

## Article

# An Effective Fault Detection Method for HVAC Systems Using the LSTM-SVDD Algorithm

Hui Zhu <sup>1</sup>, Wen Yang <sup>2</sup> , Shihong Li <sup>3</sup> and Aiping Pang <sup>1,\*</sup><sup>1</sup> School of Electrical Engineering, Guizhou University, Guiyang 550025, China; gs.huizhu19@gzu.edu.cn<sup>2</sup> Key Laboratory of Space Launching Site Reliability Technology, Xichang Satellite Launch Center, Xichang 615000, China; whutyw@126.com<sup>3</sup> School of Automation, Air Force Aviation University, Changchun 130012, China; b20210201@st.nuc.edu.cn

\* Correspondence: appang@gzu.edu.cn

**Abstract:** Fault detection in heating, ventilation and air-conditioning (HVAC) systems can effectively prevent equipment damage and system energy loss, and enhance the stability and reliability of system operation. However, existing fault detection strategies have not realized high effectiveness, mainly due to the time-delay characteristics of HVAC system faults and the lack of system-fault operation data. Therefore, aiming at the time delay of system faults and the lack of actual system-fault operation data, this paper proposes a fault detection method that combines a system simulation model and an intelligent detection algorithm. The method first uses the Modelica modeling language to build a scalable simulation model of the system to obtain fault data that are not easily accessible in practice. The long short-term memory-support vector data description (LSTM-SVDD) algorithm is then applied to detect faults in real time by dynamically adjusting the fault residuals according to the absolute difference between the predicted and actual values. The experimental results show that the LSTM-SVDD method improves the average detection accuracy by 9.675% and 9.85% over the classical LSTM network and the extreme gradient boosting (XGBoost) method, respectively, under different fault levels.



**Citation:** Zhu, H.; Yang, W.; Li, S.; Pang, A. An Effective Fault Detection Method for HVAC Systems Using the LSTM-SVDD Algorithm. *Buildings* **2022**, *12*, 246. <https://doi.org/10.3390/buildings12020246>

Academic Editors: Shi-Jie Cao and Fabrizio Ascione

Received: 31 December 2021

Accepted: 18 February 2022

Published: 20 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/>).

**Keywords:** HVAC systems; fault detection; long short-term memory; support vector data description; Modelica modeling

## 1. Introduction

The modern HVAC systems are complex non-linear systems with large inertia and large hysteresis. They are now often applied in significant locations, such as data centers and communication base stations [1]. Faults in such large-scale complex systems are usually difficult to detect in the initial stages, due to the lack of effective detection techniques, which can lead to excessive energy losses and even significant property damage and social impact. There is, therefore, an urgent need to further improve the reliability and safety of HVAC systems in order to reduce the adverse effects caused by faults. To address this issue, real-time fault detection of HVAC systems is required to identify potential system faults so that timely maintenance measures can be planned to ensure the normal operation of the system. In recent years, a series of fault-detection research has been developed for HVAC systems. There are two primary approaches: model-based methods and data-driven methods. The model-based method first establishes an indicator model that is more sensitive to a particular fault. The deviation between the expected and actual values of the detected physical object is then used to detect system faults. Zabala et al. [2] developed a physical model to diagnose faults in surface air cooler of HVAC systems, and the results showed that the method was able to accurately detect faults. Although model-based methods can perform well to detect faults, these methods are dependent on the accuracy of the physical model. However, it is very difficult to establish an accurate

model of the HVAC system by using traditional modeling techniques [3]. Therefore, data-driven methods have become a research hotspot because they can directly identify faults from the operational data of HVAC systems without modeling the physical model [4]. The historical HVAC system operation data are used to train the learning model. Fault alarms are raised as soon as the new system operation data are inconsistent with the output data of the learning model. Du et al. [5] detected faults of air handling units by using a subtractive cluster analysis algorithm as an auxiliary neural network integrated with a standard neural network. The detection results, with false-alarm and missed-alarm rates as evaluation indicators, show that the detection accuracy of the method is significantly better than that of the detection method that only uses one neural network. Magoules et al. [6] introduced a fault detection method which is based on a recursive deterministic perceptron neural network algorithm for HVAC systems. Fan faults, coil faults, pump faults and chiller faults were detected by using this method. Yan et al. [7] argue that sensor faults of HVAC systems are not easily detected in the early stages. In order to accurately detect early faults of the system in real time, a method based on a combination of kernel principal component analysis and bi-directional two-layer long-short time memory networks was proposed. The method effectively improves the efficiency of detecting early faults of the system by using the sensitivity of the long- and short-time memory network to the time series of fault data. Although data-driven techniques perform well for fault detection of HVAC systems, the performance of these techniques relies heavily on a balanced dataset containing a large number of fault and normal data points [8]. However, gathering a sufficient number of fault training samples is usually a challenging task in realistic situations where HVAC systems are not allowed to operate under fault conditions. These fault samples are necessary for the construction of a balanced training dataset. An unbalanced training dataset degrades and even invalidates the performance of most data-driven fault detection methods [9,10]. To address the problem of inadequate fault samples, Yan et al. [11] proposed that the problem of unbalanced fault detection data in air handling units could be addressed by using the generative adversarial network (GAN). The method rebalances the training dataset by augmenting a few fault samples with GAN as a preprocessing step. Zhang et al. [12] simulated operational faults of HVAC systems in the EnergyPlus simulation platform. These faults include sensor deviations, fouled coils and broken air filters. Kim et al. [13] developed and validated a 25-fault model of a small commercial building, using OpenStudio software. All of their research has generated simulation datasets for fault detection based on the developed fault models [14]. However, the Modelica modeling language can perform the modeling and fault simulation of such systems faster and more accurately than other simulation tools [15]. Hefni et al. [16] used the Modelica language to build a power plant model that provides convincing simulation results. The developed model can cover the whole system engineering lifecycle, from pre-liminary design to commissioning, operation and maintenance. Therefore, this paper proposes a modeling approach based on the Modelica language. The HVAC system faults are simulated on the OpenModelica simulation platform to generate fault samples.

It was discovered that the operational data of HVAC systems tend to have strong time-series characteristics. Iffat [17] argues that different deep learning techniques can be used to mine these data in order to achieve automatic fault detection in the system. Kayacan [18] studied the prediction of time-series data by using deep learning methods. Meanwhile, LSTM networks possess a strong extraction capability for time-series features that is well suited for the prediction of time-series monitoring variables of HVAC systems. Li et al. [19] showed that the LSTM model gave a superior performance to the back-propagation neural network and multilayer perception to predict the temperature of HVAC systems. Zhang et al. [20] significantly improved the fault-prediction accuracy of line tripping within power systems by using the stronger learning ability of LSTM networks for time series. Meanwhile, LSTM-network-based approaches may obtain better results in the prediction of HVAC system variables compared to other neural networks. However, the results of real-time fault detection in HVAC systems are not only related to

the prediction performance of the neural network, but are often also closely related to the fault threshold setting [21]. Traditional methods of setting fixed thresholds often result in missed opportunities to detect faults, delayed detection of faults or false alarms [22–24]. If the threshold is set too high, faults may be missed, or if the threshold is too low, false alarms may occur. Therefore, it is important to set a reasonable threshold for prediction errors in LSTM networks. The prediction error of the LSTM network is considered as the fault residual. When the system is operating normally, the fault residuals tend to be small, and when the system is faulty, the fault residuals tend to be larger than normal, and this can be viewed as a classification problem. The SVDD algorithm is an excellent single classification algorithm. It can change itself for sensitivity to the anomaly by changing the spherical boundary to a more flexible boundary [25]. As a result, the radius of the SVDD algorithm hypersphere can be used well as an automatic threshold for fault detection. Li et al. [26] employed the SVDD algorithm to detect the fault of sensor faults in chillers with very accurate detection results.

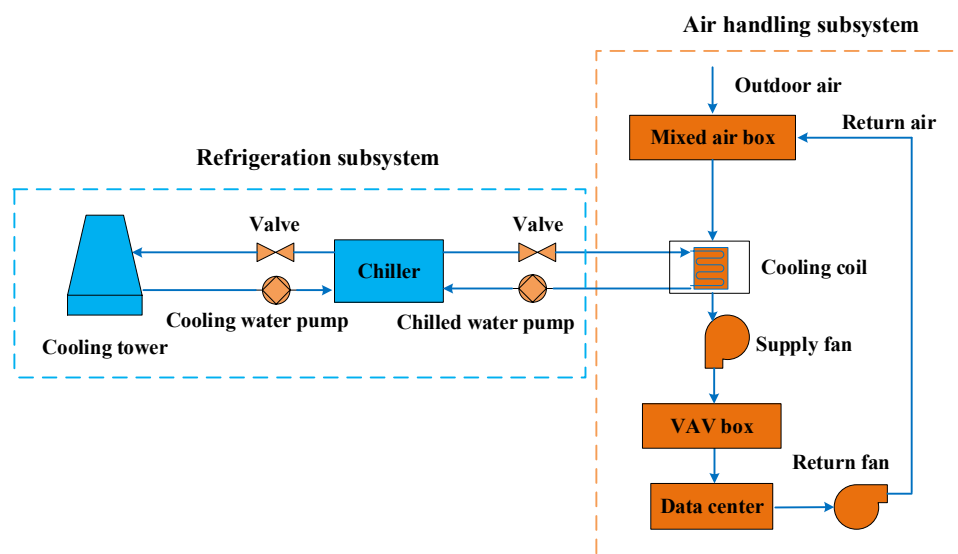
Therefore, we propose the LSTM-SVDD method to achieve real-time and accurate fault detection in HVAC systems after studying the fault characteristics of HVAC systems through simulation systems. The method first utilizes the strong learning capability of the double-layer LSTM network for the monitored variables of the HVAC system to predict the workshop temperature of the system in real time. Then, a residual value is generated by comparing the output of the LSTM network with the actual value from the workplace temperature sensor. Finally, this residual value is used as input to the SVDD method with the radius of the SVDD hypersphere used as the fault detection threshold to detect a fault in the system.

## 2. HVAC System Modeling and Fault Simulation

This section describes the development of a simulation system model based on the structure and operating principles of the HVAC system. In addition, faults are simulated by manually injecting faults.

### 2.1. HVAC System Modeling

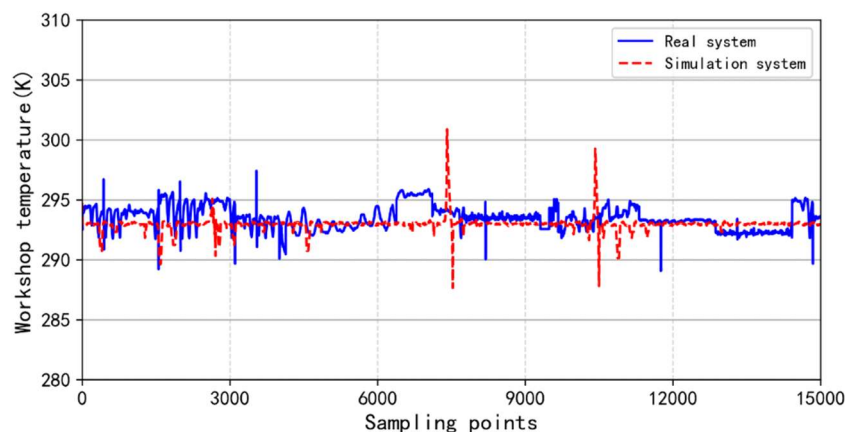
The researched HVAC system is a water-cooled HVAC system at a data center, and its system structure is similar to the subject of the literature [27]. It is shown in Figure 1.



**Figure 1.** HVAC system structure of the data center.

This HVAC system is divided into a water circuit subsystem and an air-circuit subsystem based on object-oriented modeling. The water circuit subsystem consists of a chiller model and a cooling tower model, while the air circuit subsystem consists of a fan coil

model and a data-center-workshop model. This paper establishes HVAC system simulation models on the OpenModelica platform [28] based on the Modelica standard library and the Buildings library. The parameters of the simulated system were then set and calibrated to maintain a room temperature of around 293 K based on sensor measurements of the actual HVAC system. The difference between the simulated system and the actual HVAC system is verified by using the data-center-workshop temperature variation as an evaluation indicator. At the same time, temperature data are collected from the actual data center workshop that is cooled by the HVAC system. A comparison of the collected data from the actual system with the data from the simulation model is shown in Figure 2.



**Figure 2.** Data-center-workshop temperature comparison between actual and simulated systems.

As can be seen from Figure 2, the data-center-workshop temperatures of both the simulation system and the actual system are maintained around the set point, and their average temperature values are only different with 0.58 K, so the simulation system established is relatively accurate.

## 2.2. System Monitoring Variables and Fault Simulation

HVAC system failures mean abnormal operation with reduced system performance that includes more energy being used than the normal operation or failure to maintain data-center temperature according to thermostat set points [29]. In realistic scenarios, many sensors are installed in HVAC systems to monitor the operational behavior of the system. However, some system-critical variables cannot be monitored due to the impossibility of installing sensors or the high cost of such sensors. Thus, considering the distribution of HVAC system sensors and system-important variables in the actual system, the simulation system is used to more comprehensively monitor the system operating status by constructing virtual sensors. In this research, there are eleven system-monitoring variables selected. All of these different variables, as shown in Table 1, are related to the operational status of the HVAC system.

Since the HVAC system is mainly used to regulate the temperature of the data center workshop, the change of temperature is closely related to the system failure. However, owing to the system fault propagation delay, the equipment in the data center is not working properly by the time the fault causes the temperature in the data center to rise. Therefore, the temperature of the data center workshop is chosen as the prediction parameter of the system in order to detect the failure in time. The correlation of the system monitoring variables should be considered for the accurate prediction of the temperature in the data center workshop. The correlation of each monitoring variable in Table 1 is shown in Figure 3.



**Table 1.** System-monitoring variables.

Parameter	Symbol	Unit
Chiller water supply temperature	CWST	K
Chiller water return temperature	CWRT	K
Cooling tower water supply temperature	CTWST	K
Cooling tower water return temperature	CTWRT	K
Mass flow rate of the supply air from the coil	MFRSAC	Kg/s
Coil supply air pressure	CSAP	Pa
Temperature of the mixed air	TMA	K
Coefficient of performance of the chiller	COP	\
Outlet temperature of the air from the coil	OTAC	K
Temperature of the recirculated air	TRA	K
Workshop temperature	WT	K

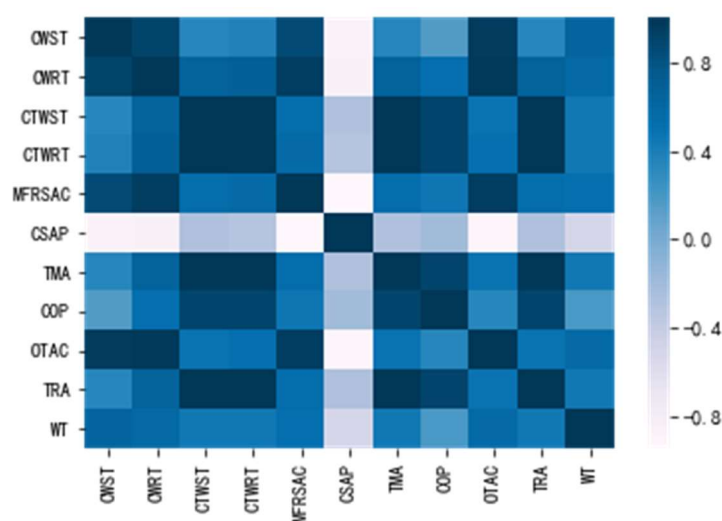
**Figure 3.** Heat map of the correlation analysis for each variable.

Figure 3 shows that the five most relevant monitoring variables relating to data-center temperature are chiller water-supply temperature, chiller water return temperature, mass flow rate of the supply air from the coil, temperature of the mixed air and outlet temperature of the air from the coil. Therefore, six monitoring variables containing the data-center-workshop temperature were selected as inputs to the fault detection model. The system fault detection can be achieved by using the proposed fault detection model to extract information from these six parameters. The six monitored variables are normalized as shown in Figure 4.

In this study, the failure of the air-supply ducts for the end of the HVAC system led to a deviation of the temperature in the data center workshop beyond its normal set point, and, thus, a ruptured air supply duct was selected for research and description. The fault data were obtained by manually injecting the fault into the simulation system. After the system had stabilized, a fault was set in the air-supply duct at a specific point in time. The fault was simulated by attaching an additional small container to the air-supply duct of the simulation model. The process simulates a rupture in the air supply duct leading to a leakage failure of the supply air volume. Different fault sizes can be simulated by connecting multiple containers, as well. The temperature of the data center workshop in the training set is shown in Figure 5. For the test set, the data-center-workshop temperatures under normal state and the faulty state are shown in Figure 6. The fault starts at the 3000th sampling point and continues until the end of the test set.

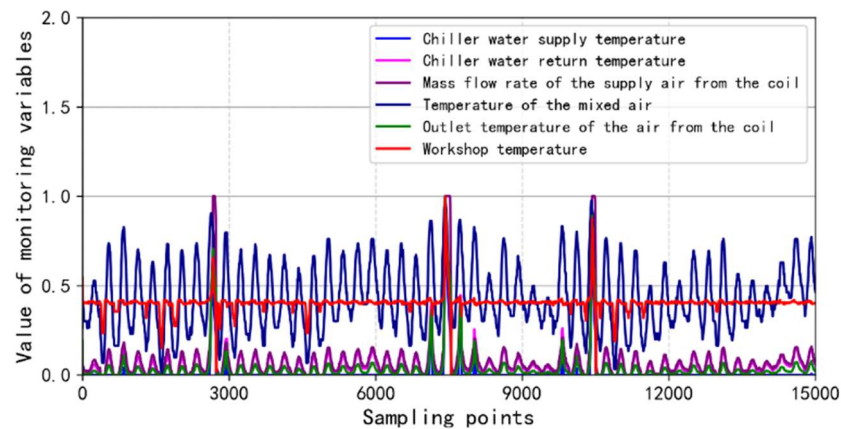


Figure 4. Six standardized monitoring variables.

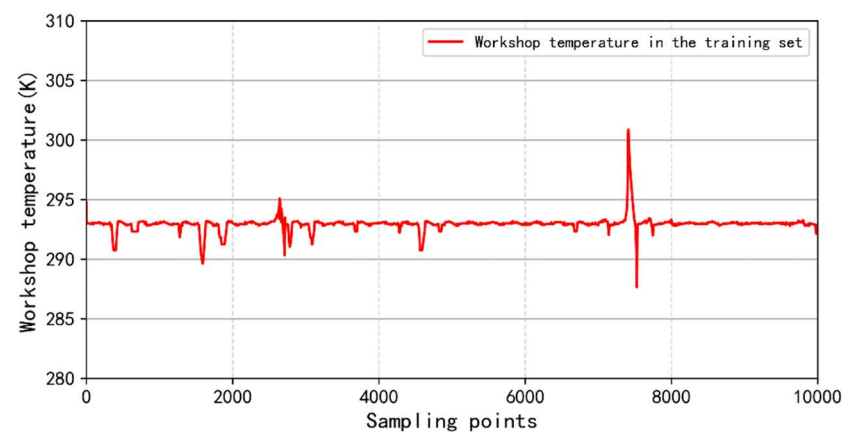


Figure 5. Data-center-workshop temperature in the training set.

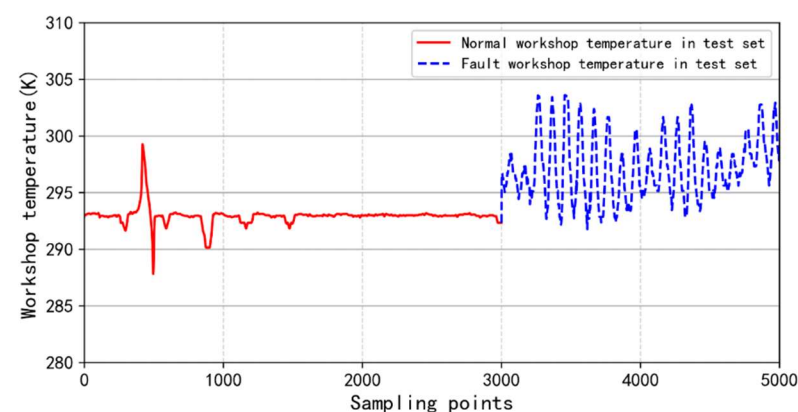


Figure 6. Data-center-workshop temperature in the test set with fault.

### 3. Architectures and Algorithms

In this section, the principles and applications of the proposed LSTM-SVDD fault detection method are described in detail. The LSTM network is an improved network based on RNN. It solves the vanishing gradient problem of RNN by introducing a memory unit. In addition, the support-vector data description performs equally well for non-Gaussian distributed data and non-linear data.

#### 3.1. Long Short-Term Memory

Long- and short-term memory networks are a variant of RNN that aim to solve the gradient disappearance problem [30]. To solve the problem of vanishing gradients, the

LSTM has three gates per cell to protect and control the cell state. Figure 7 shows the structure of the LSTM memory cell.

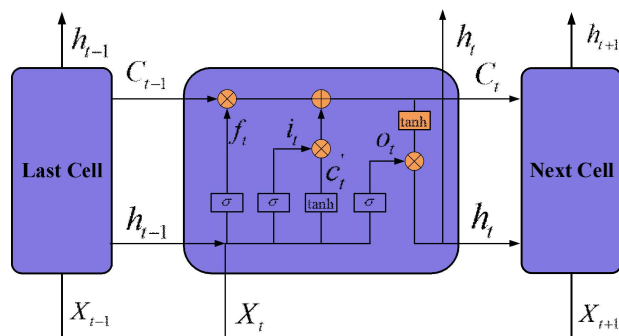


Figure 7. Structure of the LSTM memory cell.

A multiple-input LSTM model was designed to achieve accurate prediction of the data-center-workshop temperature of the HVAC system. Meanwhile, we found that the double-layer LSTM network performed the best prediction results when different structures of LSTM models were used to predict the data-center-workshop temperature experimentally.

The double-layer LSTM regression model, based on the structure of the single-layer LSTM, is shown in Figure 8. The model has one input layer, two LSTM layers and one output layer. Where  $X_L$  is the input data for the double layer LSTM regression model and  $Y_L$  is the predicted output of the model.

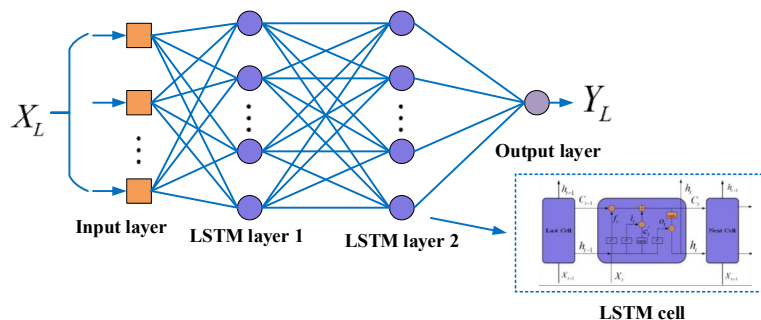


Figure 8. Double-layer LSTM regression model.

This regression prediction model enables data of the next moment to be inferred from the historical operating data of the HVAC system. Therefore, we could next develop an LSTM regression model to obtain the predicted values of the monitored parameters. In this work, six monitoring variables related to HVAC systems were used for regression prediction. Suppose the sample length of each monitored variable is  $m$ ; then the operational data of the HVAC system,  $X$ , can be represented by the matrix as follows:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \tag{1}$$

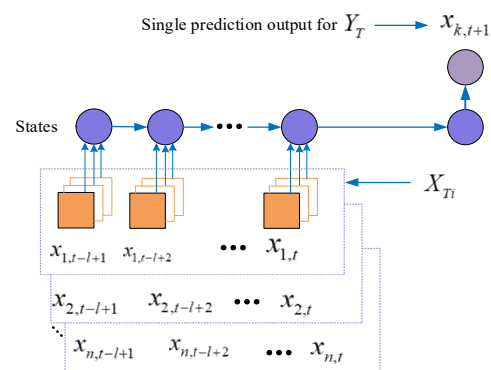
where  $n = 6$ , which is the number of the HVAC system operating parameters, and  $m$  is the number of points sampled for each monitored variable.

The workplace temperature was selected as the predicted monitoring parameter to illustrate. The operational data,  $X$ , of the HVAC system need to be reconstructed as follows in order to construct a sliding-window model of regression prediction.

$$X_{Ti} = \begin{bmatrix} x_{1,t-l+1} & x_{1,t-l+2} & \cdots & x_{1,t} \\ x_{2,t-l+1} & x_{2,t-l+2} & \cdots & x_{2,t} \\ \vdots & & & \vdots \\ x_{n,t-l+1} & x_{n,t-l+2} & \cdots & x_{n,t} \end{bmatrix} \quad (2)$$

$$Y_T = [x_{k,t+1}] \quad (3)$$

where  $X_{Ti}$  is the input of the regression model at time,  $t$ ;  $Y_T$  is the target output;  $K$  is used to indicate which parameter is being monitored;  $l$  is the length of a sliding window, which determines the length of the data sequence associated with the next moment;  $t$  increments from 1 to  $m$ ; and  $i$  increments from 1 to  $m - l$ . For moment  $t$ , the detailed sliding process is shown in Figure 9.



**Figure 9.** Sliding-window description at moment  $t$ .

After reconstruction of the operational data of the system, the input samples and corresponding outputs of the LSTM regression prediction model were as follows.

$$\begin{cases} X_L = [X_{T,1} & X_{T,2} & \cdots & X_{T,m-l}] \\ Y_L = [x_{k,l+1} & x_{k,l+2} & \cdots & x_{k,m}] \end{cases} \quad (4)$$

The regression model needs to learn the mapping function,  $f_{LSTM}(\cdot)$ , which is defined as follows:

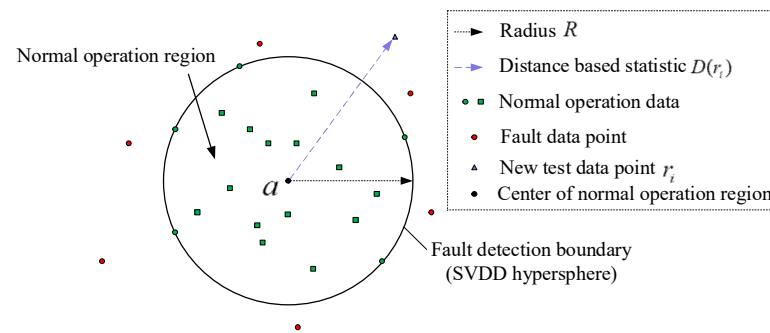
$$Y'_L = f_{LSTM}(X_L) \quad (5)$$

where  $Y'_L$  is the predicted value of  $Y_L$ . Meanwhile, the absolute value of the prediction error is defined as the residual value,  $r$ .

$$r = |Y'_L - Y_L| \quad (6)$$

### 3.2. Support-Vector-Data Description

There is a significant change between the residual values of HVAC systems in normal and fault conditions. Hence, the SVDD algorithm is used to identify the characteristics of the residual value,  $r$ , to detect whether a fault has occurred in the system. SVDD is an excellent single class classification algorithm for process monitoring and fault detection. Its basic idea is to obtain a spherical decision boundary whose minimum volume contains most normal data points, as illustrated in Figure 10.



**Figure 10.** Illustration of the SVDD hypersphere and the D statistic for fault detection.

For a given data matrix,  $D$ , containing  $N$  input vectors  $d_i$ , the desired minimum-volume hypersphere can be described by the center,  $a$ , and the radius,  $R$ , and meets the following objective function. Hence, its hypersphere radius can often be used as a great fault detection threshold.

$$\min_{a, R, \zeta} R^2 + C \sum_{i=1}^N \zeta_i \quad (7)$$

$$s.t. \quad \|\varphi(d_i - a)\|^2 \leq R^2 + \zeta_i, \zeta_i \geq 0, \forall i = 1, 2, \dots, N \quad (8)$$

where  $\zeta_i$  is a relaxation factor,  $\varphi$  is a mapping function and  $C$  is a penalty factor.

Equations (7) and (8) can be regarded as an optimization problem, so the above optimization problem can be transformed as follows:

$$\max_{\alpha_i} \sum_{i=1}^N \alpha_i G(d_i, d_i) - \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j G(d_i, d_j) \quad (9)$$

$$s.t. \quad 0 \leq \alpha_i \leq C, \sum_{i=1}^N \alpha_i = 1 \quad (10)$$

Here,  $G$  is a kernel function which is chosen as the Gaussian radial kernel function shown in Equation (11), with  $g$  as the only variable parameter; and  $\alpha_i$  is the Lagrange multiplier of each sample. If  $0 \leq \alpha_i \leq C$ , the sample  $d_i$  is denoted as the support vector. Defining the support-vector dataset  $S$ , the center ( $a$ ) and radius ( $R$ ) of the hypersphere can be calculated by Equations (12) and (13), respectively.

$$G(d_i, d_j) = \varphi(d_i) \cdot \varphi(d_j) = \exp(-g\|d_i - d_j\|^2) \quad (11)$$

$$a = \sum_{i=1}^N \alpha_i \varphi(d_i) \quad (12)$$

$$R = \sqrt{G(d_k, d_k) - 2 \sum_{i=1}^N \alpha_i G(d_k, d_i) + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j G(d_i, d_j)} \quad (13)$$

where  $d_k \in S$ .

To detect faults in HVAC systems, the residuals from the normal operation of the system are used to train the SVDD model. For any new residual test sample  $r$ , the trained SVDD model can calculate the distance  $D(r)$  from its hypersphere sphere center  $a$  to the new test sample  $R$ .

$$D(r) = \sqrt{G(r, r) - 2 \sum_{i=1}^N \alpha_i G(r, d_i) + \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j G(d_i, d_j)} \quad (14)$$



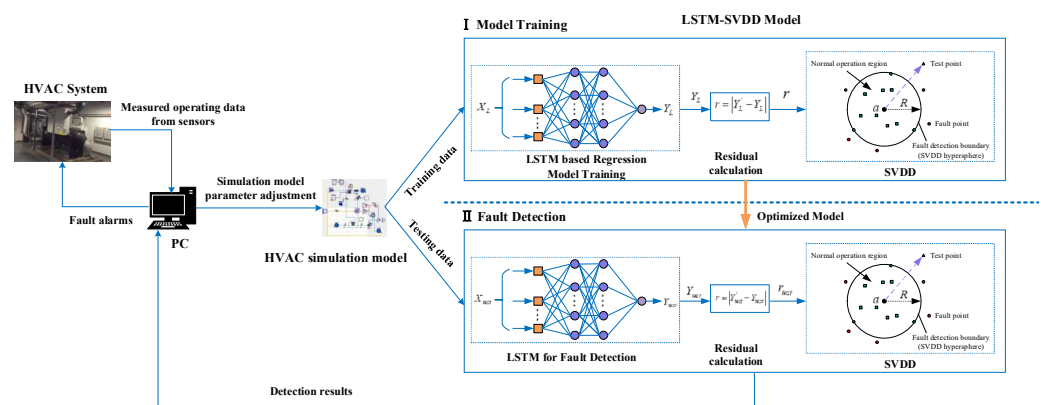
If  $D(r) > R$ , the new test-sample  $r$  is regarded as a faulty sample, and, thus, the system is detected to have faults.

#### 4. Fault Detection Framework Based on LSTM-SVDD

In this section, we describe in detail the general framework of the LSTM-SVDD-based fault detection method and then give the parameters of the fault detection model. Finally, we experimentally analyze the performance of the proposed method in this paper. The program for method validation was implemented in Python 3.7 environment, and the LSTM network and SVDD algorithm programs were written in Jupyter Notebook in combination with Python language to build the fault detection model of HVAC system. The experiment was run on a Windows computer equipped with an Intel® Core™ i5-9300U CPU clocked at 2.40 GHz and 8.0 GB of physical memory.

##### 4.1. Fault Detection Framework

The fault detection framework consists of two main components: simulation system establishment based on actual sensor-measurement data and system fault detection based on LSTM-SVDD. The general illustration of the fault detection model is shown in Figure 11. First, the server of the HVAC system in the realistic situation collects data from the sensors of the system during normal operation. The parameters of the simulated system are calibrated and set based on the data collected, so that the performance of the simulated system approximates the actual system. Subsequently, we carry out fault injection on the simulation system to obtain fault-operation data of the system. Finally, the LSTM-SVDD model extracts and identifies features from the data of the simulated system to detect system faults and provide fault warning in real time.



**Figure 11.** A fault detection framework based on LSTM-SVDD.

Faults in HVAC systems depend not only on changes in operating data at the current moment, but also on the state at previous moments. The LSTM-SVDD method makes full use of the characteristics of the time-series data to avoid false positives and improve detection performance at a certain level. As shown in Figure 12, the LSTM-SVDD fault detection steps are as follows.

Step 1: Data collection. The dataset was collected from the HVAC system of the simulation platform and consisted of six characteristic variables. The data-center-workshop temperature was the predicted output variable.

Step 2: We normalized the data to eliminate differences between data features, while the reconstructed data were converted to a sequence with a sliding window.

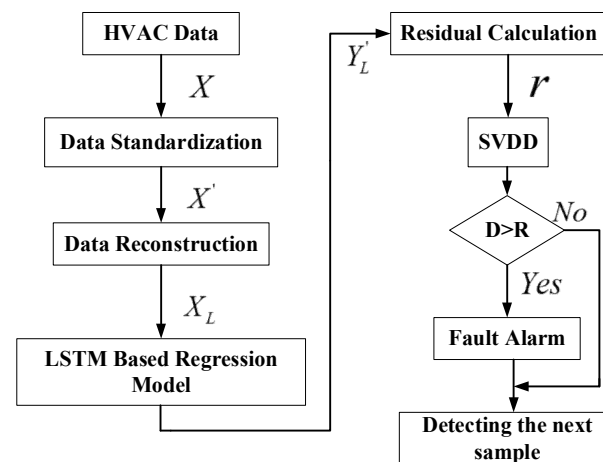


Figure 12. LSTM-SVDD fault-detection data-stream diagram.

Step 3: The processed data were taken as the input data of the double-layer LSTM network, and thus a double-layer LSTM model was developed to achieve the predicted output. This LSTM network was trained with the goal of minimizing the sum of the total loss functions. The loss function is the mean square error given in Equation (15). The final loss function is about 0.02 and remains stable after training, thus indicating that the double-layer LSTM meets the requirements with training.

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (Y'_L - Y_L)^2} \quad (15)$$

where  $m$  is the number of predicted points,  $Y_L$  is the actual output and  $Y'_L$  is the model predicted output.

Step 4: The residual value,  $r$ , is taken as input data for the SVDD algorithm. The SVDD model uses its hypersphere radius,  $R$ , as the fault threshold of fault detection to identify whether a fault occurs in the system.

#### 4.2. Parameters Setting for LSTM-SVDD Model

In this paper, two LSTM layers of the LSTM-SVDD based fault detection model are designed to increase the depth of the model and obtain better prediction results. There are 20 hidden nodes per LSTM layer. In addition, the LSTM sliding window length is set to 20 to obtain better time series prediction performance. Table 2 summarizes the parameters used in the proposed fault detection model.

Table 2. LSTM-SVDD model parameters.

Parameter	Symbol	Value
Sliding window	l	20
Input dimension	u	6
First layer nodes	S1	20
Second layer nodes	S2	20
Output dimension	v	1
Learning rate	b	0.01
Penalty weighting factor	C	0.8
Gaussian kernel width parameter	g	0.04

## 5. Fault Detection Results and Discussion

This section demonstrates the effectiveness of the proposed method based on the given fault detection evaluation indicators and comparative experimental results.

### 5.1. Fault Detection Evaluation Indexes

We need to introduce fault-detection evaluation indicators to estimate the performance of the fault detection model. There are four possible results of the fault detection, commonly referred to as true positive ( $TP$ ), false positive ( $FP$ ), false negative ( $FN$ ) and true negative ( $TN$ ) [31]. Then the detection accuracy of the model was calculated by using Equation (16).

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (16)$$

Accuracy ( $Acc$ ) indicates the proportion of all prediction samples that are correctly predicted. Accuracy is commonly used as an evaluation metric. Generally speaking, higher accuracy means a better performance for the detection model. However, in the case of unbalanced positive and negative samples and unbalanced data distribution, accuracy as an evaluation metric is not sensitive to misclassification of fewer samples. Precision and recall rate can compensate for this deficiency. Meanwhile, the F1-measure value takes into account the precision rate and recall rate comprehensively. Therefore, the  $Acc$  and F1-measure were selected as evaluation metrics in order to evaluate the model performance overall.

The precision ( $P$ ) is calculated as follows:

$$P = \frac{TP}{TP + FP} \quad (17)$$

The recall rate ( $R$ ) is calculated as follows:

$$R = \frac{TP}{TP + FN} \quad (18)$$

The F1-measure is the harmonic mean of precision and recall rate.

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (19)$$

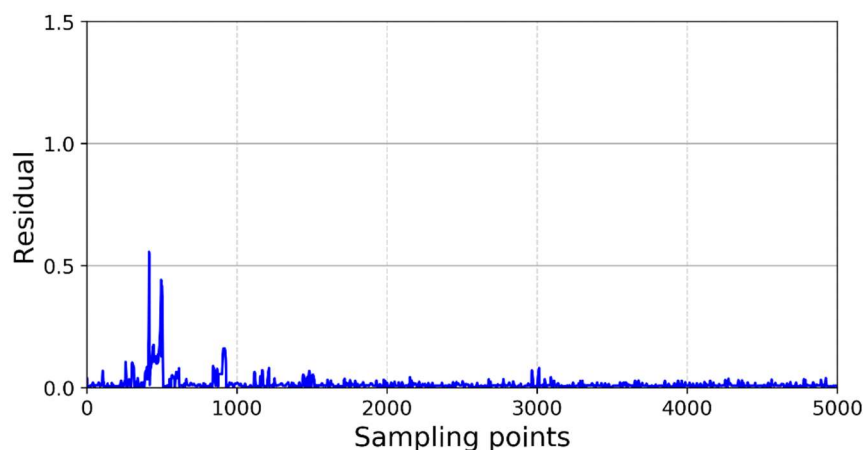
### 5.2. Fault Detection Results and Analysis Based on LSTM-SVDD

To comprehensively analyze the performance of the proposed fault detection model, we validated it from two perspectives: prediction performance and fault detection performance for different levels of faults. The predictive performance of the proposed LSTM network is not perfect for the model predicted values, even in the fault-free state, due to modeling errors. Therefore, these predicted values must be further processed before identifying the occurrence of system faults. The most efficient processing method is to impose a fixed threshold on the difference between the predicted and actual values. If the difference is more than a threshold value, a fault is identified. We evaluate its predictive performance for data-center-workshop temperature with the residual value,  $r$ .

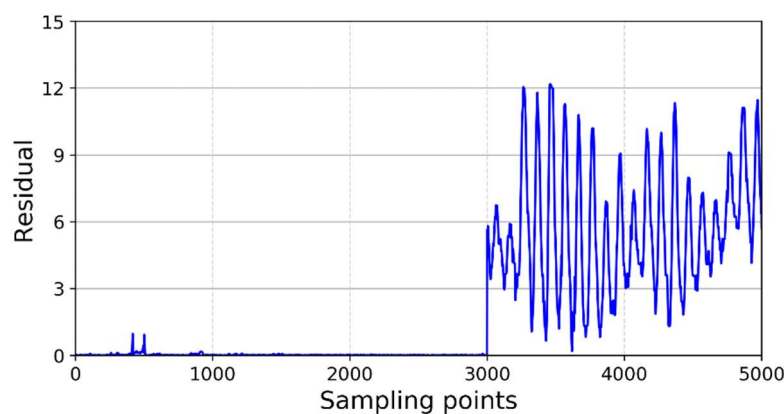
$$r = |Y'_L - Y_L| \quad (20)$$

The residual value,  $r$ , of this LSTM network during normal operation of the system is shown in Figure 13.

Assuming that the system suffers a rupture of the supply air duct at the 3000th sampling point, the residual value,  $r$ , will change as shown in Figure 14. The fault detection process is generally divided into two phases: the first phase is the fault-free phase consisting of the first 3000 sample data, and the second phase is the fault phase consisting of the last 2000 sample data. In the fault-free phase, the smaller the residual value, the better the fault detection performance. Conversely, in the fault phase, the higher the residual value the better the fault detection performance. The predictive accuracy of the different methods directly influences the subsequent fault detection results.



**Figure 13.** Data-center-workshop temperature-prediction error curve under normal operating conditions.



**Figure 14.** Data-center-workshop temperature prediction error curve with 1% fault deviation.

As can be seen from Figure 14, the residuals are small for the first 3000 samples when the system is operating normally, whereas the magnitude of the residuals grows and fluctuates considerably when the system fails.

The variation characteristics of the residual values can often be taken as a fault detection baseline during the normal and fault operation phases of the system. In the fault detection phase, all residual samples from the normal operation phase of the system are considered as positive samples and all residual samples from the fault phase operation are considered as negative samples. The SVDD algorithm is used to automatically detect the residual samples in both phases. The results of its detection are shown in Figure 15.

As can be seen from the Figure 15, the proposed LSTM-SVDD-based fault detection method has residuals all below the fault threshold in the no-fault phase, which means that there are no false alarms in this phase; meanwhile, most residuals are above the fault threshold in the fault phase with only a few below the fault threshold causing few missed alarms. This indicates that LSTM-SVDD works well for this fault detection.

The amplitude of the fault can also significantly affect the fault detection results. To validate the effectiveness of the proposed method for fault magnitude detection, the level of supply duct leakage was varied on the HVAC system simulation model to introduce different levels of faults that would cause the data-center-workshop temperature to deviate from the normal range. The fault level is based on the percentage of deviation from the average temperature of the data center workshop in normal operation. The ACC and F1-measure curves of the LSTM-SVDD method are shown in Figure 16 under the conditions of five fault levels of 2%, 1.5%, 1%, 0.5% and 0 deviation rate, respectively.

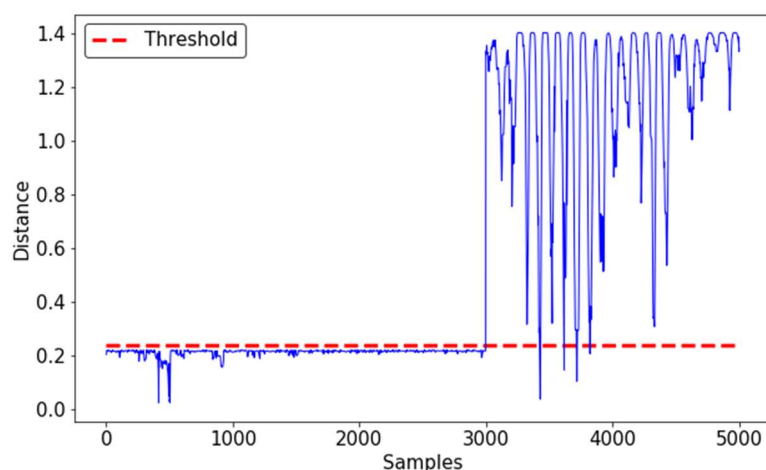


Figure 15. LSTM-SVDD detection results with 1% fault deviation.

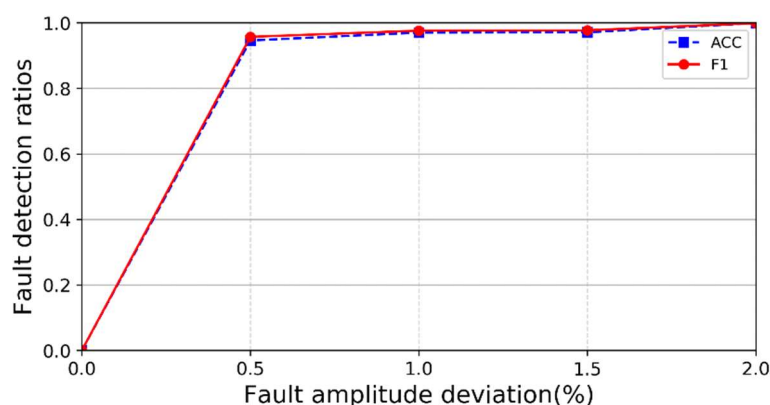


Figure 16. Detection efficiency at different fault levels.

In Figure 16, the LSTM-SVDD method still maintains a high detection efficiency for small deviation faults. Its detection efficiency reaches 0.946 even at a fault deviation rate of 0.5%, while at a fault deviation of 2%, its detection efficiency reaches 0.999. As well, there is a higher degree of accuracy in fault detection as the fault gets larger. It is easier to identify the data as anomalous, because the larger magnitude of the fault deviation means that the sample deviates from the center of the SVDD minimum hypersphere with greater distance.

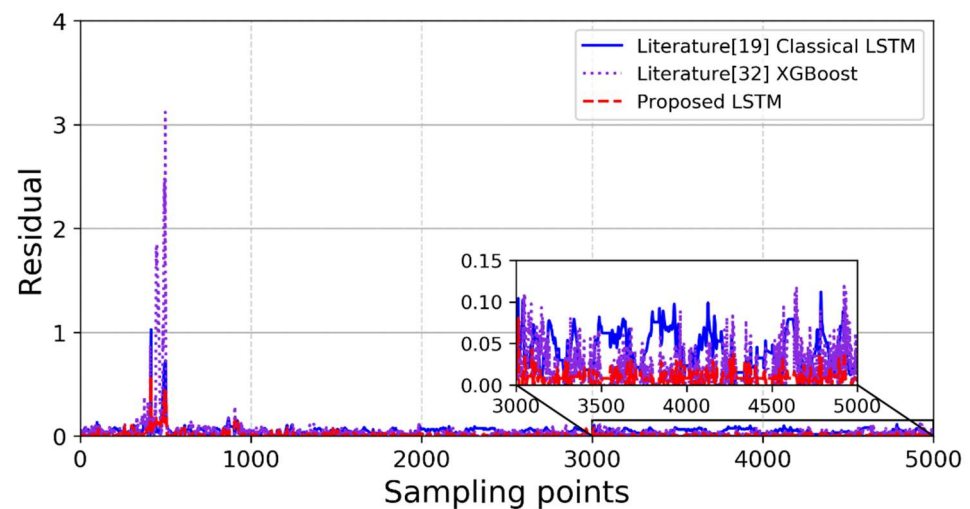
### 5.3. Methods Comparison Analysis

To verify the performance of the proposed method in this paper, regression prediction performance and fault detection efficiency were experimented upon with other methods. For prediction performance, the absolute difference,  $r$ , between the predicted and actual values on the test set was chosen as the prediction performance evaluation metric under the same time-series data that are divided between the same test and training sets. The prediction errors of the LSTM network proposed in the literature [19], the XGBoost method proposed in the literature [32] and the proposed LSTM network in this paper toward the data-center-workshop temperature are shown in Figure 17.

As can be seen from Figure 17, the prediction error of the proposed LSTM network in this paper is smaller than the prediction error of the LSTM network proposed in the literature [19] and the XGBoost method proposed in the literature [32]. Therefore, the proposed LSTM network in this paper performs the best in prediction performance, while the XGBoost method has the worst prediction performance when the time-series data fluctuates greatly. The reason is that LSTM networks tended to have an advantage over other neural networks for predicting time-series data. Since the classical LSTM network



uses only a single LSTM layer, the proposed LSTM network prediction model has a higher accuracy under the same number of training times.



**Figure 17.** Prediction results of the data-center-workshop temperature with the literature [19] LSTM, the literature [32] XGBoost and the proposed LSTM.

For fault detection efficiency, the literature [19] proposed LSTM network and the literature [32] proposed XGBoost method combined with a fixed threshold given in the literature [32] to detect faults. The final detection efficiency of the three methods under different fault levels is shown in Table 3. It can be found that the LSTM-SVDD detection model has a higher detection accuracy and F1 score than the other two methods, so it has the best fault detection performance.

**Table 3.** Detection results of different methods.

Fault Level	Methods	Prediction Errors	Acc	F1
L1 (0.5%)	Literature [30] Classical LSTM	0.0421	0.813	0.865
	Literature [31] XGBoost	0.0487	0.811	0.863
	<b>Proposed LSTM-SVDD</b>	<b>0.0165</b>	<b>0.946</b>	<b>0.957</b>
L2 (1%)	Literature [30] Classical LSTM	0.0403	0.846	0.886
	Literature [31] XGBoost	0.0464	0.845	0.885
	<b>Proposed LSTM-SVDD</b>	<b>0.0162</b>	<b>0.970</b>	<b>0.976</b>
L3 (1.5%)	Literature [30] Classical LSTM	0.0431	0.897	0.921
	Literature [31] XGBoost	0.0453	0.895	0.919
	<b>Proposed LSTM-SVDD</b>	<b>0.0167</b>	<b>0.971</b>	<b>0.977</b>
L4 (2%)	Literature [30] Classical LSTM	0.0423	0.943	0.953
	Literature [31] XGBoost	0.0446	0.941	0.952
	<b>Proposed LSTM-SVDD</b>	<b>0.0163</b>	<b>0.999</b>	<b>0.999</b>

In actual engineering, the delays of the HVAC system make it difficult to detect faults in time, thus further causing system damage and energy loss. The method proposed in this research can avoid such instances to some extent.

## 6. Conclusions

HVAC systems work with complex processes and delayed fault propagation. A fault detection method combining a simulation model with the LSTM-SVDD algorithm was proposed in order to achieve the real-time fault detection of HVAC systems to identify faults. The introduction of fault modes in the system simulation model solves the difficulty in collecting sufficient fault data samples due to the fact that the HVAC system operates

under normal conditions most of the time. Due to the time-series of the operational data of the HVAC system, the LSTM-SVDD fault detection method was proposed to perform real-time fault detection on the system, and the detection efficiency of the method was discussed by introducing different fault sizes. The results show that the proposed method still has a detection accuracy of 0.946 for a small deviation fault of 0.5%. An average increase of 0.284 and 0.29 in the F1-measure was achieved with the proposed method in comparison to the traditional LSTM network and XGBoost methods with a fixed threshold, respectively. Consequently, the proposed fault detection method is expected to timely detect system faults, prevent further deterioration of equipment and prolong the equipment life. The method can also contribute to the research of fault detection methods for other HVAC systems in the absence of fault samples.

**Author Contributions:** Methodology, H.Z. and W.Y.; software, H.Z.; validation, H.Z., W.Y. and A.P.; formal analysis, A.P.; investigation, S.L.; data curation, H.Z.; writing—original draft preparation, H.Z.; writing—review and editing, H.Z.; supervision, A.P.; funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the National Natural Science Foundation of China Regional Project 12162007, in part by the Guizhou University Incubation Program, Guizhou University Incubation (2019) 60; in part by the Guizhou Provincial Education Department Young Talent Growth Project, QiankeheKY (2021) 100; and in part by the Science and Technology Fund of the Department of Science and Technology of Guizhou Province, Qian Kehe Foundation (2020) 1Y273. MOE Key Laboratory of TianQin Project, Satellite Active Disturbance Rejection High-gain drag-free Control for Gravitational Wave Detection.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hu, X.; Du, Z.; Jin, X.; Chen, Z. Fault diagnosis based operation risk evaluation for air conditioning systems in data centers. *Build. Environ.* **2019**, *163*, 106319.1–106319.14.
2. Zabala, L.; Febres, J.; Sterling, R.; López, S.; Keane, M. Virtual testbed for model predictive control development in district cooling systems. *Renew. Sustain. Energy Rev.* **2020**, *129*, 109920. [[CrossRef](#)]
3. Afroz, Z.; Shafiullah, G.M.; Urmee, T.; Higgins, G. Modeling techniques used in building HVAC control systems: A review. *Renew. Sustain. Energy Rev.* **2018**, *83*, 64–84. [[CrossRef](#)]
4. Mirmaghi, M.S.; Haghighat, F. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy Build.* **2020**, *229*, 110492. [[CrossRef](#)]
5. Du, Z.; Fan, B.; Jin, X.; Chi, J. Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis. *Build. Environ.* **2014**, *73*, 1–11. [[CrossRef](#)]
6. Magoules, F.; Zhao, H.X.; Elizondo, D. Development of an RDP neural network for building energy consumption fault detection and diagnosis. *Energy Build.* **2013**, *62*, 133–138. [[CrossRef](#)]
7. Yan, X.; Guan, T.; Fan, K.; Sun, Q. Novel double layer BiLSTM minor soft fault detection for sensors in air-conditioning system with KPCA reducing dimensions. *J. Build. Eng.* **2021**, *44*, 102950. [[CrossRef](#)]
8. Zhao, Y.; Li, T.; Zhang, X.; Zhang, C. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renew. Sustain. Energy Rev.* **2019**, *109*, 85–101. [[CrossRef](#)]
9. Shi, Z.; O'Brien, W. Development and implementation of automated fault detection and diagnostics for building systems: A review. *Autom. Constr.* **2019**, *104*, 215–229. [[CrossRef](#)]
10. Fan, Y.; Cui, X.; Han, H.; Lu, H. Chiller fault diagnosis with field sensors using the technology of imbalanced data. *Appl. Therm. Eng.* **2019**, *159*, 113933. [[CrossRef](#)]
11. Yan, K.; Huang, J.; Shen, W.; Ji, Z. Unsupervised learning for fault detection and diagnosis of air handling units. *Energy Build.* **2020**, *210*, 109689.1–109689.10. [[CrossRef](#)]
12. Zhang, R.; Hong, T. Modeling of HVAC operational faults in building performance simulation. *Appl. Energy* **2017**, *202*, 178–188. [[CrossRef](#)]
13. Kim, J.; Frank, S.; Braun, J.E.; Goldwasser, D. Representing small commercial building faults in energyplus. Part I: Model development. *Buildings* **2019**, *9*, 233. [[CrossRef](#)]

14. Leach, M.; Kim, J. Curated Modeled Fault Data Set. 2019. Retrieved. Available online: <https://openei.org/doe-opendata/dataset/curated-modeled-fault-data-set> (accessed on 26 December 2021).
15. Hernandez-Albaladejo, G.; Urquia, A. Modelling of Low-Temperature Solar Thermal Systems with Modelica. *IFAC-Pap.* **2018**, *51*, 783–788. [[CrossRef](#)]
16. Hefni, B.E.; Bouskela, D. *Heat Exchanger Modeling: A Theoretical Introduction and a Practical Guide*; Springer: Berlin/Heidelberg, Germany, 2019.
17. Ridwana, I. Modeling of Building Energy Consumption by Integrating Regression Analysis and Artificial Neural Network with Data Classification. *Buildings* **2020**, *10*, 198. [[CrossRef](#)]
18. Kayacan, E.Y. Deep Learning for Time Series Forecasting. In *Selected Topics in Applied Econometrics*; Akay, E.Ç., Korkmaz, Ö., Eds.; Peter Lang: Berlin, Germany, 2019; pp. 243–254.
19. Li, M.; Li, Y.; Min, X. Practice and Application of LSTM in Temperature Prediction of HVAC System. In Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 12–14 June 2020; IEEE: Piscataway, NJ, USA.
20. Zhang, S.; Wang, Y.; Liu, M.; Bao, Z. Data-based Line Trip Fault Prediction in Power Systems Using LSTM Networks and SVM. *IEEE Access* **2017**, *6*, 7675–7686. [[CrossRef](#)]
21. Ding, S.X. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*; Springer-Verlag: Berlin/Heidelberg, Germany, 2008.
22. Kim, W.; Katipamula, S. A review of fault detection and diagnostics methods for building systems. *Sci. Technol. Built Environ.* **2017**, *24*, 3–21. [[CrossRef](#)]
23. Hyvarinen, J.; Karki, S. *Building Optimization and Fault Diagnosis Source Book, IEA Annex 25*; International Energy Agency: Paris, France, 1995.
24. Reddy, T.A. Automated Fault Detection and Diagnosis for HVAC&R Systems: Functional Description and Lessons Learnt. In *ASME 2008 2nd International Conference on Energy Sustainability Collocated with the Heat Transfer, Fluids Engineering, and 3rd Energy Nanotechnology Conferences*; American Society of Mechanical Engineers: Jacksonville, FL, USA, 2008; pp. 589–599.
25. Wang, F.; Lin, W.; Liu, Z.; Wu, S.; Qiu, X. Pipeline Leak Detection by Using Time-Domain Statistical Features. *IEEE Sens. J.* **2017**, *17*, 6431–6442. [[CrossRef](#)]
26. Li, G.; Hu, Y.; Chen, H.; Li, H.; Hu, M.; Guo, Y.; Shi, S.; Hu, W. A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based D-statistic and DV-contribution plots. *Energy Build.* **2016**, *133*, 230–245. [[CrossRef](#)]
27. Choi, W. Adjustment of Multiple Variables for Optimal Control of Building Energy Performance via a Genetic Algorithm. *Buildings* **2020**, *10*, 195.
28. Filonenko, K.; Ljungdahl, V.B.; Yang, T.; Veje, C. Modelica implementation of phase change material ventilation unit. In Proceedings of the 2020 6th IEEE International Energy Conference (ENERGYCon), Gammarth, Tunisia, 28 September–1 October 2020; pp. 464–467. [[CrossRef](#)]
29. Wang, S.; Qiang, Z.; Fu, X. A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults. *Energy Build.* **2010**, *42*, 477–490. [[CrossRef](#)]
30. Zou, S.; Zhao, W.; Wang, C.; Chen, F. Fault Detection Strategy of Vehicle Wheel Angle Signal via Long Short-Term Memory Network and Improved Sequential Probability Ratio Test. *IEEE Sens. J.* **2021**, *21*, 17290–17299. [[CrossRef](#)]
31. Guo, K.; Liu, L.; Shi, S.; Liu, D.; Peng, X. UAV Sensor Fault Detection Using a Classifier without Negative Samples: A Local Density Regulated Optimization Algorithm. *Sensors* **2019**, *19*, 771. [[CrossRef](#)] [[PubMed](#)]
32. Chakraborty, D.; Elzarka, H. Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold. *Energy Build.* **2019**, *185*, 326–344. [[CrossRef](#)]