machines

MDPI

Article

The Diagnosis and Recovery of Faults in the Workshop Environmental Control System Sensor Network Based on Medium-to-Long-Term Predictions

Shaohan Xiao, Fangping Ye *, Xinyuan Zhang, Mengying Tan and Canwen Zhang

Key Lab of Modern Manufacture Quality Engineering, Hubei University of Technology, Wuhan 430068, China; xiaoshaohan@hbut.edu.cn (S.X.); 102310043@hbut.edu.cn (X.Z.); temy@hbut.edu.cn (M.T.); 102310047@hbut.edu.cn (C.Z.)

* Correspondence: yefangping@hbut.edu.cn

Abstract

For the fault issues in the workshop environmental control system sensor network, a fault diagnosis and recovery method based on medium-to-long-term predictions is proposed. Firstly, a temperature observer based on the Informer model is established. Then, the predicted data temporarily replaces the missing real data, and the model predicts the state of the sensor system within the step size. Secondly, the predicted data is combined with the measured temperature series, and residuals are utilized for real-time detection of sensor faults. Finally, the predicted data at the time of the fault replaces the real data, enabling the recovery of fault data; experiments are conducted to verify the effectiveness of the proposed method. The results indicate that when the prediction horizon is 1, 5, 10, 20, and 50, the average fault diagnosis rates under four fault levels are 94.40%, 95.28%, 94.79%, 92.52%, and 93.35%, respectively. The average coefficients of determination for data recovery are 0.999, 0.997, 0.995, 0.985, and 0.915, respectively. This achieves medium-to-long-term predictions in the field of sensor fault diagnosis.

Keywords: sensor fault; fault diagnosis; data recovery; fault prevention; air conditioning

check for updates

Academic Editor: Nicola Ivan Giannoccaro

Received: 10 September 2025
Revised: 15 October 2025
Accepted: 21 October 2025
Published: 22 October 2025

Citation: Xiao, S.; Ye, F.; Zhang, X.; Tan, M.; Zhang, C. The Diagnosis and Recovery of Faults in the Workshop Environmental Control System Sensor Network Based on Medium-to-Long-Term Predictions. *Machines* **2025**, *13*, 975. https://doi.org/10.3390/machines13110975

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A large amount of precision equipment is continuously being deployed, placing higher demands on workshop environmental control systems [1,2]. The sensor network, serving as a critical component within the workshop's environmental control system, undertakes the vital task of environmental information acquisition, and its reliability directly impacts the performance and stability of the entire system [3]. The application of sensors in the workshop environmental control system is particularly crucial, as they are responsible for monitoring and feeding back key parameters of the system, playing a vital role in ensuring the normal operation of the environmental control system [4,5]. Therefore, research on fault diagnosis and recovery for sensor networks in workshop environmental control systems is highly significant for promoting the advancement of smart manufacturing.

In recent years, many scholars have carried out many innovative studies on fault diagnosis techniques within environmental control systems. Data-driven methods do not require the establishment of mathematical models; they utilize operational data from environmental control systems for fault diagnosis. Methods included are Support Vector Machines (SVM) [6,7], Principal Component Analysis (PCA) [8,9], and deep learning-based approaches [10], among others.

Gins et al. [11] consider the application of classification algorithms for data-driven fault diagnosis of batch processes. Classification-based methods often require a large amount of fault data to train models, whereas industrial databases typically lack sufficient fault data [12]. However, prediction models do not require fault data; instead, they use only historical normal operating data to complete fault diagnosis. Li et al. [13] propose an improved Bayesian method that combines principal component analysis and an improved Bayesian network, using a predictive approach to diagnose sensor faults in Heating Ventilation and Air Conditioning (HVAC) systems. The aforementioned methods perform well in single-step prediction tasks but are poor at long-sequence time series prediction tasks. As the demand for long-term stable system performance increases, medium- and long-term prediction has become a key task for sensor network fault diagnosis in environmental control systems.

By analyzing the historical time series collected by sensors to predict the time series of sensors for a relatively long time in the future, this task is known as the long-sequence time prediction task [14]. The Informer model, as a new Transformer-based sequence modeling method, has emerged in the field of long-sequence time series prediction with its efficient and accurate performance [15]. Guo et al. [16] develop a stacked Informer model for predicting power line trip faults, enhancing the accuracy of fault sequence prediction. Klimek et al. and Xu et al. [17,18] propose a sparse attention mechanism based on the Informer model, which improves prediction capability to some extent. Therefore, the advantages of the Informer model in long sequence time prediction are leveraged to address the fault domain within workshop environmental control system sensor networks in this article, thereby enhancing the long-term stability of system performance.

In this investigation, a fault diagnosis and recovery method for the sensor network of a workshop environmental control system based on medium- and long-term predictions is proposed. The method is composed of three parts: state prediction, fault diagnosis, and fault recovery, achieving real-time detection of faults in the sensor network and recovery of fault data. The effectiveness of the method is verified through experiments.

2. The Framework for Fault Diagnosis and Recovery

A fault diagnosis and recovery method for the sensor network of the workshop environmental control system is proposed based on medium-to-long-term forecasting. To implement this method, several key issues must be addressed: Long sequence prediction generation: Predictive sensor sequences must be generated by the model. Fault diagnosis: After the actual values are acquired, the residual between the predicted values and the measured values for each sensor is calculated. Fault localization: Faulty data points within the sensor network must be accurately located. Fault data recovery: Faulty data is recovered by utilizing the predicted data output from the Informer model to overwrite the fault data. This ensures the environmental control system can be maintained for short-term operation.

(1) Data Preprocessing: Missing data are initially filled using a one-dimensional interpolation method, followed by the elimination of unstable data points. Exponential Weighted Moving Average (EWMA) is employed for noise reduction in the acquired data to enhance data quality. Subsequently, the data are normalized to mitigate the influence of varying scales. For the experiment, the data are partitioned into training, testing, and validation sets. The training set is utilized to train the Informer model; the testing set is used to determine the fault threshold; and the validation set is employed for conducting fault diagnosis trials.

(2) Construction of Multivariate Input Medium-to-Long-Term Prediction Model: The preprocessed data are fed into the encoder of the Informer model. Within the encoder, critical features are extracted from the data via a sparse self-attention mechanism,

capturing long-term dependencies within the time series and spatial correlations among different sensors. Subsequently, these extracted features are utilized by the decoder, which predicts the entire future time series through a generative decoding mechanism. During model construction, Mean Squared Error (MSE) is adopted as the loss function, and early stopping is implemented to prevent overfitting.

(3) State Prediction: Based on the sensor prediction sequences, these predicted values are utilized as temporary references during the current system operation. This enables the future state of sensors to be anticipated in advance, thereby preventing system failures caused by state delays or unexpected events.

(4) Fault Threshold Analysis: The residual sequences between the outputs of the diagnostic soft sensors and the physical sensors on the testing set are calculated. The method described in Section 3.2 is then applied to determine the fault threshold based on these residuals.

(5) Fault Diagnosis: Fault diagnosis is performed by applying the criterion of consecutive exceedances. A data point is identified as faulty when the residuals of a sensor exceed the fault threshold continuously for five consecutive instances.

(6) Fault Data Recovery: Following the identification of faulty data points within a sensor, the predicted values output by the fault diagnosis model are employed to overwrite the faulty values. The predicted values are thus substituted for the faulty ones during the diagnostic process, achieving fault data recovery.

3. A Long-Term Predictive Fault Diagnosis Model for the Sensor Network in the Workshop Environmental Control System

3.1. Long Sequence Prediction Model Informer

The Informer model is designed specifically for long-sequence time-series forecasting, with its core lying in a unique attention mechanism. Compared to the traditional Transformer model, three distinct features are exhibited by the Informer model. Firstly, the ProbSparse self-attention mechanism is employed by the Informer. This mechanism is achieved at O(LlogL) levels in terms of time complexity and memory usage, allowing long-term dependencies within sequences to be efficiently captured, and a significant improvement in computational complexity is realized. Secondly, through the self-attention distillation technique, extremely long input sequences can be effectively handled by the Informer, leading to enhanced model processing capability and efficiency. Finally, a generative-style decoder is adopted by the Informer model. This enables the entire long time-series to be predicted in a single forward pass, rather than being predicted step by step, significantly increasing the efficiency and speed of the prediction process. The Informer model is composed of two main parts: an encoder and a decoder. Input time series are transformed into a fixed-length representation by the encoder. Predictions are then generated by the decoder based on this representation.

(1) Informer Model Input

The primary input to the Informer model consists of a time series. After determining the input window size, the model input at each time point is represented as in Equation (1).

$$X^t = \{x_1^t, x_2^t, \ldots x_i^t\} \tag{1}$$

where i represents the length of the current input sequence, and the predicted sequence output at time step t is as shown in Equation (2).

$$Y^t = \left\{y_1^t, y_2^t, \ldots y_j^t\right\} \tag{2}$$

where j represents the length of the predicted sequence output by the model at time t.

(2) Probability Sparse Self-Attention Mechanism

Scale the dot product of the Query vector, Key vector, and Value vector of the traditional self-attention mechanism, as shown in Equation (3).

$$Attention(Q, K, V) = Softmax\left(\frac{(QK^T)}{\sqrt{d}}\right)V \tag{3}$$

where $Q \in R^{L_{Q\times d}}$ represents the Query vector, $K \in R^{L_{K\times d}}$ represents the Key vector, and $V \in R^{L_{V\times d}}$ represents the Value vector, with d being the input dimension.

The attention for the i-th Query vector is as shown in Equation (4).

$$Attention(q_i, K, V) = \sum_j \frac{k(q_i, k_j)}{\sum_l k(q_i, k_l)} v_j = E_{p(k_j|q_i)}[v_j] \tag{4}$$

Self-attention probability scores exhibit sparse characteristics and a long-tail distribution, which means only a few attention scores have high values—and it is precisely these high-value scores that contribute most significantly to the overall attention. Therefore, one may consider ignoring the parts with smaller weights, thereby avoiding the calculation of the contribution of Query vectors with smaller influence on the Value vectors. The sparsity evaluation formula for Query vectors is as shown in Equation (5).

$$M(q_i, K) = ln\sum_{j=1}^{L_K} e^{\frac{q_i k_j^T}{\sqrt{d}}} - \frac{1}{L_K}\sum_{j=1}^{L_K}\frac{q_i k_j^T}{\sqrt{d}} \tag{5}$$

where the first term on the right side represents taking the maximum value for each Key component, and the second term is the calculation of the arithmetic mean.

Based on the above evaluation formula, the mathematical expression for ProbSparse self-attention is as shown in Equation (6).

$$Attention(Q, K, V) = Softmax\left(\frac{(\overline{Q}K^T)}{\sqrt{d}}\right)V \tag{6}$$

where $\overline{Q}$ is a sparse matrix of the same size as q, and it only exists in the components of Queries with ranks among the top u according to the sparsity evaluation metric M (q, K).

The u is determined by the sampling coefficient c, as shown in Equation (7).

$$u = c \cdot lnL_Q \tag{7}$$

Therefore, in sparse self-attention, only the dot product needs to be calculated for each Query–Key component, significantly reducing memory overhead and improving computational efficiency.

(3) Encoder

The encoder employs a multi-scale time encoder structure. This structure enables the model to consider information across different time scales simultaneously, thereby better capturing long-range dependencies in the sequence. The feature mapping in the encoder does not contain redundant information about Value; instead, it only includes Value vectors corresponding to Key–Value pairs with higher attention scores. Additionally, the encoder utilizes distillation techniques to reinforce dominant features, allowing for greater focus

in the subsequent layer of sparse self-attention. During this process, the model's input is continuously reduced, and the specific distillation operation is as shown in Equation (8).

$$X_{j+1}^{t} = MaxPool\left(ELU\left(Convld\left(\left[X_{j}^{t}\right]_{AB}\right)\right)\right) \tag{8}$$

where $[X^t{}_j]_{AB}$ includes multi-head probabilistic sparse self-attention and key operations within the attention blocks. *Convld* applies a one-dimensional convolution along the time dimension using the ELU activation function, followed by a max pooling layer that reduces the sample size to half its original length. To enhance the model's robustness, the encoder is stacked and its outputs are concatenated to obtain the final encoder output, enabling the processing of longer input sequences.

(4) Decoder

The decoder employs a generative decoder design. Unlike traditional step-by-step prediction methods, the generative decoder predicts the entire long sequence at once, significantly improving the inference speed for long sequence predictions. The decoder's input vector is as shown in Equation (9).

$$X_{feed_de}^{t} = Concat\left(X_{token}^{t}, X_{O}^{t}\right) \in R^{(L_{label}+L_{y})\times d_{model}} \tag{9}$$

where X_{token}^{t} is a subsequence of length L_{label} selected from the input sequence as a label, which means a labeled sequence is added before the formal prediction sequence. X_{O}^{t} is a placeholder that constitutes the overall target sequence, set to 0. Finally, a fully connected layer is used to obtain the model output, which can be used for univariate or multivariate prediction.

3.2. Sensor System State Prediction

After obtaining the predicted sequence from the sensor, since the actual values are not yet available, the predicted values are temporarily substituted for the actual values and passed to the downstream system. This process utilizes the long-term sequence predictions generated by the Informer model to predict the status of the sensor network in the workshop environmental control system. This ensures the continuous operation of the system within the prediction period and helps prevent potential failures. By predicting the sensor data in advance, the system maintains stable operation even without actual data input. The principle is illustrated in Figure 1.

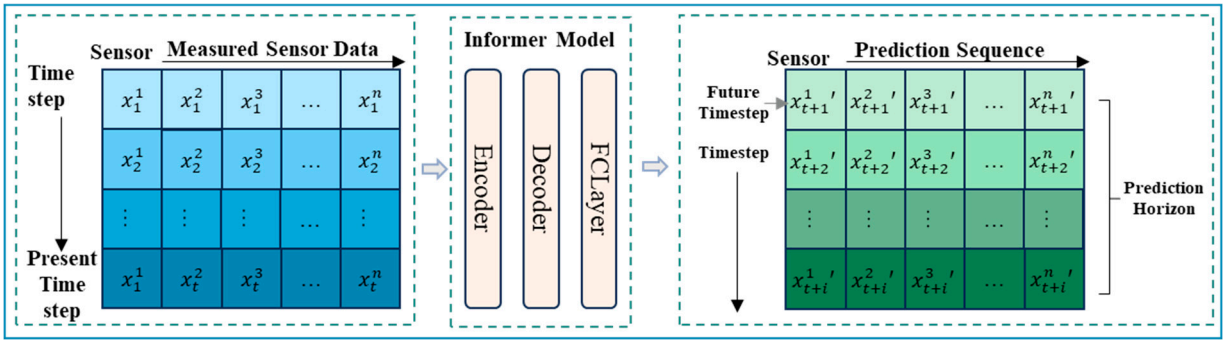


Figure 1. Principle of state prediction.

3.3. Fault Threshold Analysis

A subset is identified within the dataset for determining the fault thresholds. The residual vector of the i-th training sample at time t is given by Equation (10).

$$\varepsilon_t^{tr,i} = \hat{x_t^i} - x_t^i \tag{10}$$

where $\hat{x_t^i}$ is the time series of predicted values output by the prediction model at time t, x_t^i is the sensor time series at the t-th sampling time instant, and the vector is represented as in Equation (11).

$$x_t^i = \left(x_{t,1}^i, x_{t,2}^i, \ldots x_{t,N}^i\right)^T \in R^N \tag{11}$$

where N is the number of sensors.

The i-th threshold determines the reconstruction residual vector of the sample at time t, as shown in Equation (12).

$$\varepsilon_t^{te,i} = \left(\varepsilon_{t,1}^{te,i}, \varepsilon_{t,2}^{te,i}, \ldots \varepsilon_{t,j}^{te,i} \ldots \varepsilon_{t,N}^{te,i}\right)^T \tag{12}$$

where $\varepsilon_{t,j}^{te,i}$ is the reconstruction residual value of the j-th sensor at the t-th sampling instant, determined by the i-th threshold for the sample. The expression for the fault threshold ε_j of the j-th sensor is given in Equation (13).

$$\varepsilon_j = max\left(\left|\varepsilon_{t,j}^{te,i}\right|\right) \tag{13}$$

where N_{te} is the number of samples in the sample set for threshold determination, and $1 \leq \mathrm{I} \leq N_{te}$, $1 \leq t \leq T$. The fault threshold vector $\boldsymbol{\varepsilon}^{te,u}$ is composed of the fault thresholds of the N sensors, and its expression is given in Equation (14).

$$\varepsilon^{te,u} = (\varepsilon_1, \varepsilon_2, \ldots \varepsilon_N)^T \tag{14}$$

3.4. Residual-Based Sensor Fault Diagnosis

In the dataset, a validation set sample is partitioned, and different-sized bias faults are injected into each sensor. The reconstruction residual vector of the i-th test sample at time t is given by Equation (15).

$$\varepsilon_t^{va,i} = \left(\varepsilon_{t,1}^{va,i}, \varepsilon_{t,2}^{va,i}, \ldots \varepsilon_{t,j}^{va,i} \ldots \varepsilon_{t,N}^{va,i}\right)^T \tag{15}$$

where $\varepsilon_{t,j}^{va,i}$ is the residual value of the j-th sensor in the i-th test sample at time t.

For each sampling instance within the test sample, its reconstruction residual vector is compared element-wise with the corresponding entries in the fault threshold vector $\varepsilon^{va,u}$ to identify the sensors experiencing faults. The criterion for determining a faulty sensor is as provided in Equation (16).

$$\left|\varepsilon_{t,j}^{va,i}\right| > \varepsilon_j, \left|\varepsilon_{t,k}^{va,i}\right| \leq \varepsilon_k \tag{16}$$

where $k = 1, 2, \ldots, j-1, j+1, \ldots N$; if the residual value exceeds the threshold, then the j-th sensor is identified as faulty.

3.5. Data Recovery for Faulty Sensors

After identifying the j-th sensor as faulty, the first step is to accurately pinpoint the fault location in the j-th sensor's data and perform data recovery starting from that point. When three consecutive data points exceed the fault threshold, this point is considered the

starting point of the faulty data. Subsequently, the data is recovered starting from this point using the model's predicted output values. The specific process is shown in Figure 2.

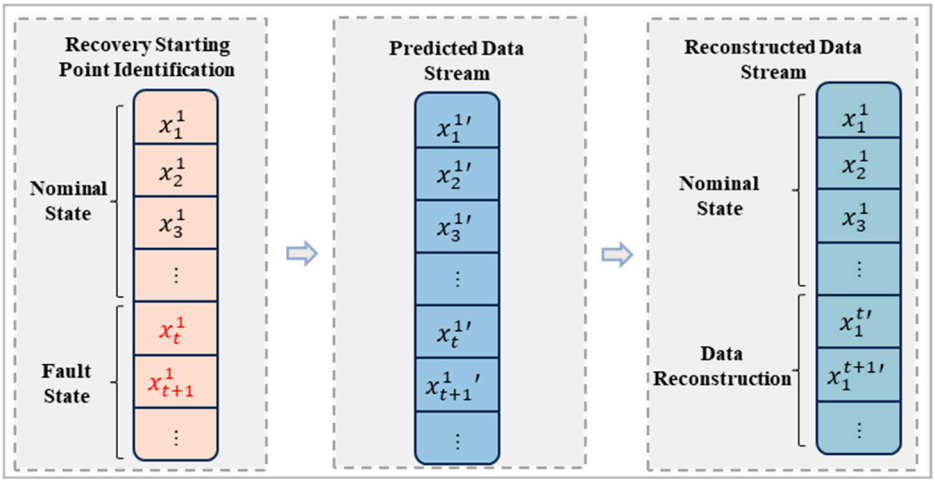


Figure 2. Initial fault data identification and recovery process.

4. Experimental Results and Analysis

4.1. Experimental Data

To verify the feasibility of the method proposed in this paper, an experiment on data acquisition from the sensor network of an environmental control system in a workshop is designed, and the experimental platform is shown in Figure 3. In the workshop environmental control system, the air conditioning system is not only regarded as the most critical component but also considered the most prone to failure. Therefore, the experiment in this study is primarily conducted on the sensor network of the air conditioning system within the workshop environmental control system. The experimental platform for the air conditioning system is primarily composed of the condenser, evaporator, compressor, expansion valve, experimental cabinet, and control cabinet. All experiments are carried out under cooling conditions in a standard enthalpy difference laboratory.

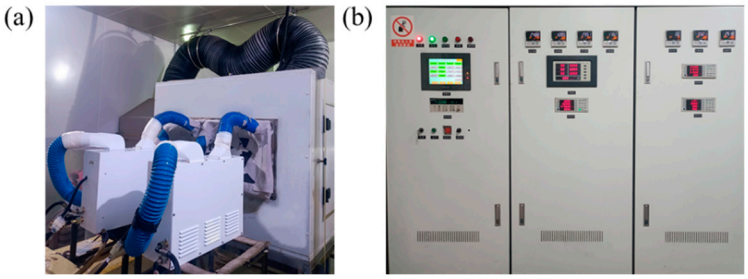


Figure 3. Environmental control system experimental platform: (**a**) experimental platform; (**b**) control system.

In the experimental platform, nine types of sensors are deployed, with details provided in Table 1. Time series data from these nine sensors are continuously collected during normal unit operation. Common summer operating conditions are simulated by adjusting the indoor and outdoor dry-bulb temperatures, with the indoor temperature regulated within a range of 23–28 °C and the outdoor temperature regulated within a range of 28–38 °C. The experimental data are collected at a frequency of 5 s. After unstable data during startup and abnormal data are removed, a total of 17,830 samples are obtained.

Table 1. Description of sensors in the environmental control system.

Number	Test Point	Symbol	Unit
1	Indoor side air inlet dry-bulb temperature	T1	°C
2	Outdoor side air inlet dry-bulb temperature	T2	°C
3	Compressor inlet temperature	T3	°C
4	Compressor outlet temperature	T4	°C
5	Condenser inlet temperature	T5	°C
6	Condenser outlet temperature	T6	°C
7	Evaporator inlet temperature	T7	°C
8	Evaporator outlet temperature	T8	°C
9	Expansion valve outlet temperature	T9	°C

Due to the involvement of equipment security, trade secrets, and other factors in obtaining industrial data specific to this application scenario and containing real faults, this study followed the common practice of preliminary research in related fields and used manual fault injection to preliminarily verify the feasibility of the method [18,19]. Among these, 15,330 samples are allocated to the training set for model training, 1500 samples are used as the test set for determining fault thresholds, and 1000 samples are designated as the validation set for conducting fault diagnosis experiments.

Since the simultaneous failure of multiple sensors in the environmental control system is considered highly improbable, this study is focused exclusively on scenarios where no more than one sensor fails at the same time. Primary emphasis is placed on the diagnosis of sensor bias faults. During the operation of the environmental control system, substantial amounts of sensor fault data are difficult to obtain. Therefore, in the validation set, we introduce biases of varying magnitudes to simulate sensor fault signals caused by bias. The minimum fault level is set at 0.5% of the mean measured value of each sensor. Faults are injected at levels corresponding to 0.5%, 1%, 1.5%, and 2% of the mean measured values of respective sensors, with all fault values rounded to two decimal places. The fault injection details for sensors T1 to T9 are presented in Table 2.

Table 2. Simulated conditions of sensors with varying levels of bias faults.

Sensors	Level 1 (0.5%)	Level 2 (1%)	Level 3 (1.5%)	Level 4 (2%)
T1	−0.13, 0.13	−0.25, 0.25	−0.38, 0.38	−0.51, 0.51
T2	−0.17, 0.17	−0.34, 0.34	−0.50, 0.50	−0.67, 0.67
T3	−0.09, 0.09	−0.17, 0.17	−0.26, 0.26	−0.34, 0.34
T4	−0.50, 0.50	−1.00, 1.00	−1.50, 1.50	−2.00, 2.00
T5	−0.39, 0.39	−0.78, 0.78	−1.17, 1.17	−1.56, 1.56
T6	−0.23, 0.23	−0.46, 0.46	−0.69, 0.69	−0.93, 0.93
T7	−0.20, 0.20	−0.40, 0.40	−0.60, 0.60	−0.80, 0.80
T8	−0.04, 0.04	−0.08, 0.08	−0.11, 0.11	−0.15, 0.15
T9	−0.06, 0.06	−0.13, 0.13	−0.19, 0.19	−0.25, 0.25

4.2. Evaluation Indicators

The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are employed to evaluate the performance of the Informer model, as shown in Equations (17) and (18). RMSE is utilized to indicate the dispersion degree of the samples, while MAE is used to assess the discrepancy between the predicted values and the actual values.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - y_i)^2} \tag{17}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|z_i - y_i| \tag{18}$$

where N is the sample quantity, z_i is the predicted value, and y_i is the true value.

To quantitatively evaluate sensor fault diagnosis performance, the Fault Detection Rate (FDR) is introduced to evaluate the detection status of sensor fault samples, as shown in Equation (19). The False Alarm Rate (FAR) is introduced to evaluate the scenario where normal sensor samples are misdiagnosed as faulty, as shown in Equation (20).

$$FDR = \frac{TP}{TP + FN} \tag{19}$$

$$FAR = \frac{FP}{FP + TN} \tag{20}$$

where FP means False Positives, TP means True Positives, FN means False Negatives, and TN means True Negatives.

To quantitatively evaluate the sensor fault data recovery performance, the Root Mean Square Error (RMSE) and the Coefficient of Determination R^2 are introduced to measure the data recovery effect, as shown in Equations (17) and (21). RMSE is used to measure the distance between the reconstructed data and the true values. R^2 is used to measure the similarity between two signals. A value of R^2 closer to 1 indicates a higher degree of similarity between the two signals, meaning that the reconstructed sensor data well characterizes the true values.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \overline{y_i})^2}{\sum_{i=1}^{n}(y_i - \dot{y})^2} \tag{21}$$

where y_i is the true value, $\overline{y_i}$ is the predicted value, and $\dot{y}$ is the average of the true values.

4.3. Analysis of Fault Diagnosis Results

4.3.1. Prediction Results Analysis

To fully exploit the correlations within the nine-dimensional sensor data from the environmental control system, the training set of nine-dimensional sensor data is input into the Informer model for training. Multivariate predictions are performed on both the test set and validation set, with prediction steps of 1, 5, 10, 20, and 50 data points conducted. The input sequence length and prediction length are maintained at a 5:1 ratio. The evaluation results of the test set predictions are presented in Table 3.

Table 3. Evaluation Metrics of the Informer Model for Different Forecasting Sequence Lengths.

Sensor	1		5		10		20		50	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
T1	0.0417	0.0316	0.0536	0.0419	0.0808	0.0636	0.1259	0.1054	0.3566	0.2674
T2	0.0524	0.0383	0.1206	0.0999	0.1252	0.0970	0.1168	0.0911	0.2910	0.2101
T3	0.0248	0.0218	0.0431	0.0349	0.0466	0.0387	0.0680	0.0568	0.1195	0.1017
T4	0.0423	0.0317	0.0549	0.0439	0.0742	0.0579	0.1111	0.0840	0.2504	0.1923
T5	0.0392	0.0299	0.0909	0.0682	0.1317	0.1011	0.2075	0.1608	0.5060	0.4106
T6	0.0663	0.0533	0.0892	0.0727	0.1210	0.0929	0.1840	0.1380	0.4586	0.3204
T7	0.0440	0.0350	0.0961	0.0734	0.1285	0.0964	0.1994	0.1515	0.5260	0.3852
T8	0.0275	0.0222	0.0366	0.0285	0.0494	0.0386	0.0680	0.0511	0.2080	0.1627
T9	0.0252	0.0186	0.0596	0.0470	0.0991	0.0791	0.1671	0.1230	0.3852	0.2829

Figure 4 shows the prediction results of the LSTM, Transformer, and Informer models for sensor T1 under different prediction steps. The Informer model is observed to achieve optimal performance across all prediction steps (1, 5, 10, 20, and 50). As the prediction step is increased, the performance of all three models is found to degrade, though the Informer's effectiveness is noted to decline more gradually. This observation demonstrates the Informer's superior stability and robustness when handling long-term sequence predictions. Under all tested prediction steps (1, 5, 10, 20, and 50), the average RMSE and MAE values of the Informer model are lower than those of the comparative models. Specifically, the average RMSE and MAE of the Informer are reduced by 38.34% and 41.46%, respectively, when compared with those of the LSTM model.

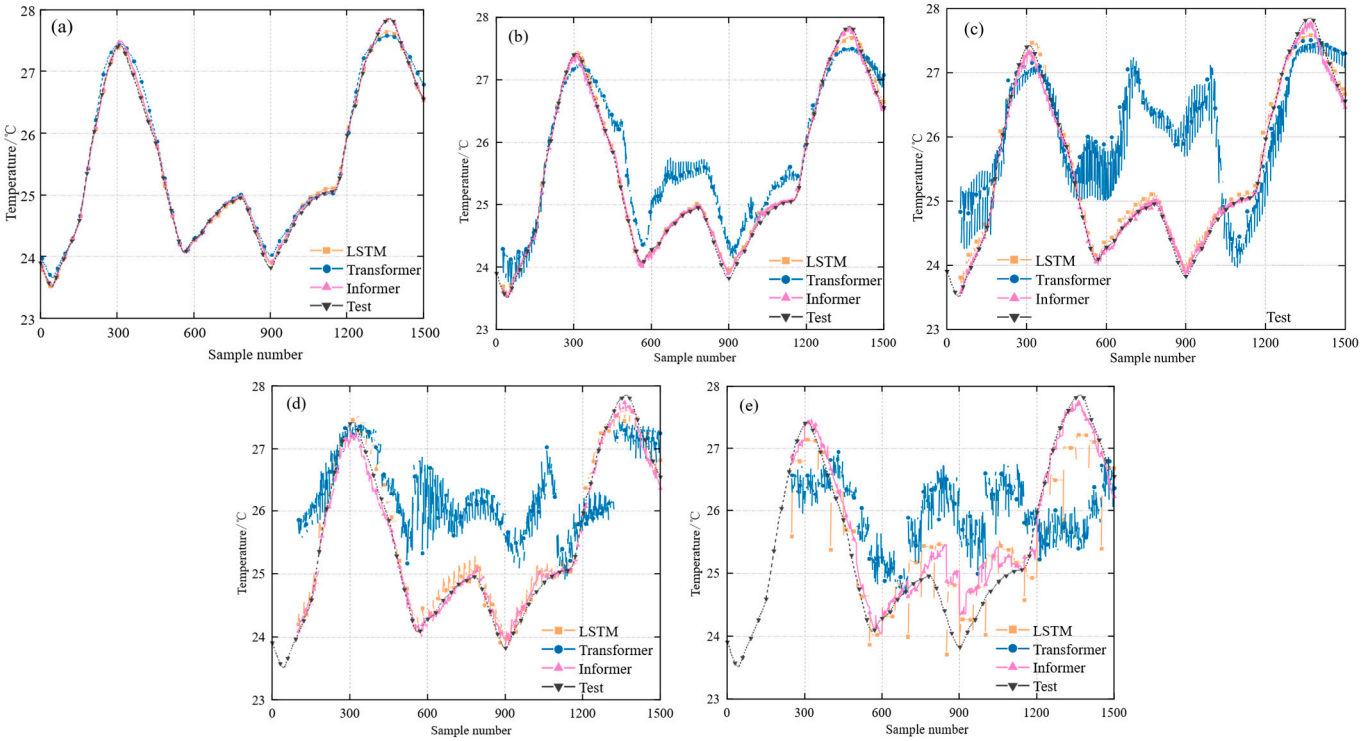


Figure 4. Prediction results of the LSTM, Transformer, and Informer models for sensor T1 under different prediction steps: (**a**) 1; (**b**) 5; (**c**) 10; (**d**) 20; (**e**) 50.

As can be observed from Table 3, a declining trend in prediction performance is demonstrated as the prediction step length increases. Specifically, both RMSE and MAE values are shown to increase with longer step lengths. When the prediction step is set

to 1, the model's predictive performance is observed to be optimal. At this step length, the lowest RMSE and MAE values across all sensors are recorded, indicating that highly accurate single-step predictions can be achieved by the model. At step lengths of 5 and 10, respectively, although prediction errors are found to increase compared to step 1, they still remain at relatively low levels, which demonstrates the model's maintained effectiveness in medium-term predictions. When step lengths of 20 and 50 are implemented, prediction errors are significantly increased, indicating reduced accuracy in long-term predictions. This trend is consistently observed across all sensors, suggesting that long-term prediction poses substantial challenges to the model, where error accumulation effects become significantly pronounced. The Informer model is shown to perform optimally in single-step predictions (step = 1), maintain satisfactory effectiveness in medium-term predictions (steps = 5 and 10), and exhibit noticeably increased prediction errors in long-term predictions (steps = 20 and 50).

4.3.2. Analysis of Fault Thresholds and Fault Diagnosis Results

After the predicted sequence is obtained through the Informer model, the residual between the predicted value and the actual value is compared with the fault threshold to achieve fault diagnosis. Based on the threshold determination method described in Section 3.3, Table 4 shows the fault thresholds for the sensors when the prediction steps are set to 1, 5, and 10, respectively. When the calculated fault threshold exceeds 0.5% of the mean measured value of the sensor, 0.5% of the mean measured value is selected as the fault threshold. To achieve higher fault diagnosis accuracy, the fault threshold corresponding to a prediction step of 1 is chosen for subsequent fault diagnosis experiments. The residuals of the test set for sensor T1 under prediction steps of 1, 5, 10, 20, and 50 are presented in Figure 5.

Table 4. Sensor fault threshold under different prediction steps.

Sensor	1	5	10
T1	(−0.11, 0.11)	(−0.18, 0.18)	(−0.28, 0.28)
T2	(−0.15, 0.15)	(−0.37, 0.37)	(−0.47, 0.47)
T3	(−0.06, 0.06)	(−0.13, 0.13)	(−0.14, 0.14)
T4	(−0.13, 0.13)	(−0.17, 0.17)	(−0.30, 0.30)
T5	(−0.17, 0.17)	(−0.38, 0.38)	(−0.44, 0.44)
T6	(−0.18, 0.18)	(−0.33, 0.33)	(−0.44, 0.44)
T7	(−0.12, 0.12)	(−0.34, 0.34)	(−0.46, 0.46)
T8	(−0.04, 0.04)	(−0.12, 0.12)	(−0.16, 0.16)
T9	(−0.06, 0.06)	(−0.15, 0.15)	(−0.24, 0.24)

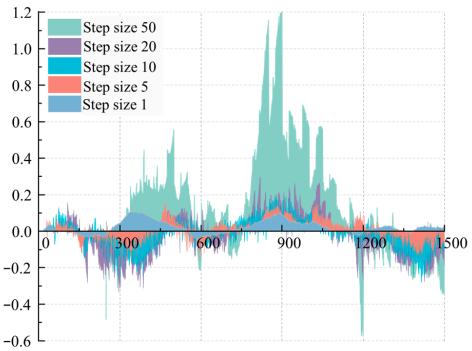


Figure 5. Residuals of Sensor T1 under different prediction steps.

By using the selected fault threshold, fault diagnosis experiments are conducted on the validation set. As shown in Figure 6, the fault diagnosis results for four sensors (T1, T4, T5, and T9) are presented. The proposed fault diagnosis method achieves satisfactory performance across different fault levels. As the fault level increases, the manifestation of faults becomes more distinct, and better fault diagnosis results are obtained. When the prediction step is set to 10, the average fault diagnosis rates for fault levels 1, 2, 3, and 4 are 85.69%, 95.17%, 98.71%, and 99.58%, respectively. The diagnostic performance is observed to improve consistently with increasing fault severity. This indicates that the proposed method not only can be effectively used to detect minor faults but also is capable of providing accurate diagnostic results under more severe fault conditions.

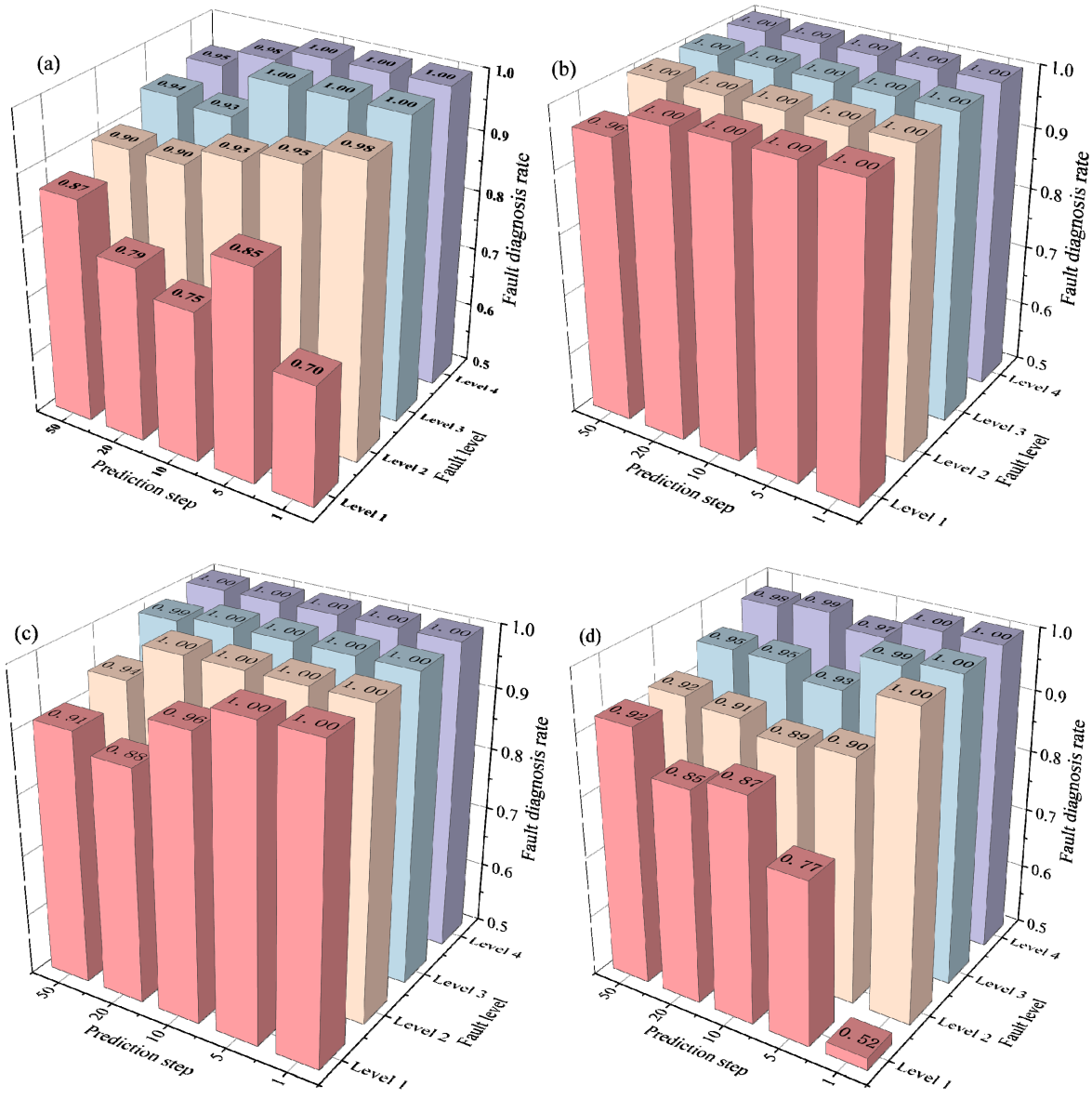


Figure 6. Sensor fault diagnosis rate for four sensors (**a**) T1; (**b**) T4; (**c**) T5; (**d**) T9.

Since the false alarm rates of sensors during fault-free phases are identical across different fault levels, Figure 7 specifically demonstrates the false alarm rates for sensors T1, T4, T5, and T9 under varying prediction steps. As the prediction step length is increased, the sensor fault diagnosis rate exhibits fluctuations but generally shows a declining trend, while the false alarm rate manifests an increasing trend. Specifically, when prediction steps are set

to 1, 5, 10, 20, and 50, respectively, the average fault diagnosis rates for T1–T9 are 94.40%, 95.28%, 94.79%, 92.52%, and 93.35%. The corresponding false alarm rates are measured at 0.00%, 0.56%, 1.78%, 20.33%, and 47.22%. When the prediction step is maintained at 10, relatively high fault diagnosis rates coupled with relatively low false alarm rates are achieved, indicating that the model possesses certain reliability and reference value for predicting sensor status within the next 50 s.

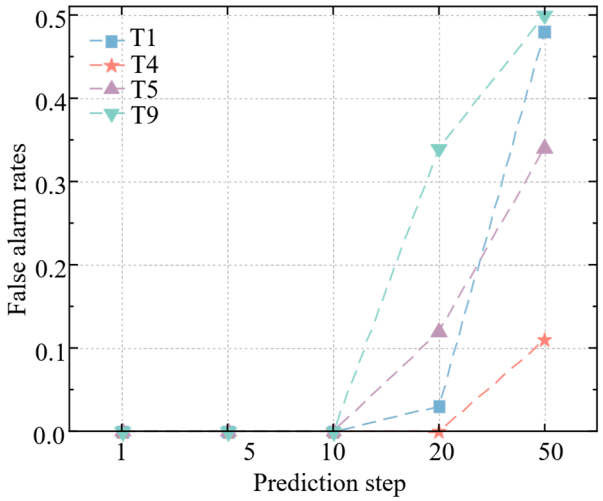


Figure 7. Distribution of False Alarm Rates for Sensor Bias Faults.

4.4. Fault Recovery Result Analysis

The recovery of sensor fault data is an effective method to enhance system reliability and stability, while also ensuring short-term normal operation. The sensor data recovery process is accomplished by first accurately identifying the initial fault point in the sensor data, then overwriting the faulty data with predicted values from the forecasting model, thereby achieving data recovery starting from that point. Figure 8 displays the fault location identification results for Sensor T5 under Level 2 conditions with prediction steps of 1, 5, and 10. In the validation set, where faults are introduced starting from the 100th data point, all sensor data are successfully maintained within the fault threshold range after recovery processing. Simultaneously, Figure 9 presents the data recovery results for Sensor T5 under prediction steps of 1, 5, and 10. The reconstructed data closely approximate the actual values, with all reconstructed data points remaining within the fault threshold range. This demonstrates that the reconstructed data can effectively recover faulty sensor data.

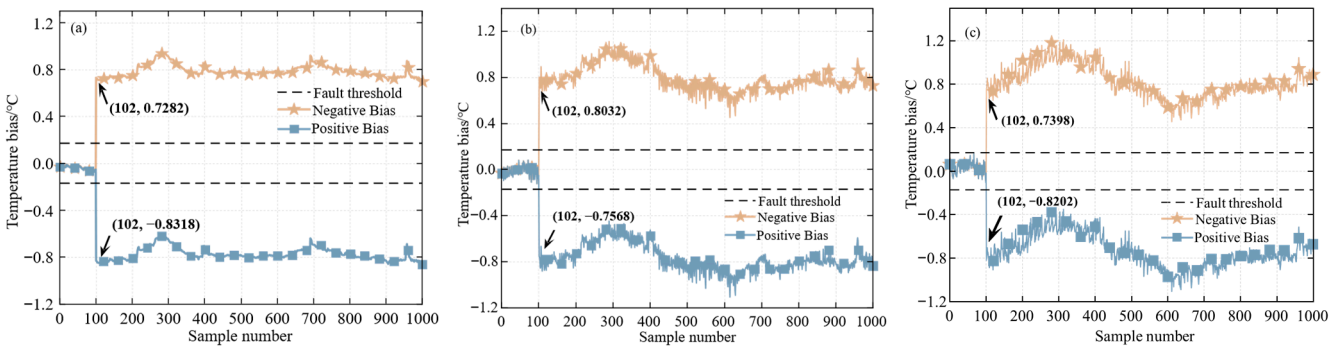


Figure 8. Fault location identification results for Sensor T5 under Level 2 conditions with different prediction steps: (**a**) 1; (**b**) 5; (**c**) 10.

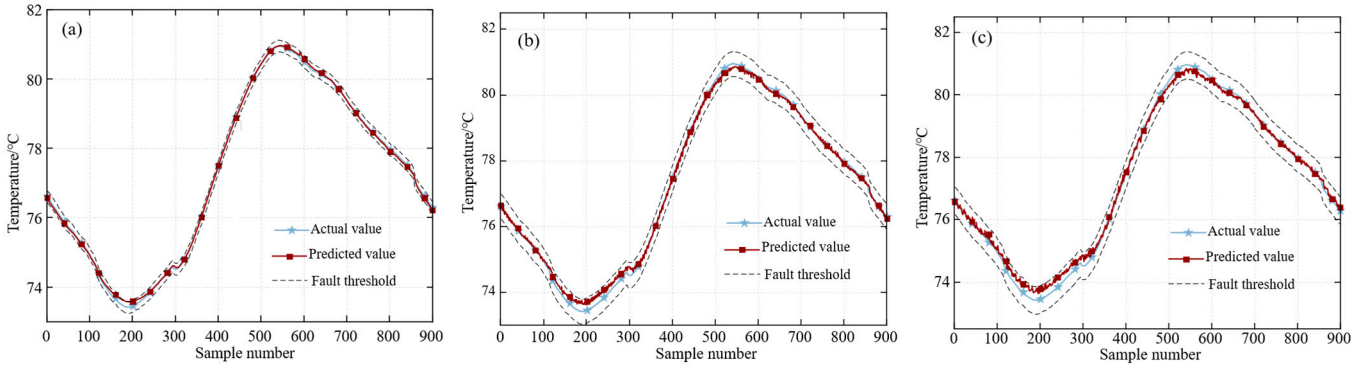


Figure 9. Data recovery results for Sensor T5 under different prediction steps: (**a**) 1; (**b**) 5; (**c**) 10.

As can be observed from Table 5, the RMSE values exhibit an increasing trend while the R^2 values demonstrate a decreasing trend as the prediction step length is extended. This pattern indicates that prediction errors are progressively increased and the model's data fitting capability is correspondingly reduced. When the prediction step length is extended beyond 20, the RMSE values are significantly elevated and R^2 values are markedly decreased, which can be attributed to the substantial challenges faced by the model in long-term predictions where forecasting accuracy is considerably diminished. However, when the prediction step length is maintained within 10, all predicted data are successfully contained within the fault threshold range. These predicted values can be effectively utilized to replace faulty data, thereby achieving successful recovery of sensor network faults in the workshop environmental control system.

Table 5. Evaluation results of faulty data recovery by fault-compensated soft sensors.

Sensor	1		5		10		20		50	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
T1	0.0565	0.9982	0.0678	0.9974	0.1111	0.9930	0.2197	0.9730	0.5595	0.8250
T2	0.0623	0.9994	0.1564	0.9965	0.1551	0.9965	0.1713	0.9958	0.3819	0.9793
T3	0.0199	0.9994	0.0484	0.9965	0.0510	0.9961	0.0924	0.9874	0.1103	0.9820
T4	0.0386	0.9992	0.0671	0.9978	0.0698	0.9976	0.0962	0.9956	0.2103	0.9790
T5	0.0500	0.9995	0.1079	0.9978	0.1428	0.9962	0.2261	0.9906	0.5633	0.9416
T6	0.0742	0.9993	0.1216	0.9982	0.1776	0.9962	0.2612	0.9918	0.6655	0.9470
T7	0.0651	0.9992	0.1392	0.9967	0.1823	0.9943	0.2776	0.9869	0.7280	0.9103
T8	0.0261	0.9987	0.0477	0.9958	0.0680	0.9915	0.1136	0.9764	0.3077	0.8274
T9	0.0231	0.9996	0.0715	0.9966	0.1192	0.9905	0.2183	0.9684	0.4905	0.8408

5. Conclusions

(1) Compared with the LSTM and Transformer models, the Informer model achieves superior performance on long time series, exhibiting the best prediction results at step lengths of 1, 5, 10, 20, and 50. This capability enables accurate state prediction within the sensor's prediction horizon. Specifically, the experiment with a prediction step of 10 achieves both high fault diagnosis rates and low false alarm rates, confirming the method's effectiveness in predicting sensor states for up to 50 s into the future.

(2) For the same sensor, improved fault diagnosis performance is observed with higher fault levels. The average fault diagnosis rates for Level 1 to Level 4 are recorded as 85.69%, 95.17%, 98.71%, and 99.58%, respectively. At prediction steps of 1, 5, 10, 20, and 50, the average diagnosis rates for T1–T9 sensors are maintained at 94.40%, 95.28%, 94.79%, 92.52%, and 93.35%, demonstrating effective fault diagnosis capability across various conditions.

(3) Through medium-to-long-term prediction, faulty data recovery is successfully achieved in this study. The reconstructed data are guaranteed to remain within the fault threshold range, enabling short-term data restoration. System operation can be maintained under sensor fault conditions through this approach. This method has been established as providing a viable medium-to-long-term prediction solution for sensor network fault diagnosis in workshop environmental control systems. Future research should be directed toward exploring methods to increase prediction step lengths while ensuring diagnostic accuracy and reliability.

(4) Adapting this method to new applications does indeed incur a one-time initial model training cost, but its long-term operation and calibration costs are low. Therefore, we believe that this method has significant competitiveness in terms of cost-effectiveness. Its one-time upfront investment can bring long-term reliability and stability improvement to industrial systems, avoiding huge economic losses caused by sudden failures, thus having high practical application value and return rate.

Author Contributions: Writing—Original Draft, S.X.; Writing—review & editing, F.Y.; Investigation, X.Z.; Validation, M.T.; Resources, C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Natural Science Foundation of Hubei Province Grants No. 2023AFB400, the Hubei University of Technology Research Fund under Grants BSQD2020009 and International Science and Technology Joint Research Project of Hubei, China (Grant No. 2024EHA007).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Negi, V.P.S.; Ranganayakulu, C. Energy-efficient three-wheel bleedless electrical environmental control system for a passenger aircraft. *J. Appl. Therm. Sci. Eng. Appl.* **2025**, *17*, 021009. [CrossRef]
2. Jennions, I.; Ali, F.; Miguez, M.E.; Escobar, I.C. Simulation of an aircraft environmental control system. *J. Appl. Therm. Eng.* **2020**, *172*, 114925. [CrossRef]
3. Verma, P.; Kumar, V.; Daim, T.; Sharma, N.K. Design thinking framework toward management control system in environmental dynamism: An innovation perspective. *J. IEEE Trans. Eng. Manag.* **2024**, *71*, 4955–4970. [CrossRef]
4. Petera, P.; Wagner, J.; Paksiova, R. The Influence of environmental strategy, environmental reporting and environmental management control system on environmental and economic performance. *J. Energ.* **2021**, *14*, 4637. [CrossRef]
5. Cozza, A. Low frequency model-based identification of soft impedance faults in cables. *J. IEEE Trans. Instrum. Meas.* **2019**, *68*, 3524–3535. [CrossRef]
6. Ginzarly, R.; Moubayed, N.; Hoblos, G.; Kanj, H.; Alakkoumi, M.; Mawas, A. Assessing HMM and SVM for condition-based monitoring and fault detection in HEV electrical machines. *J. Energ.* **2025**, *18*, 3513. [CrossRef]
7. Yan, K.; Zhong, C.W.; Ji, Z.W.; Huang, J. Semi-supervised learning for early detection and diagnosis of various air handling unit faults. *J. Energy Build.* **2018**, *181*, 75–83. [CrossRef]
8. Orellana, C.; Ortiz, L.; Aguila, A.; Garcia, M.; Ruiz, M. Hybrid algorithm for fault detection in three-phase motors using principal component analysis. *J. E-Prime—Adv. Electr. Eng. Electron. Energy* **2025**, *11*, 100916. [CrossRef]
9. Mao, Q.J.; Fang, X.; Hu, Y.P.; Li, G. Chiller sensor fault detection based on empirical mode decomposition threshold denoising and principal component analysis. *J. Appl. Therm. Eng.* **2018**, *144*, 21–30. [CrossRef]
10. Zhong, X.P.; Zhang, L.; Ban, H. Deep reinforcement learning for class imbalance fault diagnosis of equipment in nuclear power plants. *J. Ann. Nucl. Energy* **2023**, *184*, 109685. [CrossRef]
11. Gins, G.; Kerkhof, D.V.P.; Vanlaer, J.; Van Impe, J.F. Improving classification-based diagnosis of batch processes through data selection and appropriate pretreatment. *J. Process Control* **2015**, *26*, 90–101. [CrossRef]
12. Wang, X.; Shen, C.Q.; Xia, M.; Wang, D.; Zhu, J.; Zhu, Z. Multi-scale deep intra-class transfer learning for bearing fault diagnosis. *J. Reliab. Eng. Syst. Saf.* **2020**, *202*, 107050. [CrossRef]
13. Li, G.N.; Xiong, C.L.; Gao, J.J.; Zhu, H.; Wang, C.; Xiao, J. Fault detection, diagnosis and calibration of heating, ventilation and air conditioning sensors by combining principal component analysis and improved bayesian inference. *J. Build. Eng.* **2024**, *82*, 108230. [CrossRef]

14. Zhou, H.Y.; Zhang, S.H.; Peng, J.Q.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115. [CrossRef]
15. Li, F.Y.; Wan, Z.B.; Koch, T.; Zan, G.; Li, M.; Zheng, Z.; Liang, B. Improving the accuracy of multi-step prediction of building energy consumption based on EEMD-PSO-Informer and long-time series. *J. Comput. Electr. Eng.* **2023**, *110*, 108845. [CrossRef]
16. Guo, L.; Li, R.Z.; Jiang, B. A data-driven long time-series electrical line trip fault prediction method using an improved stacked-informer network. *J. Sens.* **2021**, *21*, 4466. [CrossRef] [PubMed]
17. Klimek, J.; Jakub, K.; Kraśkiewicz, W.; Topolewski, M. Long-term series forecasting with query selector-efficient model of sparse attention. *J. Softw. Impacts* **2022**, *11*, 100187. [CrossRef]
18. Xu, S.Q.; Yu, H.Y.; Wang, H.; Chai, H.; Ma, M.; Chen, H.; Zheng, W.X. Simultaneous diagnosis of open-switch and current sensor faults of inverters in IM drives through reduced-order interval observer. *J. IEEE Trans. Ind. Electron.* **2025**, *72*, 6485–6496. [CrossRef]
19. Shaheen, K.; Chawla, A.; Rossi, P.S. Trust-Enhanced Distributed Kalman Filtering for Sensor Fault Diagnosis in Sensor Networks. *J. IEEE Trans. Signal Inf. Process. Over Netw.* **2025**, *11*, 1178–1187. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.