Energy Proceedings
Vol 50, 2025
ISSN 2004-2965

Fault diagnosis for industrial gas turbines based on multi-fidelity data#

Baoyu Zhu, Shaojun Ren*, Qihang Weng, Yijia Zhang, Fengqi Si*

Key Laboratory of Energy Thermal Conversion and Control of Ministry of Education,

Southeast University, Nanjing 210096, China

(*Shaojun Ren:rsj@seu.edu.cn; *Fengqi Si: fqsi@seu.edu.cn)

ABSTRACT

Autoencoders (AEs) are widely used in industrial gas turbines for fault diagnosis. However, traditional AEs often perform poorly with limited training data, causing frequent false alarms. Insufficient data can cause a mismatch between the model and the actual system, especially under operational conditions absent from the training samples. This paper introduces a novel multi-fidelity autoencoder (MFAE) model that integrates limited high-fidelity operational data with abundant low-fidelity simulation data. The MFAE model learns latent features from a pre-trained low-fidelity AE model and establishes correlations between low- and high-fidelity data via linear and nonlinear networks. The effectiveness of the proposed methods is evaluated on an industrial gas turbine unit. MFAE effectively captures essential process characteristics and adapts to various operating conditions with limited operational data, enabling accurate gas turbine monitoring.

Keywords: gas turbine, multi-fidelity data, fault diagnosis, process monitoring, autoencoder

NONMENCLATURE

Abbreviations	
AE	Autoencoder
HFAE	High-fidelity autoencoder
MFAE	Muti-fidelity autoencoder
PINN	Physics-informed neural network
SPE	Squared prediction error

1. INTRODUCTION

Gas turbines, operating on the Brayton cycle, are crucial components in power generation systems. Effective process monitoring and early fault detection are vital for improving gas turbine safety, reliability, and efficiency[1]. The power industry's rapid digitalization has produced extensive real-time and historical data, enabling the development of data-driven monitoring and fault diagnosis methods[2]. These data-driven methods extract intrinsic correlations among process parameters from historical data, reducing the need for extensive prior knowledge[3].

Data-driven methods are widely adopted for process monitoring and fault diagnosis[4]. Gas turbines, however, exhibit complex nonlinear behavior due to their inherent process characteristics. Traditional statistical methods often fail to adequately capture these nonlinearities[5]. Autoencoders (AEs) have gained popularity due to their powerful nonlinear feature extraction capabilities[6]. AEs compress input data into a latent space via an encoder and reconstruct it using a decoder. Despite their potential, AEs require substantial training data, which is often lacking for new or modified gas turbine systems. Furthermore, when operating conditions change, AE models may fail to capture process characteristics accurately due to insufficient training samples for new operational modes.

Recent advancements, such as transfer learning[7], meta-learning[8], and data augmentation[9], have been proposed to address the challenges posed by small sample sizes. Additionally, physics-informed neural networks (PINNs) offer a new perspective by combining first principles and data-driven modeling [10]. However, for complex systems like gas turbines with incompletely understood mechanisms, PINNs may suffer from misspecified physical equations, resulting in suboptimal predictions [11].

Multi-fidelity modeling (MFM) integrates first principles and data-driven approaches by combining abundant, low-cost, low-fidelity simulation data with limited high-fidelity operational data [12]. Unlike PINNs, MFM uses data to characterize first principles instead of relying on explicit mathematical equations. In the MFM framework, low-fidelity simulation data represent first principles, which are then fused with high-fidelity operational data using various methods[13, 14]. Although

This is a paper for the 16th International Conference on Applied Energy (ICAE2024), Sep. 1-5, 2024, Niigata, Japan.

less accurate than operational data, simulation data provide information about the system's behavior under different conditions, capturing the underlying physical principles governing the gas turbine's operation. By leveraging low-fidelity data with the limited high-fidelity operational data, MFM techniques can construct more robust and generalizable models adaptable to diverse operating conditions.

While MFM has demonstrated significant potential in various fields [14, 15], its application to industrial process monitoring and fault diagnosis, especially for gas turbines, remains largely unexplored. This study offers the following key contributions:

(1) Development of a novel multi-fidelity autoencoder (MFAE) framework for fault monitoring, integrating limited high-fidelity operational data with abundant low-fidelity simulation data to improve process monitoring and fault diagnosis.

(2) Implementation of coupled parallel linear and nonlinear networks to identify and utilize correlations between low- and high-fidelity data, enabling accurate representation of high-fidelity process characteristics with limited operational data.

(3) Validation of the MFAE framework's effectiveness through a case study on an industrial gas turbine unit, demonstrating its superior performance over traditional single-fidelity

2. MATERIAL AND METHODS

2.1 The basic autoencoder

An AE typically consists of an encoder and a decoder. Given a training dataset $\boldsymbol{X} = [\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(N)}]^{\mathrm{T}} \in \mathbb{R}^{N\times n}$ with N samples and n variables, where $\boldsymbol{x}^{(i)} = \left[x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}\right] \in \mathbb{R}^n$, $(1 \le i \le N)$. The input variable $\boldsymbol{x}$ is projected into an l-dimensional latent spac through m encoding layers:

$$\boldsymbol{h}_m = f_m^{enc}(\boldsymbol{h}_{m-1}^{enc}) = f_m^{enc}(f_{m-1}^{enc}(\cdots f_1^{enc}(\boldsymbol{x}))) \quad (1)$$

with

$$f_i^{enc}(\boldsymbol{x}) = \sigma(\boldsymbol{W}_i^{enc}\boldsymbol{x} + \boldsymbol{b}_i^{enc}), \qquad 1 \le i \le m \quad (2)$$

where $\boldsymbol{h}_m \in \mathbb{R}^l$ is the latent variable, $f_{\mathrm{i}}^{enc}(\cdot)$ represents the function of the i-th encoding layer, $\boldsymbol{h}_{m-1}^{enc}$ is the output of the $(m-1)$-th encoding layer, $\boldsymbol{W}_i^{enc} \in \mathbb{R}^{h_i\times n}$ and $\boldsymbol{b}_i^{enc} \in \mathbb{R}^{h_i}$ are the weight matrix and bias vector of the i-th encoding layer, h_i denotes the number of neurons, $\sigma(\tau) = \max(0, \tau)$ is the rectified linear unit (ReLU) function.

The decoder then transforms the latent variables back to the original space:

$$\widehat{\boldsymbol{x}} = f_1^{dec}\big(f_2^{dec}(\cdots \sigma\big(\boldsymbol{W}_m^{dec}\boldsymbol{h}_m + \boldsymbol{b}_m^{dec}\big))\big) \quad (3)$$

with

$$f_i^{dec}(\boldsymbol{x}) = \sigma\big(\boldsymbol{W}_i^{dec}\boldsymbol{x} + \boldsymbol{b}_i^{dec}\big), \qquad 1 \le i \le m \quad (4)$$

where $\widehat{\boldsymbol{x}}$ is the AE output variable. f_i^{dec} represents the function of the i-th decoding layer, $\boldsymbol{W}_i^{dec} \in \mathbb{R}^{h_i\times l}$ and $\boldsymbol{b}_i^{dec} \in \mathbb{R}^{h_i}$ are the weight matrix and bias vector of the i-th decoding layer, respectively.

The AE training objective is to minimize the reconstruction error, which can be formulated as:

$$\underset{\boldsymbol{W}_i^{enc},\boldsymbol{b}_i^{enc},\boldsymbol{W}_i^{dec},\boldsymbol{b}_i^{dec},1\le i\le m}{\operatorname{argmin}} \quad \|\boldsymbol{x} - \widehat{\boldsymbol{x}}\|^2 \quad (5)$$

where $\|\cdot\|^2$ represents the L_2 norms.

2.2 Multi-fidelity autoencoder, MFAE

The MFAE model extends the basic AE framework by leveraging both low- and high-fidelity data to enhance process monitoring and fault diagnosis. Let $\boldsymbol{x}_{\mathrm{H}}$ represents the limited high-fidelity data accurately depicting the process, and $\boldsymbol{x}_{\mathrm{L}}$ represents the abundant low-fidelity data capturing essential process characteristics. The MFAE model is built upon a pre-trained low-fidelity AE model and learns the correlations between low- and high-fidelity data. The MFAE framework is concisely expressed as:

$$\widehat{\boldsymbol{x}}_{\mathrm{H}} = \mathcal{F}(\boldsymbol{h}_{\mathrm{H}}, \widetilde{\boldsymbol{x}}_{\mathrm{H}}) = \mathcal{F}(f_{\mathrm{L}}^{enc}(\boldsymbol{x}_{\mathrm{H}}), AE_{\mathrm{L}}(\boldsymbol{x}_{\mathrm{H}}; \boldsymbol{\Theta}_{\mathrm{L}})) \quad (6)$$

where $\widehat{\boldsymbol{x}}_{\mathrm{H}}$ and $\widetilde{\boldsymbol{x}}_{\mathrm{H}}$ denote the MFAE outputs and low-fidelity AE outputs for high-fidelity data $\boldsymbol{x}_{\mathrm{H}}$, respectively. $\mathcal{F}(\cdot)$ is an unknown mapping function relating low-fidelity features to high-fidelity data. $\boldsymbol{h}_{\mathrm{H}}$ denotes the latent space representation of high-fidelity data $\boldsymbol{x}_{\mathrm{H}}$ obtained from the low-fidelity AE encoder $f_{\mathrm{L}}^{enc}(\cdot)$. $AE_{\mathrm{L}}(\cdot;\boldsymbol{\Theta}_{\mathrm{L}})$ represents the pre-trained low-fidelity AE model with parameters $\boldsymbol{\Theta}_{\mathrm{L}}$ capturing low-fidelity features.

The correlation function $\mathcal{F}(\cdot)$ is decomposed into linear and nonlinear components:

$$\mathcal{F} = \alpha\mathcal{F}_l + (1-\alpha)\mathcal{F}_{nl} \quad (7)$$

where α is a hyperparameter that balances the contributions of linear and nonlinear functions. $\mathcal{F}_l$ and $\mathcal{F}_{nl}$ represent the linear and the nonlinear correlation functions, respectively.

Consequently, Eq. (6) can be rewritten as:

$$\widehat{\boldsymbol{x}}_{\mathrm{H}} = \mathcal{F}(\boldsymbol{x}_{\mathrm{H}}) = \alpha\mathcal{F}_l(\boldsymbol{h}_{\mathrm{H}}, \widetilde{\boldsymbol{x}}_{\mathrm{H}}) + (1-\alpha)\mathcal{F}_{nl}(\boldsymbol{h}_{\mathrm{H}}, \widetilde{\boldsymbol{x}}_{\mathrm{H}}) \quad (8)$$

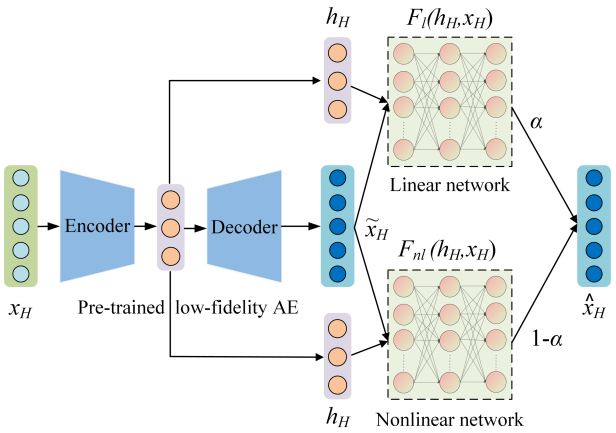

Fig. 2 Schematic diagram of the MFAE framework

MFAE learns linear and nonlinear correlations through two neural networks, with the overall structure illustrated in Fig. 2. The MFAE architecture comprises three sub-networks: a low-fidelity AE network (AE_{L}) for feature extraction, and two NN networks (NN_{H_1} and NN_{H_2}) for learning linear and nonlinear correlations, respectively. NN_{H_1} is a linear neural network without activation functions, while NN_{H_2} is a nonlinear neural network using the ReLU activation function. The outputs of NN_{H_1} and NN_{H_2} are weighted by hyperparameter α and summed to yield the final MFAE outputs $\hat{\boldsymbol{x}}_{\mathrm{H}}$. The overall MFAE framework is summarized as:

$$\tilde{\boldsymbol{x}}_{\mathrm{H}} = AE_{\mathrm{L}}(\boldsymbol{x}_{\mathrm{H}}; \boldsymbol{\Theta}_{\mathrm{L}}) \tag{9}$$

$$\boldsymbol{h}_{\mathrm{H}} = f_{\mathrm{L}}^{enc}(\boldsymbol{x}_{\mathrm{H}}) \tag{10}$$

$$\hat{\boldsymbol{x}}_{\mathrm{H}} = \alpha NN_{\mathrm{H}_1}(\boldsymbol{h}_{\mathrm{H}}, \tilde{\boldsymbol{x}}_{\mathrm{H}}; \boldsymbol{\Theta}_{\mathrm{H}_1}) + (1-\alpha) NN_{\mathrm{H}_2}(\boldsymbol{h}_{\mathrm{H}}, \tilde{\boldsymbol{x}}_{\mathrm{H}}; \boldsymbol{\Theta}_{\mathrm{H}_2}) \tag{11}$$

where $\boldsymbol{\Theta}_{\mathrm{H}_1}$ and $\boldsymbol{\Theta}_{\mathrm{H}_2}$ represent the model parameters of NN_{H_1} and NN_{H_2}, respectively.

The MFAE training process involves two steps. First, the low-fidelity AE model AE_{L} is pre-trained using low-fidelity data $\boldsymbol{x}_{\mathrm{H}}$, as outlined in Section 2.1. The loss function of AE_{L} is defined as:

$$\mathcal{L}_{\mathrm{L}} = \|\boldsymbol{x}_{\mathrm{L}} - \tilde{\boldsymbol{x}}_{\mathrm{L}}\|^2 + \lambda_{s,\mathrm{L}} \|\boldsymbol{\Theta}_{\mathrm{L}}\|^2 \tag{12}$$

where $\mathcal{L}_{\mathrm{L}}$ represents the total loss of AE_{L}. $\tilde{\boldsymbol{x}}_{\mathrm{L}}$ denotes the reconstructed low-fidelity data obtained from AE_{L}, and $\boldsymbol{x}_{\mathrm{L}}$ is the input low-fidelity data. The second term represent the L_2 regularization loss of model parameters $\boldsymbol{\Theta}_{\mathrm{L}}$, used to mitigate overfitting. The hyperparameter $\lambda_{s,\mathrm{L}}$ controls the weight of the regularization term.

Second, the linear and nonlinear networks NN_{H_1} and NN_{H_2} are trained to learn correlations. It is worth noting that the network parameters of the pre-trained AE_{L} are transferred and frozen during this stage. The loss function of the MFAE framework is defined as:

$$\mathcal{L}_{\mathrm{M}} = \|\boldsymbol{x}_{\mathrm{H}} - \hat{\boldsymbol{x}}_{\mathrm{H}}\|^2 + \lambda_{s,\mathrm{H}} (\|\boldsymbol{\Theta}_{\mathrm{H}_1}\|^2 + \|\boldsymbol{\Theta}_{\mathrm{H}_2}\|^2) \tag{13}$$

where $\mathcal{L}_{\mathrm{M}}$ represents the total loss of the MFAE framework, and $\lambda_{s,\mathrm{H}}$ is a hyperparameter that balances the trade-off between regularization and regression losses.

The BP algorithm is utilized to optimize the network parameters in the above steps.

2.3 Fault diagnosis framework of MFAE

The Squared Prediction Error (SPE) statistic serves as the monitoring metric to capture the MFAE model's residual space. Fault diagnosis using the MFAE framework comprises two main stages: offline modeling and online fault monitoring. Algorithm 1 outlines the detailed steps.

Algorithm 1. fault diagnosis procedure by the MFAE
Step 1: Offline modeling: collect and initialize the low-fidelity dataset and the high-fidelity dataset.
Step 2: Develop a MFAE model in two steps, as described in Section. 2.2, and obtain the confidence limit δ_α^2.
Step 3: Online monitoring: collect a new sample $\boldsymbol{x}$.
Step 4: calculate the SPE. If $SPE > \delta_\alpha^2$, the process is considered abnormal. If $SPE < \delta_\alpha^2$, the process is in a normal state, and return to Step 2 until $SPE > \delta_\alpha^2$.
Step 5: End of the algorithm.

3. RESULTS AND DISCUSSION

3.1 Model development

This section applies the proposed MFAE to monitor a single-shaft gas turbine in a natural gas combined cycle (NGCC) power plant. The MFAE model for the target gas turbine is developed using two data sources. The first is a low-fidelity dataset generated from a simulated model based on Advanced Process Simulator (APROS) software, covering a broad range of operating conditions.

Table 1 Selected variables and datasets for establishing monitoring models of the gas turbine

Variable	Operating data ranges	Simulated data ranges	Unit
Generator power	95.1-105.2	10.4-152.87	MW
Ambient temperature	9.1-19	-20-40	℃
Compressor outlet temperature	316.2-333.9	231.1-379.6	℃
Compressor outlet pressure	13.48,14.49	6.001-14.29	bar
Turbine inlet temperature	1083.8-1105.2	720.5-1136.3	℃
Turbine outlet temperature	545.8-563.1	409.9-593.5	℃

The second is a high-fidelity dataset collected from the supervisory information system (SIS) of an NGCC power plant. This high-fidelity dataset comprises 1440 samples collected over one day at one-minute intervals, which may not capture the full range of operating dynamics. Additionally, 4,320 samples, collected over three days, were designated as the test dataset. Data normalization was performed to ensure all variables were on a comparable scale. Table 1 presents the details of the selected variables, including temperatures and pressures associated with operational safety[1].

Two models are developed: (1) a high-fidelity autoencoder (HFAE) based on the high-fidelity dataset; (2) a multi-fidelity autoencoder (MFAE) utilizing both low- and high-fidelity datasets. The MFAE architecture combines an autoencoder (AE) with both linear and nonlinear neural networks. The AE component features a symmetrical structure: six input and output neurons corresponding to operating parameters, a two-dimensional latent space, and three hidden layers with 20, 10, and 5 neurons respectively. Within the MFAE model, both linear and nonlinear neural networks employ a structure comprising a 9-neuron input layer, two hidden layers (10 and 20 neurons), and a 6-neuron output layer. To ensure a fair comparison, the HFAE model adopts the same framework as the AE in the MFAE model. All neural networks, except the linear network, use the rectified linear unit (ReLU) as the activation function. Network structures are optimized using Optuna, a hyperparameter optimization framework.

3.2 *Monitoring performance analysis and comparison*

Figure 3 illustrates the monitoring metrics for both HFAE and MFAE models. Both models exhibited excellent monitoring capabilities on the training dataset, achieving SPE values near zero. However, significant performance disparities emerged when applied to the test dataset. The MFAE model maintained consistently low and stable SPE values throughout the testing phase. In contrast, the HFAE model displayed significant SPE fluctuations during the testing phase.

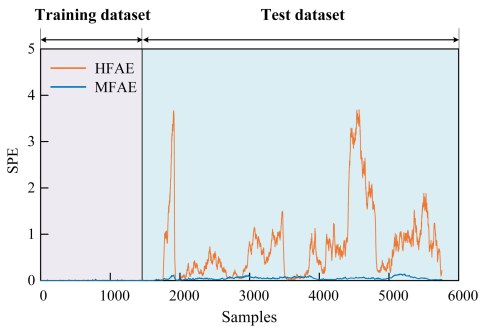

Fig. 3 Monitoring metrics for different models

Fig. 4 compares the measured and predicted values of the monitored operating parameters for the MFAE model. The normal operating ranges in the figure are established based on the outputs of the monitoring model. Measured values exceeding these intervals during normal operation trigger a false alarm, indicating suboptimal model monitoring performance. Figure 4 shows that measured values consistently fall within the normal operating ranges of the MFAE, demonstrating effective process dynamics monitoring without false alarms.

Fig. 5 reveals that the measured values notably deviated from the predicted values in the test dataset, consistent with the SPE analysis. These deviations suggest that the HFAE model fails to capture the normal operating ranges for most variables, resulting in frequent false alarms. Consequently, the HFAE model struggles to adapt to the process with wide operating ranges. The limited high-fidelity training data hinders the model's generalization ability, leading to poor monitoring performance.

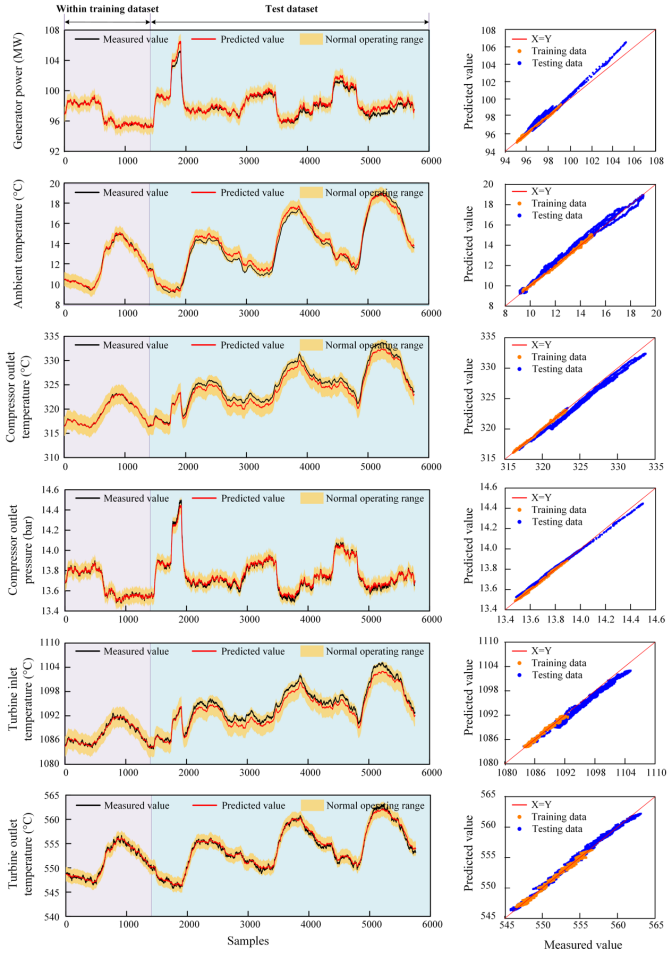

Fig. 4 Comparison of predicted and measured values for the MFAE Model

3.3 *Monitoring performance analysis and comparison*

In real-world gas turbine operations, abnormal turbine outlet temperature is a common fault, typically caused by decreased heat exchange efficiency or incomplete combustion[1]. This temperature anomaly may present as either sudden shifts or gradual changes. A continuous subset of the test dataset was selected as the fault samples to evaluate the performance of the MFAE monitoring model. The control limit for the SPE was established based on the 5,760 pieces of operating data, ensuring a comprehensive assessment of the model's fault detection capabilities under practical operating condition.

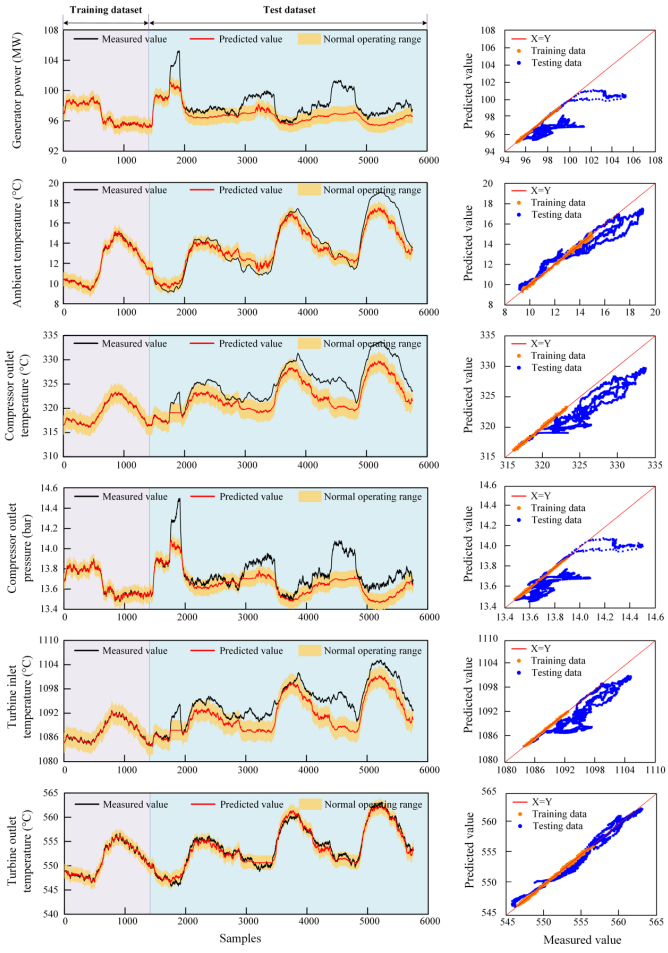

Fig. 5 Comparison of predicted and measured values for the HFAE Model

In fault case 1, we simulated a 10°C step increase between samples 2000 and 3000. Figure 6 shows that the MFAE model's SPE exhibited a sudden increase, exceeding the control limit and remaining anomalous throughout the affected sample range. This rapid response to the abrupt temperature change demonstrates the model's ability to swiftly detect and react to sudden faults. Case 2, in contrast, introduced a more subtle 0.015°C linear increase over the same sample range. Figure 7 illustrates that despite the gradual nature of this anomaly, the MFAE model's SPE rapidly exceeded the control limit and continued to rise sharply, demonstrating its high sensitivity to minor temperature deviations. These results showcase the MFAE model's robust capability to promptly detect both abrupt and subtle faults, establishing it as an effective tool for early fault warning in gas turbine systems.

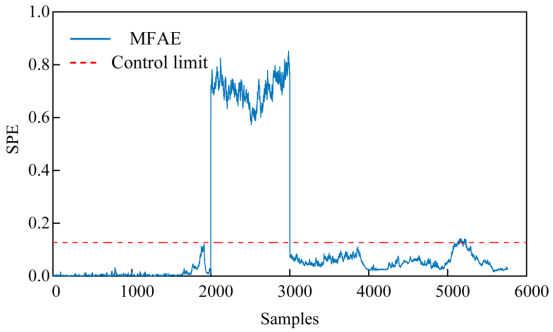

Fig. 6 Monitoring metrics of the MFAE model for fault case 1

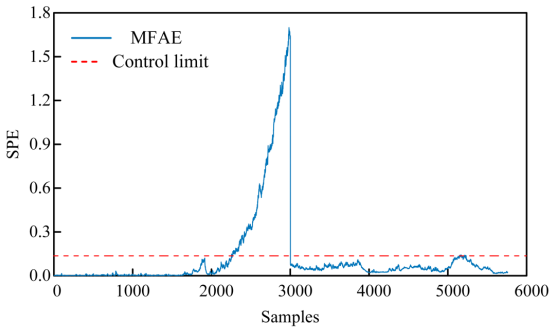

Fig. 7 Monitoring metrics of the MFAE model for fault case 2

4. CONCLUSIONS

In this study, a novel MFAE model was developed for nonlinear process monitoring and fault diagnosis. The MFAE framework extracts latent features from low-fidelity data through the AE structure and learns their correlations with high-fidelity features through the linear and nonlinear networks. This hierarchical network structure enables the MFAE to leverage abundant low-fidelity data for improved feature extraction while requiring only limited high-fidelity samples to calibrate the model, effectively ensuring its accuracy and robustness. The industrial case study highlights the MFAE's superior performance compared to single-fidelity models in process monitoring and early fault warning.

Future research directions include exploring fault isolation techniques within the MFAE framework, investigating the incorporation of prior physical knowledge, and developing adaptive strategies for online model updating as new high-fidelity data becomes available. These advancements have the potential to

extend the MFAE's applicability across diverse industrial domains, enhancing process monitoring capabilities in complex, data-constrained environments.

ACKNOWLEDGEMENT

Financial support for this research was provided by the National Key R&D Program of China (No. 2022YFB4100700) and National Natural Science Foundation of China (No. 52306230), and are gratefully acknowledged.

REFERENCE

[1] Rahme S, Meskin N. Adaptive sliding mode observer for sensor fault diagnosis of an industrial gas turbine. Control Engineering Practice 2015;38:57-74.

[2] Wang P, Ren S, Wang Y, Zhu B, Fan W, Si F. Quality-related nonlinear process monitoring of power plant by a novel hybrid model based on variational autoencoder. Control Engineering Practice 2022;129:105359.

[3] Ge ZQ. Review on data-driven modeling and monitoring for plant-wide industrial processes. Chemometrics and Intelligent Laboratory Systems 2017;171:16-25.

[4] Weese M, Martinez W, Megahed FM, Jones-Farmer LA. Statistical Learning Methods Applied to Process Monitoring: An Overview and Perspective. Journal of Quality Technology 2016;48(1):4-24.

[5] Joe Qin S. Statistical process monitoring: basics and beyond. Journal of Chemometrics 2003;17(8-9):480-502.

[6] Qian J, Song Z, Yao Y, Zhu Z, Zhang X. A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes. Chemometrics and Intelligent Laboratory Systems 2022;231:104711.

[7] Pan SJ, Yang Q. A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering 2010;22(10):1345-1359.

[8] Su H, Xiang L, Hu A, Xu Y, Yang X. A novel method based on meta-learning for bearing fault diagnosis with small sample learning under different working conditions. Mechanical Systems and Signal Processing 2022;169:108765.

[9] Chen ZS, Hou KR, Zhu MY, Xu Y, Zhu QX. A virtual sample generation approach based on a modified conditional GAN and centroidal Voronoi tessellation sampling to cope with small sample size problems: Application to soft sensing for chemical process. Applied Soft Computing 2021;101:107070.

[10] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nature Reviews Physics 2021;3(6):422-440.

[11] Zou Z, Meng X, Karniadakis GE. Correcting model misspecification in physics-informed neural networks (PINNs). Journal of Computational Physics 2024;505:112918.

[12] Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. Siam Review 2018;60(3):550-591.

[13] Partin L, Geraci G, Rushdi AA, Eldred MS, Schiavazzi DE. Multifidelity data fusion in convolutional encoder/decoder networks. Journal of Computational Physics 2023;472:111666.

[14] Li J, Li Y, Liu T, Zhang D, Xie Y. Multi-fidelity graph neural network for flow field data fusion of turbomachinery. Energy 2023;285:129405.

[15] Shi X, Liu Y, Xue L, Chen W, Chyu MK. Prediction of supercritical CO2 heat transfer behaviors by combining transfer learning and deep learning based on multi-fidelity data. International Journal of Heat and Mass Transfer 2024;218:124802.