Electron. Eng.*, vol. 14, no. 1, pp. 148–156, Jan. 2019.
[20] X. Wang, Q. Kang, M. Zhou, L. Pan, and A. Abusorrah, "Multi-scale drift detection test to enable fast learning in nonstationary environments," *IEEE Trans. Cybern.*, early access, Jun. 16, 2020, doi: 10.1109/TCYB.2020.2989213.
[21] S. F. Liu, *Grey System Theory and Its Applications (Eighth Edition)*. Beijing, China: Science Press 2017.
[22] R. Adhikari and R. K. Agrawal, "An introductory study on time series modeling and forecasting," 2013, *arXiv:1302.6613*. [Online]. Available: http://arxiv.org/abs/1302.6613
[23] A. G. Kavaz and B. Barutcu, "Fault detection of wind turbine sensors using artificial neural networks," *J. Sensors*, vol. 2018, pp. 1–11, Dec. 2018.
[24] X. Xu, J. W. Hines, and R. E. Uhrig, "Sensor validation and fault detection using neural networks," in *Proc. Maintenance Rel. Conf. (MARCON)*, May 1999, pp. 10–12.
[25] E. Kayacan, B. Ulutas, and O. Kaynak, "Grey system theory-based models in time series prediction," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1784–1789, Mar. 2010.
[26] S. Liu, Q. Hu, P. Li, J. Zhao, M. Liu, and Z. Zhu, "Speckle suppression based on weighted nuclear norm minimization and grey theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2700–2708, May 2019.
[27] Y. Chen, J. Yang, Y. Xu, S. Jiang, X. Liu, and Q. Wang, "Status self-validation of sensor arrays using gray forecasting model and bootstrap method," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 7, pp. 1626–1640, Jul. 2016.
[28] J. P. Ma, "Methods and systems for fault diagnosis in nuclear power plants," Ph.D. dissertation, Univ. Western Ontario, London, ON, Canada, 2015.
[29] S. Cho and J. Jiang, "Detection and estimation of sensor drifts using Kalman filters with a demonstration on a pressurizer," *Nucl. Eng. Des.*, vol. 242, pp. 389–398, Jan. 2012.
[30] S. Li and J. Wen, "A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform," *Energy Buildings*, vol. 68, pp. 63–71, Jan. 2014.
[31] S. Zhang, Q. Tang, Y. Lin, and Y. Tang, "Fault detection of feed water treatment process using PCA-WD with parameter optimization," *ISA Trans.*, vol. 68, pp. 313–326, May 2017.
[32] M. Madakyaru, F. Harrou, and Y. Sun, "Improved data-based fault detection strategy and application to distillation columns," *Process Saf. Environ. Protection*, vol. 107, pp. 22–34, Apr. 2017.
[33] R. X. Gao and R. Q. Yan, *Wavelets: Theory and Applications for Manufacturing*. New York, NY, USA: Springer, 2010.
[34] P. K. Ray, B. K. Panigrahi, P. K. Rout, A. Mohanty, F. Y. Eddy, and H. B. Gooi, "Detection of islanding and fault disturbances in microgrid using wavelet packet transform," *IETE J. Res.*, vol. 65, no. 6, pp. 796–809, 2019.
[35] T.-L. Tien, "A new grey prediction model FGM (1, 1)," *Math. Comput. Model.*, vol. 49, nos. 7–8, pp. 1416–1426, Apr. 2009.
[36] K. M. Tsang, "Sensor data validation using gray models," *ISA Trans.*, vol. 42, no. 1, pp. 9–17, Jan. 2003.
[37] K. Li, L. Liu, J. Zhai, T. M. Khoshgoftaar, and T. Li, "The improved grey model based on particle swarm optimization algorithm for time series prediction," *Eng. Appl. Artif. Intell.*, vol. 55, pp. 285–291, Oct. 2016.

XIAOJIA HAN is currently pursuing the Ph.D. degree in detection technology and automatic equipment with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. She was a Visiting Graduate Student with Western University, from 2018 to 2019. Her research interests include sensor and actuator fault detection methods, smart grid, and signal processing.

JING JIANG (Fellow, IEEE) received the Ph.D. degree from the University of New Brunswick, Fredericton, NB, Canada, in 1989.

Since 1991, he has been with the Department of Electrical and Computer Engineering, The University of Western Ontario, London, ON, Canada, where he is currently a Distinguished University Professor and a Senior Industrial Research Chair Professor. He also works closely with the International Atomic Energy Agency on modern control and instrumentation for nuclear power plants. His research interests include the fault-tolerant control of safety-critical systems, the advanced control of electrical power plants, and power systems, especially microgrids involving renewable energy resources.

Dr. Jiang is a Fellow of the Canadian Academy of Engineering. He is also a member of the International Electrotechnical Commission 45A Subcommittee to develop industrial standards on instrumentation and control for nuclear facilities.

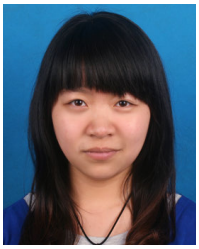
AIDONG XU received the M.E. degree in computer science from Northeastern University, in 1998, and the Ph.D. degree in electromechanical engineering from the Shenyang Institute of Automation (SIA), Chinese Academy of Sciences, in 2012.

He is currently a Research Fellow with SIA. His research interests include industrial digital communications, cyber security, and reliability. He is a member of the International Electrotechnical Commission/Industrial Process Measurement and Control Committee/Industrial Network (IEC/TC65/SC65C).

ATAUL BARI received the B.E.Sc. degree in mechanical engineering from the University of Rajshahi, Rajshahi, Bangladesh, in 1986, and the B.Sc., M.Sc., and Ph.D. degrees in computer science from the University of Windsor, ON, Canada, in 2004, 2006, and 2010, respectively.

He has over a decade of experience in the chemical and sheet metal industries, working in the capacity of both an Engineer and a Manager. His research interest includes wireless sensor networks (WSNs), particularly those used in industrial settings. His current research interests include the application of WSNs for plant and equipment condition monitoring, fault diagnosis, and for the remote monitoring of industrial environments. He has been awarded several provincial and federal level scholarships during the course of his graduate studies, including NSERC PGD-D, NSERC CGS-M, and OGS.

CHAO PEI is currently pursuing the Ph.D. degree in control theory and control engineering with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. He is currently a Visiting Ph.D. Student with the Department of Computer Science, The University of Alabama. His research interests include cyber physical security of smart grid, power system state estimation, and signal processing.



YUE SUN is currently pursuing the Ph.D. degree in detection technology and automatic equipment with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. Her research interests include rotating machinery fault diagnosis technology, and prognostic and health management technology.