Energy and AI 16 (2024) 100364

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Fault detection and diagnosis of electric bus air conditioning systems incorporating domain knowledge and probabilistic artificial intelligence

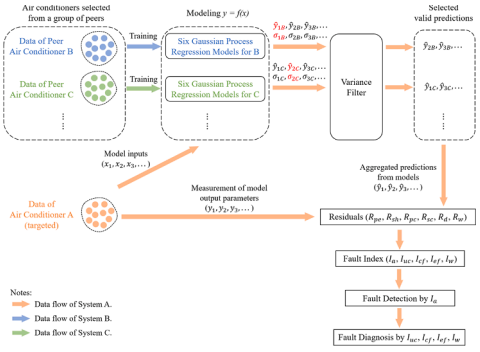
Fangzhou Guo [a,b], Zhijie Chen [a], Fu Xiao [a,c,*]

[a] Department of Building Environment and Energy Engineering, Hong Kong Polytechnic University, Hong Kong
[b] Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[c] Research Institute for Smart Energy, Hong Kong Polytechnic University, Hong Kong

HIGHLIGHTS

- FDD is performed by comparing features among peer electric bus air conditioners.
- Gaussian process regression is adopted to quantify the uncertainty of predictions.
- Domain expertise is used for feature selection and performing fault diagnosis.
- The method can label faulty systems efficiently with low false positives/negatives.
- Even 1/3 of bus systems are faulty, accuracy of the method can still be guaranteed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Fault detection and diagnosis
Predictive maintenance
Air conditioner
Electric vehicle
Data-driven model
Gaussian process

ABSTRACT

The air conditioning systems in electric city buses usually operate in rapidly changing ambient conditions and are more likely to suffer from mechanical faults. Although many fault detection and diagnosis (FDD) methods have been developed for building air conditioning systems, they are difficult to be applied to bus air conditioners since its operation is highly dynamic and fault-free data are usually unavailable. Therefore, this paper proposes an FDD method for electric bus air conditioners to tackle the above issues. First, the method identifies faults in an unsupervised manner by comparing selected features among a group of peer systems. Then, considering the features are influenced by the operating conditions, Gaussian process regression (GPR) models are established to find the relationships between each feature and its influential parameters. The probabilistic nature of the GPR is used to differentiate predictions with large uncertainty, which are then excluded from FDD. In this way, robustness of the method is evidently improved. Finally, fault indexes are defined to detect and diagnose mechanical faults. We applied the method to a group of air conditioners in a city bus fleet. Results showed that it can effectively identify refrigerant undercharge and indoor and outdoor fan problems with low false positive/genitive rates. Also, the method is highly robust and not sensitive to the faulty systems in the bus fleet.

* Corresponding author.
E-mail address: linda.xiao@polyu.edu.hk (F. Xiao).

https://doi.org/10.1016/j.egyai.2024.100364

Available online 27 March 2024

Nomenclatures

Control signals

f_{comp}	compressor frequency (Hz)
f_{fani}	indoor fan frequency (Hz)
f_{fano}	outdoor fan frequency (Hz)

Physical parameters from measurement

P_e	evaporating pressure (kPa)
P_c	condensing pressure (kPa)
$T_{c,o}$	condenser outlet temperature (°C)
T_d	compressor discharge temperature (°C)
T_i	indoor (cabin) temperature (°C)
T_o	outdoor temperature (°C)
T_{sat}	saturation temperature (°C)
T_{sc}	liquid-line subcooling (°C)
T_{sh}	suction-line superheat (°C)
T_{suc}	compressor suction temperature (°C)
W	compressor power (kW)

Residuals

R_y	referring to any residuals among R_{pe}, R_{sh}, R_{pc}, R_{sc}, R_d, and R_w
R_{pe}	residual of evaporating pressure (kPa)
R_{sh}	residual of superheat (°C)
R_{pc}	residual of condensing pressure (kPa)
R_{sc}	residual of subcooling (°C)
R_d	residual of compressor discharge temperature (°C)
R_w	residual of compressor power (kW)

Fault indexes

I_a	index of abnormality
I_{uc}	index of undercharge
I_{cf}	index of condenser fault
I_{ef}	index of evaporator fault
I_w	index of abnormally high power
I'_{uc}	transformed index of undercharge
I'_{cf}	transformed index of condenser fault
I'_{ef}	transformed index of evaporator fault

Abbreviations

COP	coefficient of performance
CV(RMSE)	coefficient of variation of root mean square error
FDD	fault detection and diagnosis
GPR	Gaussian process regression
IQR	interquartile range
MAPE	mean absolute percentage error
Q3	the 75th percentile of data, the upper quartile
RBF	radial basis function (kernel)
RQ	rational quadratic (kernel)
SND	standard normal distribution

Other symbols

x	training input of the Gaussian process regression model
x_*	test input of the Gaussian process regression model
y	training output of the Gaussian process regression model
f_*	test output of the Gaussian process regression model
μ	mean function of the Gaussian process regression model
Σ	covariance matrix of the Gaussian process regression model
K	kernel matrix in Gaussian process
$N(\mu,\sigma^2)$	normal distribution with a mean of μ and a variance of σ^2
σ	standard deviation
ε	white noise
$\widehat{y}$	predicted values of any parameters among P_e, T_{sh}, P_c, T_{sc}, T_d, and W
y	measured values of any parameters among P_e, T_{sh}, P_c, T_{sc}, T_d, and W
i	a data point of the system to be diagnosed
j	the number of a peer system
k	degree of freedom in a statistical distribution

1. Introduction

The proliferation of electric buses in densely populated urban areas has surged in popularity. When contrasted with traditional diesel buses, the integration of electric buses yields reductions in energy consumption [1], greenhouse gas emissions [2] and air pollution [3], thereby advancing the pursuit of carbon neutrality and enhancing urban environmental quality.

Nevertheless, electric buses face a few drawbacks such as limited driving range and extended charging durations, potentially leading to lengthy queues at charging stations [4]. To overcome this issue, researchers have focused on finding the optimal locations and capacities of bus charging stations [5], refining the public transit schedules [6] and enhancing charging timetables [7]. Besides, another useful method is to reduce energy consumption during the operation of electric buses and improve the driving range. Although energy consumption mainly comes from the electric motor which already has a relatively high efficiency [8], the auxiliary systems can also consume 10–30 % of the total energy [1]. Moreover, among all auxiliary systems of an electric bus, the air conditioning system consumes the largest amount of electrical energy, and thus could significantly reduce the driving range [9]. What's more, bus air conditioners operate in poor and fast-changing indoor and outdoor environments. Thus, compared to air conditioning systems installed in buildings, the operation of bus air conditioners is highly dynamic, causing the equipment to wear out faster and being more likely to have faults.

To ensure bus air conditioning systems operate at optimal efficiency and consume less energy, predictive maintenance through fault detection and diagnosis (FDD) is critical. Nowadays with the advancement of the Internet of Things (IoT) technology, an increasing number of city bus fleets are opting to integrate IoT sensors for fleet management and predictive maintenance. For example, Killeen et al. [10] proposed an IoT architecture for a bus fleet which can diagnose faulty buses using IoT sensor measurements combined with semi-supervised machine learning algorithms. Massaro et al. [11] developed an IoT system with smart electronic control units that can extract all data from a bus. The IoT system was connected with an artificial intelligence engine in the cloud which can perform FDD for the bus fleet using artificial neural network algorithms. Following this idea, if the IoT sensors are installed on the air conditioning systems of the city buses, FDD algorithms can be executed on a remote platform to continuously monitor the operating conditions of each system. Consequently, any operating faults in bus air conditioners can be promptly identified, enhancing the effectiveness of predictive maintenance strategies.

While FDD techniques specifically tailored for vehicle air conditioning systems remain scarce, extensive FDD methodologies have been developed for small commercial and residential building air conditioners, offering valuable insights. Katipamula and Brambley [12] categorized the FDD methods for building energy systems into quantitative model-based, qualitative model-based, and process history-based. Quantitative model-based methods rely on detailed or simplified physical models to perform FDD. Because of the complexity and the

significant effort to develop a physical model, very few studies adopted this approach. Qualitative model-based methods mainly include rule-based methods and qualitative physics-based methods. The rule-based methods are used in many studies. For example, in [13], the researchers measured the steady-state values of key physical parameters in the refrigeration cycle before and after faults were manually imposed, and then generated rules and thresholds based on the measurements to label different types of faults. For qualitative physics-based methods, a typical approach is virtual sensor-based, which involves using low-cost temperature and pressure sensors along with simple physics to estimate critical physical quantities such as refrigerant charge [14], mass flow rate [15], power consumption and cooling capacity [16].

In recent years, process history-based (also called data driven-based) methods have become increasingly popular for FDD of air conditioning systems [17]. The methods can be further categorized into classification-based, regression-based, and unsupervised learning-based [18]. They rely on data collected from sensor records to develop artificial intelligence-based models (e.g. machine learning [19], deep learning models [20]) or statistical models, and domain expertise is not necessarily required. Some researchers focus on specific types of data driven-based FDD methods to solve practical problems. For example, Wang et al. [21] adopted a growing Gaussian mixture regression model to continuously learn a system's behavior as the performance of the system changes. Li et al. [22] proposed deep transfer learning strategies to improve the performance of the proposed FDD algorithm among different systems. Fan et al. [23] adopted self-supervised learning approaches to tackle FDD problems with data availability issues (i.e., limited labeled data). Eom et al. [24] used convolutional neural networks to predict the refrigerant charge level of air conditioners at faulty conditions, in order to guide repairing technicians. Guo et al. [25] applied statistical methods to perform FDD for residential systems with only thermostat data being available. Other researchers focus on analyzing the generalizability and restrictions of data-driven methods. For example, Zhong et al. [26] concluded that machine learning-based algorithms trained and tested under one season of the year may produce poor generalizability and fail to identify faults under other unseen seasons. Bode et al. [27] found that machine learning methods can still perform poorly on real-world datasets though they perform well on experimental datasets.

Generally, the data driven-based FDD methods consist of the following five steps: feature selection, steady-state detection, fault-free modeling, fault detection, and fault diagnosis [28]. After features are selected, usually steady-state operations should be detected and the data should be filtered accordingly. Then, fault-free data are required to model the normal behavior of the system and generate feature thresholds. After that, at the current operating conditions, the measured feature values are compared with the thresholds. If the measured values are beyond the thresholds, then the system is identified as faulty. Finally, the fault will be diagnosed by rules or classifiers.

The present study employs data-driven FDD methodologies to assess the performance of bus air conditioners, leveraging the abundant data collected by IoT sensors on a daily basis. However, these methodologies encounter several challenges when applied to bus air conditioning systems. First, the operational behavior of bus air conditioners has strong dynamics. For example, electric buses have door openings, passengers entering and leaving at each bus stop, and receive intermittent shadings from different angles when traveling on the road. These disturbances cause the air conditioner to change its control signals frequently, so that steady-state operating periods are hard to find. Second, the fault-free data are usually collected in a lab test at a few fixed operating conditions. However, the operating conditions of bus air conditioners vary a lot and the fault-free data collected in lab conditions may not be enough in practice. As a result, models describing the fault-free behavior of bus air conditioners are usually unavailable. Third, after a fault is detected by an FDD algorithm, technicians are expected to repair the system based on the detection results of the algorithm. However, because the majority of artificial intelligence-based models are not interpretable to technicians, the technicians may not trust the FDD algorithm and prefer to examine the faulty system by themselves. This situation increases the labor cost since the technicians need to go back and forth a few times to bring the measuring instruments, new parts, or refrigerant tanks.

Due to the above difficulties, service companies currently still rely on reactive maintenance, which means technicians will only be dispatched after receiving customer complaints (usually from the bus driver) of a complete equipment breakdown. Note that because of the limited labor resources, preventive maintenance through regularly conducting field tests for every bus air conditioner is not feasible. Recently, Guo et al. [29] and Chen et al. [30] developed two methods separately to benchmark the energy performance of a large group of electric bus air conditioners. Both of the methods could identify faulty air conditioners with abnormally high energy consumption. However, because the two methods did not analyze other useful physical parameters in the refrigeration cycle, they are still in the stage of preliminary fault detection and cannot identify the cause of faulty behavior.

This paper proposes a fault detection and diagnosis method to further promote predictive maintenance for electric bus air conditioners. The method incorporates both machine learning methods for establishing models and domain expertise for developing rule-based classifiers, which can solve the abovementioned challenges as follows. First of all, because fault-free data are not available, the method relies on comparing the physical parameters of a system to its peer systems of the same model operating at similar conditions. Because city bus fleets usually own a lot of electric buses of the same model operating on different routes, these bus air conditioners can be regarded as a large group of peers. Assuming the majority of systems are fault-free, a system is considered faulty if its performance deviates from the average performance of its peers.

Second, since the operation of bus air conditioners is highly dynamic and operating conditions vary a lot, the method directly uses the unsteady-state data for analysis. However, in unsteady-state conditions, the physical parameters change with operating conditions, control signals, and other physical parameters. Therefore, the method identifies the crucial relationships among physical parameters using domain expertise and then develops data-driven models for each system to find the relationships. For a group of peer systems, the relationships should be very similar in fault-free conditions. But when a fault occurs in a system, the relationships are expected to change, and the feature values measured from the faulty system will deviate from the predicted values of the majority of models, given that most models are developed by normal data. Also, considering the systems operate at different conditions and data-driven models may not be generalizable with extrapolation, Gaussian process regression (GPR) is applied to establish the models. GPR is a probabilistic artificial intelligence technique capable of delivering precise predictions from low volumes of historical data. One advantage of the Gaussian process is that it can quantify the uncertainty of the prediction as well as the prediction itself. A larger dissimilarity between the test inputs and training inputs will cause the GPR model to output higher prediction error (quantified by the prediction variance outputted by GPR) of the test data. The predictions with relatively large errors are considered invalid and can be eliminated.

Third, the method performs fault diagnosis by setting rules and thresholds based on the change of physical parameters according to domain expertise. For example, condenser faults result in higher condensing pressure and discharge temperature. These rules can be fully understood by technicians and provide useful information for them to repair the system. The thresholds are determined by assuming the residuals between measured and predicted values of fault-free systems are normally distributed with zero mean. In this way, statistical analysis can be conducted for the residuals and faults can be identified with a confidence level. For example, the diagnosis result could be "a condenser fault is detected with 99.9 % confidence".

Last but not least, the method can be applied daily and monitor the

health of each vehicle air conditioner in real-time. Thus, incipient faults can be detected and complete equipment failure can be avoided by predictive maintenance. The remainder of this paper will detail the proposed FDD method (Section 2) and then apply it to a group of real electric bus air conditioners to illustrate its effectiveness and robustness (Section 3).

2. Proposed real-time fault detection and diagnosis method

This section details the real-time FDD method for electric bus air conditioning systems. Section 2.1 provides a framework of the method and summarizes the procedures to detect and diagnose a fault using the performance of its peer systems. Then, the following sections discuss each crucial step of the method. Because the FDD method relies on reference models of peer systems to assess the operating conditions of each target system, Section 2.2 lists critical physical parameters that are sensitive to faults as modeling outputs according to previous research papers, and Section 2.3 determines the modeling input parameters according to domain knowledge. Then, Section 2.4 introduces the Gaussian process regression to model relationships between the inputs and outputs, and explains how the Gaussian process helps improve the modeling accuracy and robustness. After that, Section 2.5 defines the index of each type of fault based on the residuals between measured physical parameters and their corresponding model outputs. Finally, in Section 2.6, the fault indexes are applied to identify if a system has fault and diagnose the specific type of fault.

2.1. Overview of the FDD method

The process to perform FDD for an air conditioner within a group of peers is shown in Fig. 1. Take *air conditioner A* as the target. Firstly, 6 GPR models are developed for each peer air conditioner (i.e. *B, C*, etc.) to model the relationships between operating parameters in the refrigeration cycle. Each of the 6 models predict one physical parameter, namely the evaporating pressure (P_e), suction temperature (T_{suc}), condensing pressure (P_c), condenser outlet temperature ($T_{c,o}$), discharge temperature (T_d), and total power consumption (W). Secondly, the regression models of peer air conditioners are employed to make predictions using the input data of *air conditioner A*. Because the Gaussian process regressor quantifies the uncertainty (represented by the standard deviation σ) of each prediction, a variance filter is then applied to eliminate the predictions with high uncertainty (marked as red in Fig. 1) caused by inappropriate model extrapolation. The remaining predictions from all peer systems are aggregated by taking the median. Thirdly, the physical parameters of *air conditioner A* are subtracted by the model predictions to obtain residuals, including the residual of condensing pressure, the residual of subcooling, etc. Fourthly, the index of each type of fault is calculated using the residuals according to specific rules obtained from previous experiments, including the index of abnormality (I_a), index of undercharge (I_{uc}), index of condenser fault (I_{cf}), index of evaporator fault (I_{ef}), and index of abnormally high power (I_w). Finally, the fault detection and diagnosis results can be obtained based on statistical analysis of the fault indexes. The value of I_a is used to detect if *air conditioner A* has a fault. If a fault is detected, then the values of other fault indexes are used to diagnose the type of fault. After *air conditioner A* is examined, the other air conditioners in this group of peers can be diagnosed similarly.

Additionally, as shown in Fig. 2, the FDD method can be executed in real-time on a daily basis. At the end of each day, the operating data of the air conditioners are retrieved and preprocessed. GPR models are retrained for each system using data in the last few days. Then, for each target system, the GPR models of its peers are used for predictions of the

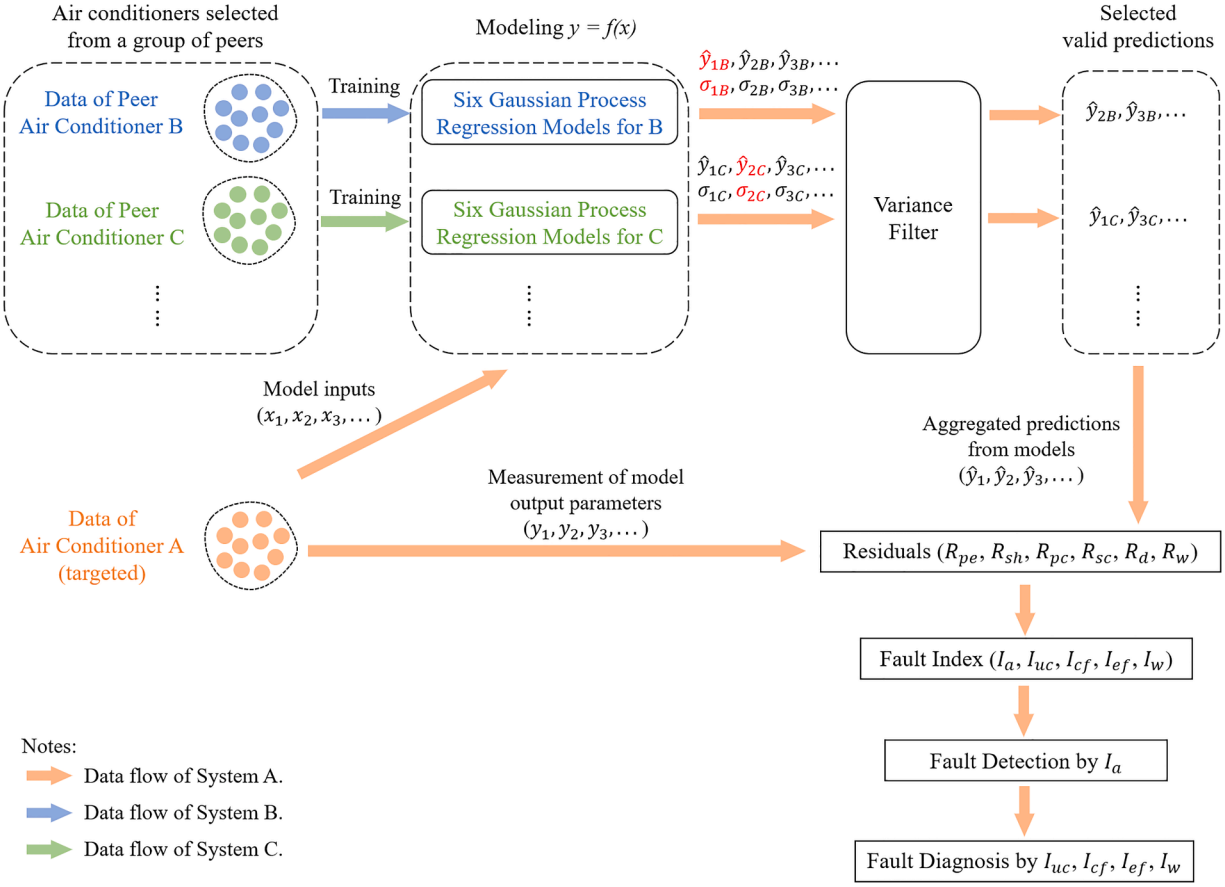


Fig. 1. Framework for fault diagnosis of *air conditioner A*. Each air conditioner in the group can be diagnosed with this framework.

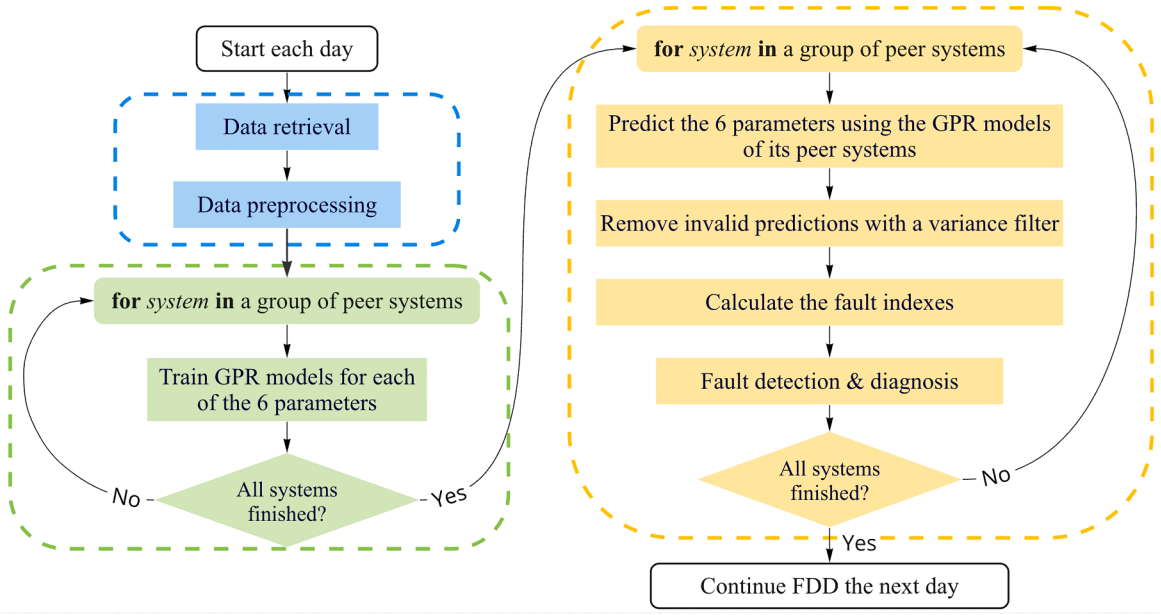


Fig. 2. Flowchart of the proposed real-time FDD method for a group of peer vehicle air conditioners.

6 physical parameters, and fault indexes are calculated using data of the target system on the current day. In this way, the performance of every system could be actively monitored. If a system has fault, then the corresponding fault index will rise above the detection thresholds for a consecutive of days and raise alarm for maintenance (see Section 2.6 for the details of the fault detection and diagnosis step). After a faulty system is repaired, the real-time FDD method continues monitoring the performance and evaluates if the system is fully repaired or whether another maintenance needs to be scheduled.

2.2. Feature selection

Fig. 3 shows the configuration of the electric bus air conditioner studied in this paper. The compressor, indoor and outdoor fans are variable-speed and are adjusted by the proportional controller. The control signals including the compressor frequency, the indoor fan frequency, and the outdoor fan frequency are used for FDD. Besides, the available sensors to perform FDD include five temperature sensors, two pressure sensors, one voltage meter, and one current meter. These sensors and control signals are all required for the proposed FDD method.

A key task in developing an FDD method is feature selection. In [31–33], eight features which are very sensitive to faults are used to diagnose system faults, including the evaporating temperature, condensing temperature, superheat, subcooling, discharge temperature, temperature drop in the liquid line, air temperature change across the condenser and the evaporator. Some researchers also proposed other features, especially the ones which could be measured from virtual sensors [28]. Considering the availability of sensors installed on the bus air conditioners, this paper chooses 6 features, namely the evaporating pressure (P_e), suction-line superheat (T_{sh}), condensing pressure (P_c), liquid-line subcooling (T_{sc}), discharge temperature (T_d), and total power consumption (W). These features can all be derived from the sensor measurements shown in Fig. 3.

Table 1 summarizes the impact of 7 types of faults on the 6 features. An up arrow indicates the fault causes a rise of the feature values, while a down arrow indicates the fault causes a decrease of the feature values. These expert rules are discovered by previous studies, which used the rules to develop rule-based FDD algorithms. For instance, if T_{sc} is lower than a predefined threshold in the testing condition, then the system is labeled with refrigerant undercharge [31]. In comparison, to be discussed in Section 2.5, our study will define several fault indexes based on

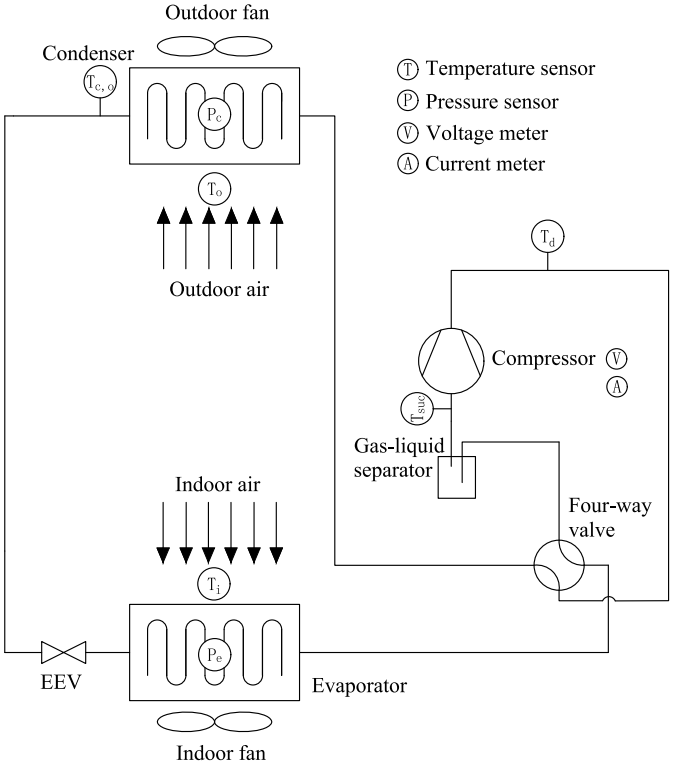


Fig. 3. Configuration of the vehicle air conditioning system (only the indoor cooling cycle is depicted). The sensors shown in the figure are required to perform the proposed FDD method.

these expert rules and perform fault diagnosis by integrating machine learning and statistical models. In this way, the proposed FDD method can be applied to a wide range of system operating conditions (not restricted to several testing conditions) and the thresholds to label faulty systems could be determined by rigorous statistical analysis.

This study only analyzes the first four types of faults in Table 1, including undercharge, condenser fault, evaporator fault, and abnormally high power, because they are more common and have been

Table 1
Fault characteristics of air conditioning systems [13,31,34]*.

	P_e	T_{sh}	P_c	T_{sc}	T_d	W
Refrigerant undercharge	↓	↑	↓	↓	↑	
Condenser fault			↑		↑	
Evaporator fault	↓		↓			
Abnormally high power						↑
Compressor wear	↑		↓			
Refrigerant overcharge			↑	↑	↑	
Non-condensable gas				↑	↑	

* Note: Researchers may have different opinions on the change of some physical parameters when a fault occurs. Thus, this table only shows the conclusive and evident fault characteristics for each type of fault. Only the rules shown in this table are used to develop the fault indexes.

detected in a few bus air conditioners studied in this paper. As to bus air conditioners, undercharge is often caused by active refrigerant leakage, condenser fault can be caused by condenser fouling or outdoor fan issues, and evaporator fault can be caused by evaporator fouling or indoor fan issues. The occurrence of abnormally high power is a symptom of faults typically caused by inefficiency of compressors and fans, but it can be analyzed along with other faults. The last three types of faults, namely compressor wear, overcharge, non-condensable gas, are less common and examples have not been found among bus air conditioners. Diagnosis of these three types of faults will not be discussed in this paper, but the proposed FDD method is still applicable.

2.3. Modeling the relationships between physical parameters

When the data are unsteady, the six selected features will change with time and thus directly comparing the magnitude of each feature is not feasible. For example, the total power consumption changes with the compressor and fan speed, evaporating temperature changes with the fluctuation of the cabin temperature, and the superheat and subcooling also vary with the frequent regulation of the compressor speed. Therefore, relationships between these features and other physical parameters should be modeled. Table 2 shows the proposed modeling inputs and outputs. In total six models are established, with each model predicting one parameter. Four features, P_e, P_c, T_d, and W, are predicted directly by the models, while the other two features, T_{sh} and T_{sc}, are calculated as follows:

$$T_{sh} = T_{suc} - T_{sat}(P_e) \tag{1}$$

$$T_{sc} = T_{sat}(P_c) - T_{c,o} \tag{2}$$

where $T_{sat}(P_e)$ and $T_{sat}(P_c)$ are obtained from thermal properties of the refrigerant, T_{suc} and $T_{c,o}$ are directly predicted from the models.

The main idea to identify proper input and output parameters is that the relationships between the inputs and outputs should not change under similar operation conditions when the system is normal, but should change when a fault occurs. Thus, the domain expertise, especially understanding of the refrigeration cycle and the control logic, is crucial. For example, because the outdoor fan speed is fully controlled by the condensing pressure with the same control logic for all systems, these two parameters always have a fixed relationship (the fan speeds are usually obtained from control signals instead of sensor measurements). That is to say, if the outdoor fan frequency is selected as an input to predict the condensing pressure, strictly speaking the prediction error will be zero and the condensing pressure will always seem to be normal no matter if a condenser fault occurs or not. However, indoor fan speed must be included as an input to predict the evaporating pressure. This is because the indoor fan speed can be either controlled by the difference between the indoor temperature and its setpoint, or by the bus driver manually. When an evaporator fault occurs, the evaporating pressure will probably change even when the indoor fan speed (the control signal) remains the same. But if the indoor fan speed is not included in the input features, then deviation of measured evaporating pressure from the predicted values may result from the difference in the indoor fan speed rather than faults. Additionally, the compressor frequency is included as an input feature for all models because it can significantly impact all output features.

Table 2
The table shows the modeling inputs and outputs for each air conditioning system. In total six models are developed for each system, and each model uses its respective input parameters (checked in the table) to predict a single output parameter.

		Output parameter for each of the six models					
		P_e	T_{suc}	P_c	$T_{c,o}$	T_d	W
Input parameters	P_e			✓	✓	✓	
	P_c	✓	✓				
	f_{comp}	✓	✓	✓	✓	✓	✓
	f_{fani}	✓	✓				✓
	f_{fano}						✓
	T_i	✓	✓				✓
	T_o			✓	✓	✓	✓

2.4. Model development with Gaussian process regression (GPR)

After the model inputs and outputs are determined according to the domain knowledge, Gaussian process is applied to develop regression models for each system. In comparison with other machine learning methods, GPR can produce probabilistic predictions which are assumed to follow the normal distribution. The mean of the normal distribution is used as the model prediction, and the variance of the normal distribution is often used to quantify the confidence or uncertainty of the prediction. When the test input is within the range of the training inputs, the GPR model has more confidence of the prediction and returns a lower variance. On the contrary, when the test input is out of the range of the training inputs, the model has to extrapolate, resulting in a higher variance with less confidence.

In this study, each regression model is developed by data from a single system and then used to evaluate the performance of other systems excluding itself, so the ability of the Gaussian process to quantify the prediction uncertainty becomes very useful. Because the bus air conditioners may not operate at the same conditions, each developed model has to extrapolate in the process of predicting features of other systems. However, data-driven black box models often have poor performance on extrapolation. Therefore, the extrapolated predictions possibly have relatively large uncertainty and eventually affect the overall accuracy of the FDD method. Gaussian process effectively solves this issue. In GPR models, the prediction uncertainty is characterized by the variance, and predictions with high variance can be identified and removed by a threshold.

GPR models the training set ($\boldsymbol{x}$, $\boldsymbol{y}$) with M data points as follows:

$$y = f(x) + \varepsilon \tag{3}$$

where $\boldsymbol{x}$ has N features and ε is the white noise in normal distribution, i.e. $\varepsilon \sim N(0,\ \sigma_y^2)$.

Then, considering a test set $\boldsymbol{x}_*$ which has M_* data points and its predictions denoted by f_*, a GPR model assumes $\begin{pmatrix} y & f_* \end{pmatrix}^T$ is in multivariate normal distribution:

$$\begin{pmatrix} y \\ f_* \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x) \\ \mu(x_*) \end{pmatrix}, \begin{pmatrix} K_y & K_* \\ K_*^T & K_{**} \end{pmatrix}\right) \tag{4}$$

and

$$K_y = K + \sigma_y^2 I_M \tag{5}$$

where $\boldsymbol{y}$ is in the size of $M \times\ 1$, f_* is in the size of $M_* \times\ 1$, $\mu(x)$ is the

mean function of x, $\mu(x_*)$ is the mean function of x_*, I_M is an identity matrix of size M, the K matrix is in the size of $M \times M$ defined as the kernels between x and x, the K_* matrix is in the size of $M \times M_*$ defined as the kernels between x and x_*, and the K_{**} matrix is in the size of $M_* \times M_*$ defined as the kernels between x_* and x_*.

Finally, the conditional probability of f_* can be proved also in normal distribution:

$$P(f_* \mid x,\ x_*,\ y) = N(f_* \mid \mu_*,\ \Sigma_*) \tag{6}$$

where the mean function (μ_*) and the covariance matrix (Σ_*) are calculated by:

$$\mu_* = \mu(x_*) + K_*^T K_y^{-1}(y - \mu(x)) \tag{7}$$

$$\Sigma_* = K_{**} - K_*^T K_y^{-1} K_* \tag{8}$$

where μ_* is the mean of each predicted f_* and the diagonal of Σ_* is the variance of each predicted f_*.

One critical assumption of the Gaussian process is that two similar data points should have similar target values. This similarity in the target values is defined by kernels, also known as covariance functions in Gaussian process. In this study, we adopt a combination of the radial basis function (RBF) kernel, rational quadratic (RQ) kernel, and the white kernel:

$$kernel(x_i, x_j) = \theta_1 \exp\left(-\frac{\| x_i - x_j \|^2}{2\theta_2^2}\right) + \theta_3 \left(1 + \frac{\| x_i - x_j \|^2}{2\theta_4 \theta_5^2}\right)^{-\theta_6} + \theta_7 \tag{9}$$

where θ_1 to θ_6 are parameters to be optimized, θ_7 is zero when $i = j$ and is a parameter to be optimized when $i \neq j$. The first term in the kernel function is the RBF kernel, which is very commonly-used in Gaussian process. Because it is infinitely differentiable, GPR models with this kernel are very smooth. The second term is the RQ kernel, which can be regarded as an infinite sum of RBF kernels with different length-scale parameters (i.e., the θ_2 parameter in the RBF kernel). The last term is the white kernel, which estimates the global noise level from the data. Because the FDD method develops a large number of GPR models, using a combination of multiple kernels ensures the model is generalizable to more training datasets. In the cases where not all three kernels are needed, the corresponding parameters will approach zero or infinite after optimization. For example, if the RQ kernel is not useful in a GPR model, then the training results will show that $\theta_3 \to 0$ or $\theta_6 \to +\infty$.

The GPR models developed by each system are used to diagnose other systems excluding itself. For example, in order to diagnose *air conditioner A*, all GPR models of its peer air conditioners are used to predict the 6 physical parameters (i.e. P_e, P_c, T_d, W, plus T_{sh} and T_{sc} calculated by Eqs. (1) and (2)) at the operating conditions of *air conditioner A*. Then, as shown in Fig. 4, for each of the 6 physical parameters, variances of all predictions are ranked and the 75th percentile (Q3) of the ranked variance data is calculated. After that, a variance filter eliminates all predictions with their variance greater than Q3. In other words, one fourth of the predictions are considered having large uncertainty due to model extrapolation and not used for the final step of diagnosis.

2.5. Defining the fault indexes

For a target system to be diagnosed, after GPR model predictions of its peer systems are filtered according to the prediction variances (σ^2), a few fault indexes are computed by the remaining prediction means (μ). First of all, for each data point of the target system, since there are usually predictions from the GPR models of more than one peer system, the median of these predictions (from different peer systems) is calculated:

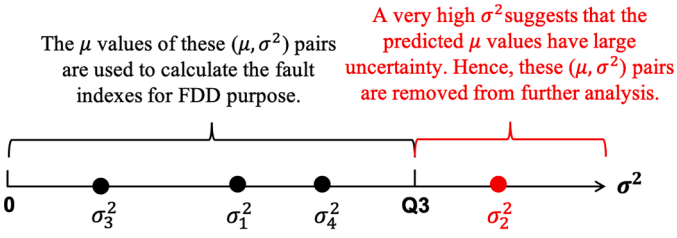


Fig. 4. For each of the 6 physical parameters, GPR model predictions with variances greater than Q3 are removed from further analysis. Predictions with high variances are considered having large uncertainty due to model extrapolation. Note: because the GPR model outputs a Gaussian distribution for each prediction instead of a single value, in this figure, μ stands for the mean of the Gaussian distribution which is equivalent to the predicted variable $\hat{y}$, and σ^2 stands for the variance of the Gaussian distribution (variance of prediction).

$$\hat{y}_i = median(\hat{y}_{ij}) \tag{10}$$

where i denotes a data point of the system to be diagnosed, j denotes the number of a peer system, and $\hat{y}$ refers to the predicted values (i.e., mean (μ) of the Gaussian distributions outputted by GPR models) of any physical parameters among P_e, T_{sh}, P_c, T_{sc}, T_d, and W. The median is used in Eq. (10) instead of the mean because the peers may contain faulty systems, causing a minority of predictions of the physical parameters are anomalies. The median is more robust than the mean in that a few abnormal values have much less influence on the median.

Next, the residuals for each of the 6 parameters are calculated from $\hat{y}_i$:

$$R_y = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i) \tag{11}$$

where i denotes a data point of the system to be diagnosed, N denotes the total number of data points, y refers to the measured values, and R_y refers to any residuals among R_{pe}, R_{sh}, R_{pc}, R_{sc}, R_d, and R_w. After the 6 residuals of the system are obtained, 5 metrics termed *fault indexes* are proposed. The 5 fault indexes are calculated according to Eqs. (12)–(16) using the expert rules in Table 1:

- Index of abnormality

$$I_a = \left(\frac{R_{pe}}{IQR(R_{pe})} \cdot IQR(SND)\right)^2 + \left(\frac{R_{sh}}{IQR(R_{sh})} \cdot IQR(SND)\right)^2 + \left(\frac{R_{pc}}{IQR(R_{pc})} \cdot IQR(SND)\right)^2 + \left(\frac{R_{sc}}{IQR(R_{sc})} \cdot IQR(SND)\right)^2 + \left(\frac{R_d}{IQR(R_d)} \cdot IQR(SND)\right)^2 + \left(\frac{R_w}{IQR(R_w)} \cdot IQR(SND)\right)^2 \tag{12}$$

- Index of undercharge

$$I_{uc} = (-1)\cdot sgn(R_{pe})\cdot\left(\frac{R_{pe}}{IQR(R_{pe})}\cdot IQR(SND)\right)^2 +(+1)\cdot sgn(R_{sh})\cdot\left(\frac{R_{sh}}{IQR(R_{sh})}\cdot IQR(SND)\right)^2 +(-1)\cdot sgn(R_{pc})\cdot\left(\frac{R_{pc}}{IQR(R_{pc})}\cdot IQR(SND)\right)^2 +(-1)\cdot sgn(R_{sc})\cdot\left(\frac{R_{sc}}{IQR(R_{sc})}\cdot IQR(SND)\right)^2 +(+1)\cdot sgn(R_{d})\cdot\left(\frac{R_{d}}{IQR(R_{d})}\cdot IQR(SND)\right)^2 \quad (13)$$

- Index of condenser fault

$$I_{cf} = (+1)\cdot sgn(R_{pc})\cdot\left(\frac{R_{pc}}{IQR(R_{pc})}\cdot IQR(SND)\right)^2 +(+1)\cdot sgn(R_{d})\cdot\left(\frac{R_{d}}{IQR(R_{d})}\cdot IQR(SND)\right)^2 \quad (14)$$

- Index of evaporator fault

$$I_{ef} = (-1)\cdot sgn(R_{pe})\cdot\left(\frac{R_{pe}}{IQR(R_{pe})}\cdot IQR(SND)\right)^2 +(-1)\cdot sgn(R_{pc})\cdot\left(\frac{R_{pc}}{IQR(R_{pc})}\cdot IQR(SND)\right)^2 \quad (15)$$

- Index of abnormally high power

$$I_{w} = (+1)\cdot sgn(R_{w})\cdot\left(\frac{R_{w}}{IQR(R_{w})}\cdot IQR(SND)\right)^2 \quad (16)$$

where $sgn(\cdot)$ is the sign function, $IQR(SND)$ denotes the interquartile range of standard normal distribution which equals 1.349, $IQR(R_y)$ denotes the interquartile range of the residual R_y which could be calculated from the R_y of a large number of peer systems within a relatively long time (e.g. one month).

In Eqs. (12)–(16), each R_y is first normalized using the IQR of itself and the IQR of SND. Note that IQR instead of the standard deviation of R_y is used because IQR is more robust to outliers. Since the majority of peer systems are assumed to be fault-free, this normalization procedure can scale the standard deviation of R_y for fault-free systems to one. Then, another assumption in this study is that residuals (R_y) of fault-free systems follow the normal distribution with a mean of zero. Hence, the normalized R_y follows the standard normal distribution (SND). After that, the squares of selected normalized residuals are added or subtracted:

- Eq. (12) calculates the index of abnormality (I_a) of a system, which focuses on **fault detection**, and thus all 6 residuals are added up. If a system has a relatively large I_a, then the system has a greater chance to be faulty. In statistics, the sum of the squares of k independent SND random variables follows the χ^2 distribution with k degrees of freedom. Thus, percentiles in the χ^2 distribution can be applied to determine the magnitude of I_a. Fig. 5 shows the probability density function of the χ^2 distribution with 6 degrees of freedom, which is used to model the distribution of I_a. This study uses the 99th and 99.9th percentiles in the χ^2 distribution as thresholds, and labels the systems with I_a greater than the thresholds as outliers.
- Eqs. (13)–(16) calculate the index of a specific type of fault, which focus on **fault diagnosis**. Since the changes of physical variables for each type of fault have directions (i.e. either increase or decrease), each term of the residuals can be added or subtracted according to the rules in Table 1. Similarly, if a system has a relatively large fault index, then the system has a greater chance to have the specific fault. However, because not every term of the residuals is positive, these fault indexes do not follow the χ^2 distribution. Therefore, Monte Carlo simulation is applied to obtain the distribution of each fault index with different degrees of freedom. In the Monte Carlo simulation, 1,000,000 samples drawn from the standard normal distribution are used to generate the four distributions according to Eqs. (13)–(16). As shown in Table 3, the 99th and 99.9th percentiles in each distribution are used as thresholds. Note that because the residuals are normalized, the thresholds in Table 3 remain the same for any groups of peer air conditioning systems.

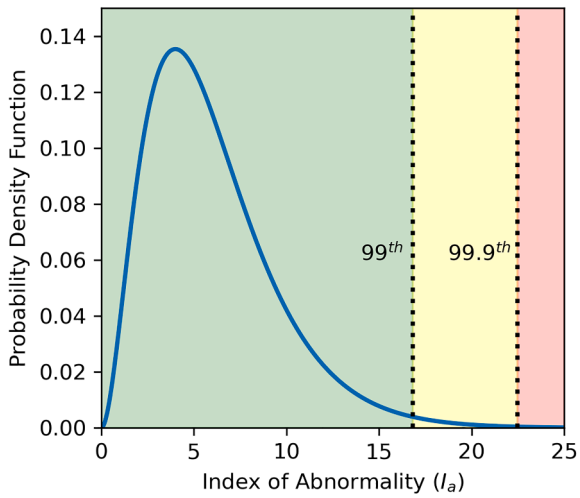


Fig. 5. Probability density function of the χ^2 distribution with 6 degrees of freedom, which is used to model the distribution of the index of abnormality (I_a) for all systems. The 99th percentile is 16.812 and the 99.9th percentile is 22.458. A system with I_a greater than the thresholds is possibly abnormal.

2.6. Fault detection and diagnosis using the fault indexes

Fig. 6 shows the fault detection and diagnosis process (as the details of the fault detection and diagnosis step in Fig. 2). Each day, detection of whether a fault occurs is based on the index of abnormality (I_a). A system will be flagged if on that particular day I_a is above the 99.9th percentile threshold (i.e. 22.458 calculated in Table 3). If the system is flagged at least twice in the latest three days, then the system is labeled as faulty and fault diagnosis will be conducted. This rule significantly reduces the false positive rate, since real faulty systems usually have very high I_a for multiple days while normal systems may have I_a slightly above the threshold temporarily under some extreme situations. Note that the days when the bus does not operate are not counted in the "latest three days" defined above. Also, if I_a is above the 99th percentile threshold which is 16.812, then the system will also be diagnosed though not labeled as faulty.

In fault diagnosis, the other four fault indexes (i.e. I_{uc}, I_{cf}, I_{ef}, and I_w) are used. If any of those four indexes are above the 99.9th percentiles for at least twice in the latest three days, the system will be labeled refrigerant undercharge, condenser fault, evaporator fault, or abnormally high power, and then repairing technicians need to be dispatched to examine the system. If none of the above conditions is true, then the system may suffer from other types of faults or the fault level is not

Table 3
Thresholds (i.e., the 99th and 99.9th percentiles) of all fault indexes calculated from their respective distributions. Note that I_a follows the χ^2 distribution, and the distributions of other fault indexes are obtained from Monte Carlo simulations. A system with a fault index greater than the thresholds is considered faulty. The thresholds are constants and do not change with air conditioning systems.

Fault index	Index of abnormality	Index of undercharge	Index of condenser fault	Index of evaporator fault	Index of abnormally high power
Symbol	I_a	I_{uc}	I_{cf}	I_{ef}	I_w
Degree of freedom (k)	6	5	2	2	1
99th percentile	16.812	10.190	7.068	7.068	5.412
99.9th percentile	22.458	15.440	11.568	11.568	9.556

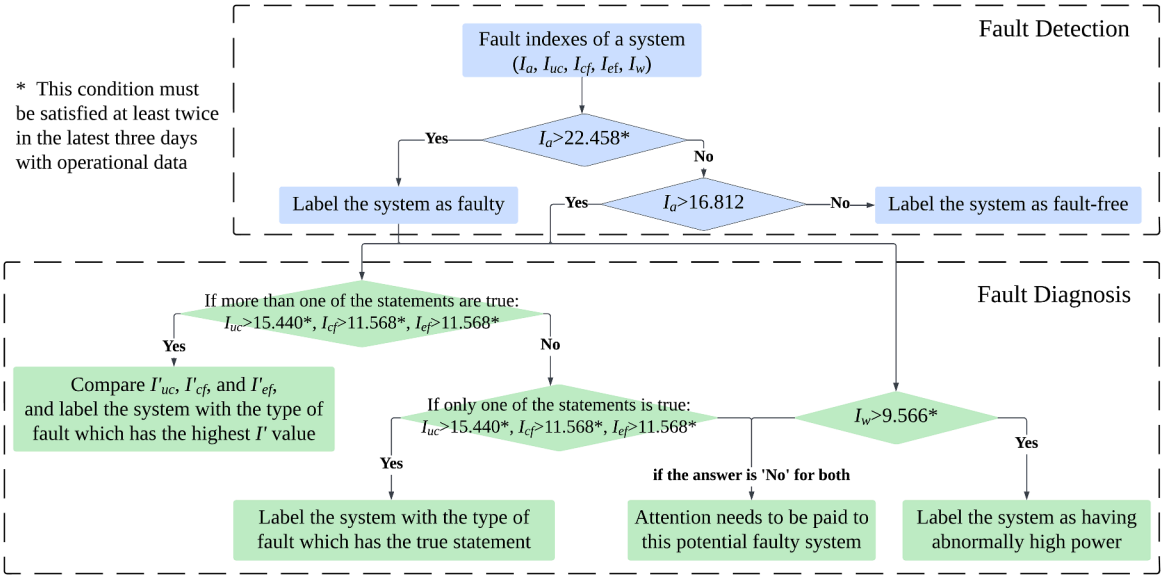


Fig. 6. The diagram shows the fault detection and diagnosis step for a system after fault indexes are calculated.

severe. In this case, an alarm is raised and more attention needs to be paid to this system. Sometimes faulty systems may have more than one of the four indexes above thresholds. The value of I_w is only calculated only by the residual of power consumption, a feature not used by other fault indexes, so the issue of abnormally high power is independent of other type of faults. But the calculation of I_{uc}, I_{cf}, and I_{ef} have used common features. These three indexes are correlated and thus the algorithm needs to distinguish them. This study prioritizes the fault types based on the value of each fault index divided by its own degree of freedom:

$$I' = \frac{I}{k} \tag{17}$$

where I refers to a fault index, k is the degree of freedom of the fault index, and I' is the transformed fault index. After transformation, I' of the three fault indexes are ranked and the most probable type of fault is the index with the highest I' value.

3. Implementation of the FDD method

The proposed FDD method is applied to 38 electric bus air conditioners in Shenzhen, China, with data from August 1st to August 24th in 2022 (24 days in total). The refrigerant used in the air conditioners is R410A. These electric buses operate on different routes in Shenzhen following a schedule. During this period, 12 faulty systems are verified by technicians: 5 undercharge systems due to refrigerant leaks, 5 systems with condenser fault due to outdoor fan problems, 1 system with evaporator fault due to indoor fan problems, and 3 systems with abnormally high power due to compressor motor problems (2 of them are among the other 11 faulty systems). A few of them are repaired during this period, but others still remain faulty.

3.1. Data preprocessing

The vehicle air conditioner has two operating modes, which are the indoor cooling mode to cool the bus cabin and the battery cooling mode that usually operate during battery charging. In this research, only the indoor cooling mode data are analyzed, which account for approximately 70 % of the total cycles. After the raw data are queried, the indoor cooling cycles of each air conditioner are extracted and resampled every one minute. The resampled data are preprocessed as follows:

1) Temperature records less than 0 °C or greater than 60 °C are removed.
2) Cycles with duration less than 10 min or with data update intervals longer than 2 min are removed.
3) The first 5 min of each indoor cooling cycle is removed considering the start-up transient behavior affects the subcooling and superheat.
4) Every input parameter for the GPR models is standardized by subtracting the mean and scaling to unit variance.

After preprocessing, the total number of time-averaged data points is 80,860. Considering the data are collected on 38 systems within 24 days and have been resampled to one minute, on average about 90 data points could be collected (i.e., 1.5-hour-long operation) for each system in each single day.

Fig. 7 shows the statistics of the 6 physical parameters that are used to calculate the fault indexes for each of the 38 bus air conditioners, with the 12 faulty systems marked with different colors. A visual comparison of the boxplots between these peer systems can identify some abnormal behaviors of the faulty systems. However, three important issues need to be addressed. First, the comparison between peer systems should be automated by an FDD method. Second, a visual comparison cannot

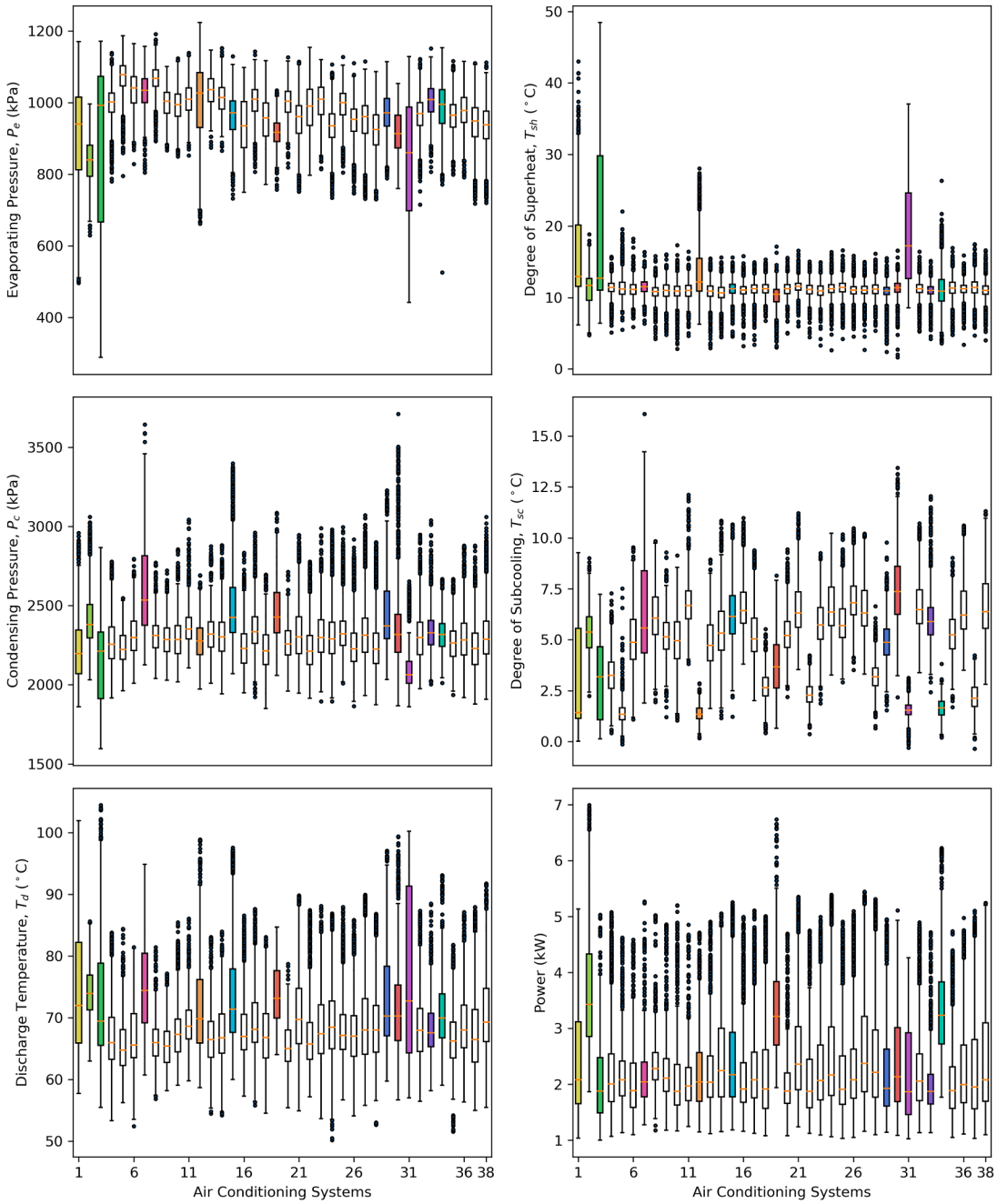


Fig. 7. The boxplots show the range of 6 physical parameters for 38 bus air conditioners in Shenzhen in August 2022. In each subplot, every vertical boxplot represents one system. The 12 faulty systems are marked with different colors.

quantify the magnitude of abnormality and then identify the faulty systems based on pre-defined thresholds. Third, the abnormal behavior of a system could be owing to its control strategy or extreme operating conditions, instead of having a fault. For example, the indoor fan could be manually switched to the low-speed level by the bus driver, causing the evaporating pressure to decrease; a condensing pressure much higher than its peers could be owing to the high outdoor temperature when the bus is parked under the sun at noon; high power consumption could also be caused by high cooing demand, etc.

The FDD method proposed in this paper can effectively solve these three issues. The GPR models build relationships between the features, environmental parameters, and control signals. The relationships should remain the same when the system is fault-free but will change significantly when a fault occurs. Hence, abnormally high or low values of the residuals (i.e., R_{pe}, R_{sh}, R_{pc}, R_{sc}, R_d, and R_w) can only be owing to faults. Also, 5 fault indexes are proposed to quantify the magnitude of different faults. The thresholds of these indexes are obtained from percentiles of statistical distributions.

3.2. Validation of the Gaussian process regression (GPR) models

In order to execute the FDD method in real-time and on a daily basis, this study proposes to use data of peer systems in the latest 14 days to establish the GPR models, and then calculate the residuals and fault indexes based on the data of the current day. For the purpose of illustration, this study uses data from August 1st to August 14th to develop the regression models, and then evaluates the 6 features of the 38 systems in each day from August 1st to August 24th. In total, 38 systems ×

6 features = 228 models are developed. In the last subsection we have already shown that on average about 90 data points can be obtained from one system in a single day. Therefore, within 14 days, the data points of a system could range from a few hundred to a few thousand, which is sufficient for model training. Generally speaking, a data-driven model such as GPR has higher robustness and generalizability with a large volume of training set. However, GPR is usually restricted to a few thousand data points because its computational cost increases quickly with the size of the training set. To improve the efficiency in training, the algorithm will randomly sample 1000 data points if the training data of a model is more than that. By this means, the total number of training data points will be strictly fewer than 38,000, which is only 47 % of the total volume of data after preprocessing. On a desktop computer, the total training time is 1.6 h with five-fold cross validation. In practice, the relatively short training time allows the models to be updated even every day. In fact, because the ambient conditions (e.g., outdoor temperature, radiation) change gradually with time, we recommend to update the model frequently.

Performance of the GPR models is assessed by mean absolute percentage error (MAPE) and coefficient of variation of root mean square error (CV(RMSE)) using five-fold cross validation, which are calculated as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{18}$$

$$CV(\text{RMSE}) = \frac{1}{\bar{y}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{19}$$

where N is the number of data, y_i are the actual values, $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the average of all actual values. MAPE and CV (RMSE) are both scale-independent metrics, allowing the modeling errors with different units (i.e., °C, kPa, and kW) being compared together. In the five-fold cross validation, the data to develop each individual model are split into five equally sized subsets. Then, the fitting and evaluation procedure is repeated for five times. In each time, a different subset is selected as the test set while the remaining four subsets serve as the training set. The performances of the five test sets are aggregated to a final score. Fig. 8 shows the cross-validation MAPE and CV(RMSE) of all regression models. About 98.7 % of models have an MAPE of lower than 5 %. To avoid overfitting, models exhibiting MAPE values exceeding 5 % are discarded.

Finally, after the regression models are developed, they are used to evaluate the performance of peer systems. Fig. 9 displays the residuals of the 6 features for the 38 bus air conditioners in all 24 days. The residuals are calculated by $y_i - \hat{y}_i$ in Eq. (11) without taking the average. Normal systems usually have all 6 residuals close to zero, while faulty systems have some residuals deviated from zero. The next two sections will detail the results of fault detection and diagnosis respectively.

3.3. Fault detection results

Fault detection is performed following the steps in Fig. 6. The index of abnormality (I_a) is calculated based on the 6 residuals each day for each of the 38 systems. Fig. 10 shows the result with the vertical axis displayed in a logarithmic scale. The two thin dotted lines are the thresholds to label a faulty system (calculated in Table 3). From this figure, the following conclusions can be drawn:

- The 12 faulty systems are marked with colors. For 11 of the 12 systems, the abnormality indexes (I_a) are well above the upper threshold (i.e., the 99.9th percentile) for at least two days in three consecutive days from the beginning of the fault durations. Thus, they are correctly labeled as faulty. There is only one exception (the dark purple line) whose I_a value is only above the lower threshold (i.e., the 99th percentile), but the system will still be diagnosed following the procedures. Some faulty systems even have I_a values up to around 10,000, much higher than the thresholds. A few faulty systems are repaired between August 13th and August 18th, and their I_a values decline significantly to much lower than the two thresholds.
- The systems without verified faults are shown in grey. None of them is labeled as faulty by this algorithm. But a few of them has I_a values above the 99th percentile threshold and will also be diagnosed later.

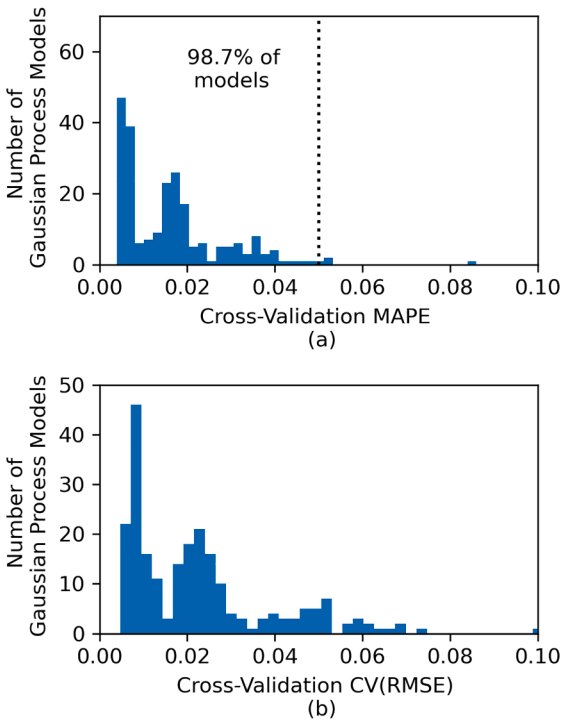


Fig. 8. Cross-validation mean absolute percentage error (MAPE) and coefficient of variation of root mean square error (CV(RMSE)) of all GPR models. 98.7 % of models have an MAPE of below 5 %. Models with MAPE above 5 % are discarded by the algorithm to avoid overfitting.

3.4. Fault diagnosis results

Continuing the procedures in Fig. 6, after fault detection is completed, fault diagnosis is conducted for systems with $I_a > 16.812$. These systems are diagnosed by the indexes of undercharge (I_{uc}), condenser fault (I_{cf}), evaporator fault (I_{ef}), and abnormally high power (I_w). The final results are displayed in Fig. 11 with the verified faulty systems shown in color in corresponding subplots and the other systems shown in grey. The results show that, 11 of 12 verified faulty systems can be correctly diagnosed. Only one system with undercharge cannot be identified, but the system has been labeled faulty in the fault detection process and will still be examined by technicians. Additionally, there are no false positives (i.e., fault-free systems erroneously labeled as faulty).

The above results of fault detection and diagnosis show the robustness of the method. Although almost one third of systems in this group of peers are abnormal, the method could still correctly identify the faulty systems and the corresponding type of fault through peer system comparison without labeled data. That is to say, the proposed FDD method is still valid even if one third of the bus air conditioners in a group of peers are faulty.

Additionally, in practice, once a system is verified as faulty, its GPR models can be considered useless and thus discarded. To illustrate this point, we also performed FDD for all 38 systems using only the GPR models of the 26 fault-free systems. We found that the FDD results are the same as the results shown in Figs. 10 and 11, i.e., no false positives

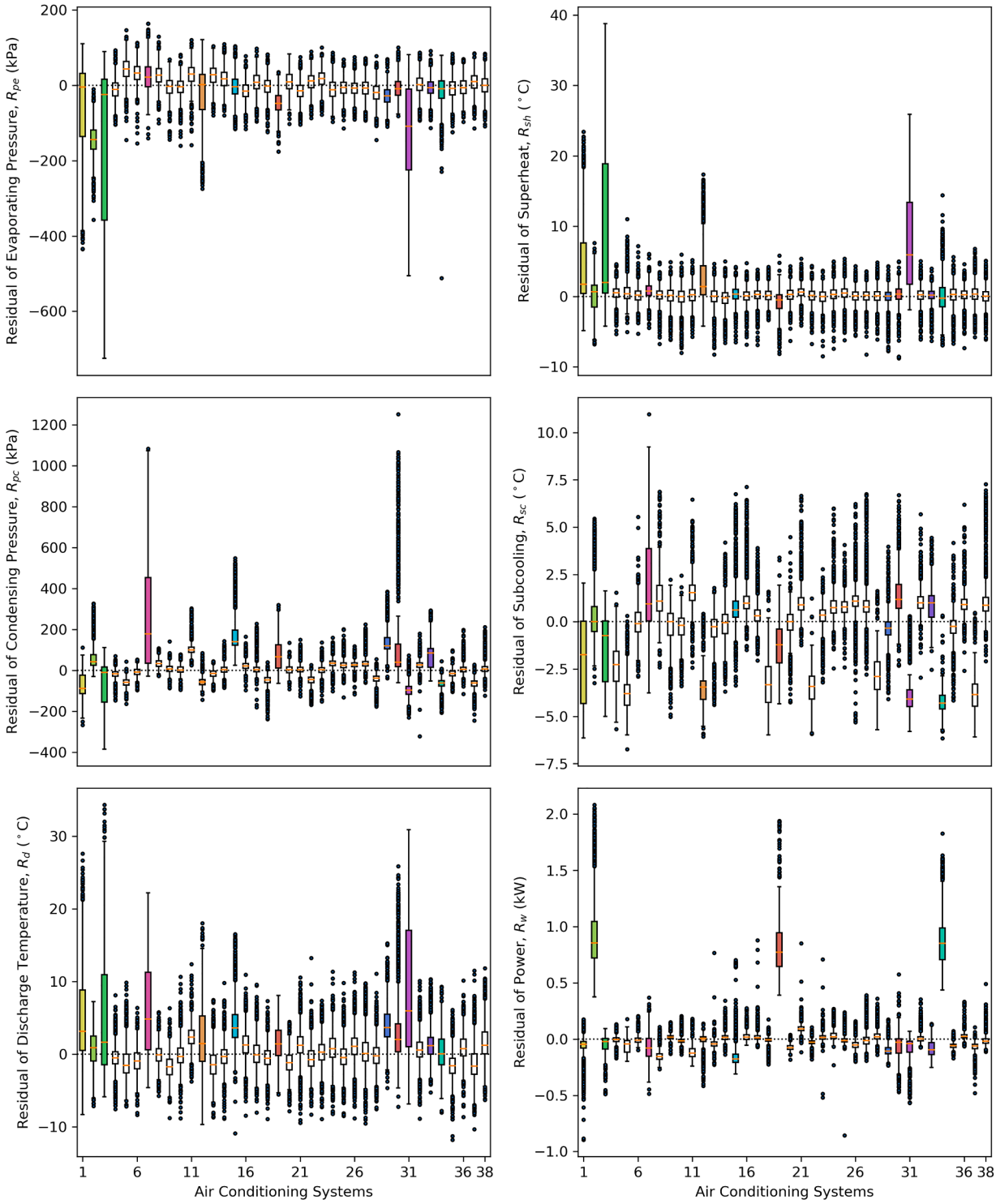


Fig. 9. The boxplots show the residuals of the 6 physical parameters for 38 bus air conditioners in Shenzhen in August 2022. In each subplot, every vertical boxplot represents one system. The 12 faulty systems are marked with different colors.

and only one false negative. In other words, no matter if the training data of the GPR models are all fault-free data, or the training data contains a fraction of faulty data (up to 1/3 of the training set), the method can correctly label faulty systems with low false positive/negative rates.

Finally, this paper shows two case studies below. The fault diagnosis results of the FDD method before and after the systems being repaired are compared.

3.4.1. *Case one: undercharge with active leakage*

Fig. 12 shows an air conditioning system with active refrigerant leakage, which was repaired on August 13th. Before repair, the transformed index of undercharge ($I_{uc}^{'}$) is higher than the other two transformed fault indexes all the time. Thus, the system is labeled with undercharge. In addition, the FDD method can also provide useful information of how this particular fault is labeled for repairing technicians. For example, this system is labeled undercharge due to low

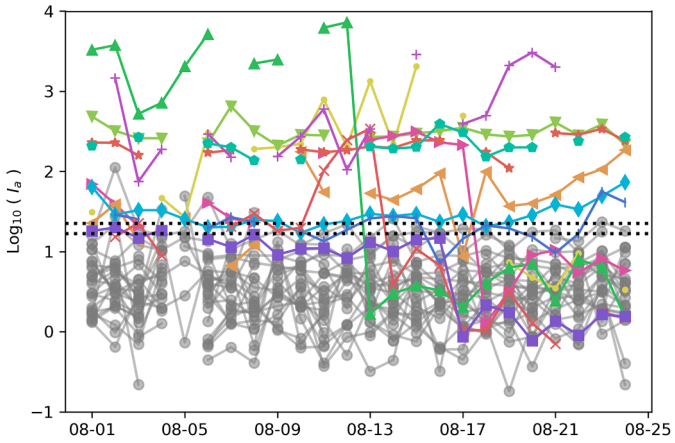


Fig. 10. Index of abnormality (I_a) of the 38 bus air conditioners in each day. The two dotted lines are the 99th and 99.9th percentiles of the χ^2 distribution with 6 degrees of freedom which are adopted as thresholds for fault detection. The 12 faulty systems are marked with different colors, while systems without verified faults are shown in grey. Some faulty systems have a significant decrease of I_a due to being repaired. Note that the vertical axis is in a logarithmic scale.

evaporator pressure, low condensing pressure, high superheat, low subcooling, and high discharge temperature, compared to the performance of other peer systems. Experienced repair technicians can understand the information and thus trust the results. Eventually, the technicians do not need to examine the system one more time by themselves, which saves the labor resources a lot. After repair, the fault indexes reduce significantly to around zero for I_{uc}, I_{ef}, and I_{cf}, which are far below the thresholds. The method will continue to monitor the system operating conditions each day.

3.4.2. Case two: condenser fault with outdoor fan problem

Fig. 13 shows a system with condenser fault because a fraction of the outdoor fans stopped working. Before repair, the index of condenser fault (I_{cf}) is much higher than the upper threshold; I_{uc} and I_{ef} are lower than zero and do not appear in the logarithmic scale. Thus, the system is labeled condenser fault, which is usually caused by outdoor fan problems or condenser fouling. Although the detail is still unknown, the repair technicians have obtained useful information and can focus on checking the outdoor fans and condenser coils. The system was repaired on August 18th by replacing the broken outdoor fans with new ones. After repair, I_{cf} drops from 195 to 0.2, much lower than the 99th percentile which equals 7.068; I_{uc} remains around 5, lower than the 99th percentile which equals 10.190; I_{ef} is still negative. The results show that the FDD method is able to correctly diagnose a faulty system and also detect the repair instantaneously.

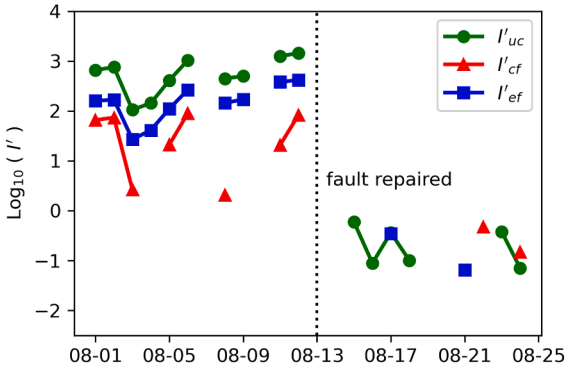


Fig. 12. Case one: the system is correctly identified with refrigerant undercharge, and has been verified with leakage by the service company. The fault was repaired on August 13th. Note that negative values of the transformed fault indexes after repair are not shown because the y-axis is in the logarithmic scale.

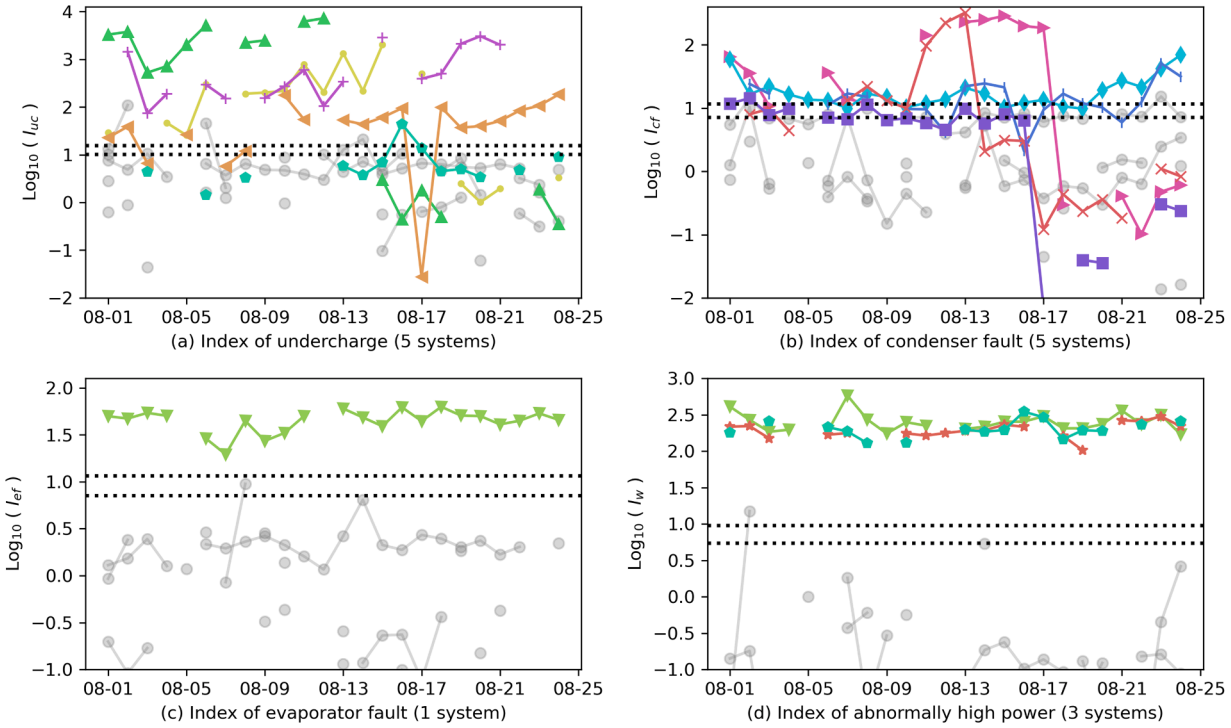


Fig. 11. Index of undercharge, condenser fault, evaporator fault, and abnormally high power (I_{uc},I_{cf}, I_{ef}, I_w) for the 12 faulty systems (in color) and systems without verified faults (in grey). The two dotted lines are the 99th and 99.9th percentiles calculated in Table 3 which are adopted as thresholds for fault diagnosis. The vertical axes are in logarithmic scales.

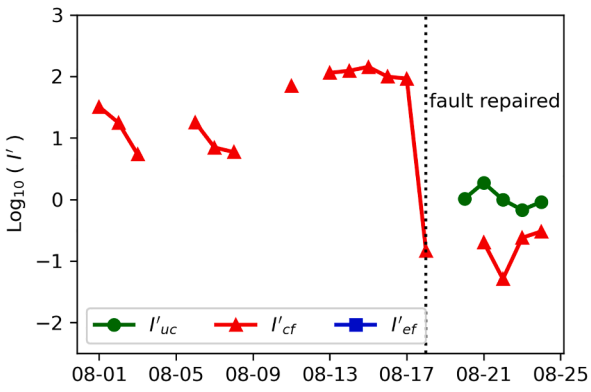


Fig. 13. Case two: the system is correctly identified with condenser fault, and has been verified with outdoor fan malfunctioning by the service company. The fault was repaired on August 18th. Note that negative values of the transformed fault indexes are not shown because the y-axis is in the logarithmic scale.

4. Conclusion

This paper designed an FDD method specially for electric vehicle air conditioners combining both artificial intelligence and domain expertise. The method identifies faults based on comparison of physical parameters between a group of peer systems which can be simulated using the same type of model and working at similar conditions.

The method extracts 6 influential features from all available measurements of the refrigeration cycle, and develops GPR models for each system to correlate the relationships between every feature and its influential parameters. Then, the performance of each system is evaluated by the regression models from its comparable peer systems. If the measurement feature values of a system deviate significantly from its comparable peer systems' model predictions, then the performance of the system is probably abnormal and requires further examination. To further diagnose the abnormal systems, a few fault indexes are proposed based on the rules verified by previous experiments.

The method has been applied to 38 field-operating electric bus air conditioners in Shenzhen, China, and is able to correctly identify 11 of the 12 verified faulty systems. The method has a few advantages. First, collection of fault-free data is not needed and detection of steady-state conditions is not required. Hence, the process of data collection is relatively easy. Second, with the application of the Gaussian process, the method is highly robust under rapidly changing operation conditions and can effectively identify different types of faults in real-time as long as the majority of peer systems are fault-free. Third, the GPR models can be updated on a daily basis, such that FDD can be performed every day for each system using the latest data. Every system is actively monitored and operational faults can be identified promptly. Fourth, the detection results are interpretable to repairing technicians, helping them quickly locate the faults and repair the systems.

Future work can focus on uncertainty and sensitivity analysis. The uncertainty of model predictions and the effect of existing faulty systems on the FDD results can be further analyzed to improve the overall performance of the proposed FDD method.

CRediT authorship contribution statement

Fangzhou Guo: Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Zhijie Chen:** Validation, Methodology, Data curation, Conceptualization. **Fu Xiao:** Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors gratefully acknowledge the support of this research by the Research Talent Hub for ITF Project (ITP/002/22LP) sponsored by Hong Kong Innovation and Technology Fund and the Research Grants Council of the Hong Kong SAR (C5018-20GF).

References

[1] Lajunen A. Energy consumption and cost-benefit analysis of hybrid and electric city buses. Transp Res Part C: Emerg Technol 2014;38:1–15. https://doi.org/10.1016/j.trc.2013.10.008.
[2] García A, Monsalve-Serrano J, Lago Sari R, Tripathi S. Life cycle CO_2 footprint reduction comparison of hybrid and electric buses for bus transit networks. Appl Energy 2022;308:118354. https://doi.org/10.1016/j.apenergy.2021.118354.
[3] Alimujiang A, Jiang P. Synergy and co-benefits of reducing CO2 and air pollutant emissions by promoting electric vehicles—a case of Shanghai. Energy Sustain Dev 2020;55:181–9. https://doi.org/10.1016/j.esd.2020.02.005.
[4] Lu F, Hua G. A location-sizing model for electric vehicle charging station deployment based on queuing theory. In: 2015 international conference on logistics, informatics and service sciences (LISS). Barcelona, Spain: IEEE; 2015. p. 1–5. https://doi.org/10.1109/LISS.2015.7369769.
[5] Uslu T, Kaya O. Location and capacity decisions for electric bus charging stations considering waiting times. Transp Res Part D: Transp Environ 2021;90:102645. https://doi.org/10.1016/j.trd.2020.102645.
[6] Zhu C, Chen X. Optimizing battery electric bus transit vehicle scheduling with battery exchanging: model and case study. Procedia - Soc Behav Sci 2013;96:2725–36. https://doi.org/10.1016/j.sbspro.2013.08.306.
[7] Wang Y, Huang Y, Xu J, Barclay N. Optimal recharging scheduling for urban electric buses: a case study in Davis. Transp Res Part E: Logist Transp Rev 2017;100:115–32. https://doi.org/10.1016/j.tre.2017.01.001.
[8] Bhatt P, Mehar H, Sahajwani M. Electrical motors for electric vehicle – a comparative study. SSRN J 2019. https://doi.org/10.2139/ssrn.3364887.
[9] Wang J, Liu K, Yamamoto T, Morikawa T. Improving estimation accuracy for electric vehicle energy consumption considering the effects of ambient temperature. Energy Procedia 2017;105:2904–9. https://doi.org/10.1016/j.egypro.2017.03.655.
[10] Killeen P, Ding B, Kiringa I, Yeap T. IoT-based predictive maintenance for fleet management. Procedia Comput Sci 2019;151:607–13. https://doi.org/10.1016/j.procs.2019.04.184.
[11] Massaro A, Selicato S, Galiano A. Predictive maintenance of bus fleet by intelligent smart electronic board implementing artificial intelligence. IoT 2020;1:180–97. https://doi.org/10.3390/iot1020012.
[12] Katipamula S, Brambley M. Review article: methods for fault detection, diagnostics, and prognostics for building systems—a review, part I. HVAC&R Res 2005;11:3–25. https://doi.org/10.1080/10789669.2005.10391123.
[13] Kim M, Kim MS. Performance investigation of a variable speed vapor compression system for fault detection and diagnosis. Int J Refrig 2005;28:481–8. https://doi.org/10.1016/j.ijrefrig.2004.11.008.
[14] Li H, Braun JE. Development, evaluation, and demonstration of a virtual refrigerant charge sensor. HVAC&R Res 2009;15:117–36. https://doi.org/10.1080/10789669.2009.10390828.
[15] Kim W, Braun JE. Development and evaluation of virtual refrigerant mass flow sensors for fault detection and diagnostics. Int J Refrig 2016;63:184–98. https://doi.org/10.1016/j.ijrefrig.2015.11.005.
[16] Cheung H, Braun JE. Virtual power consumption and cooling capacity virtual sensors for rooftop units. In: International refrigeration and air conditioning conference; 2014.
[17] Kim W, Katipamula S. A review of fault detection and diagnostics methods for building systems. Sci Technol Built Environ 2018;24:3–21. https://doi.org/10.1080/23744731.2017.1318008.
[18] 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. https://doi.org/10.1016/j.rser.2019.04.021.
[19] Tien PW, Wei S, Darkwa J, Wood C, Calautit JK. Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality – a review. Energy AI 2022;10:100198. https://doi.org/10.1016/j.egyai.2022.100198.

[20] Zhang F, Saeed N, Sadeghian P. Deep learning in fault detection and diagnosis of building HVAC systems: a systematic review with meta analysis. Energy AI 2023; 12:100235. https://doi.org/10.1016/j.egyai.2023.100235.
[21] Wang L, Braun J, Dahal S. An evolving learning-based fault detection and diagnosis method: case study for a passive chilled beam system. Energy 2023;265:126337. https://doi.org/10.1016/j.energy.2022.126337.
[22] Li G, Chen L, Liu J, Fang X. Comparative study on deep transfer learning strategies for cross-system and cross-operation-condition building energy systems fault diagnosis. Energy 2023;263:125943. https://doi.org/10.1016/j.energy.2022.125943.
[23] Fan C, Lei Y, Sun Y, Mo L. Novel transformer-based self-supervised learning methods for improved HVAC fault diagnosis performance with limited labeled data. Energy 2023;278:127972. https://doi.org/10.1016/j.energy.2023.127972.
[24] Eom YH, Yoo JW, Hong SB, Kim MS. Refrigerant charge fault detection method of air source heat pump system using convolutional neural network for energy saving. Energy 2019;187:115877. https://doi.org/10.1016/j.energy.2019.115877.
[25] Guo F, Rasmussen B. Predictive maintenance for residential air conditioning systems with smart thermostat data using modified Mann-Kendall tests. Appl Therm Eng 2023;222:119955. https://doi.org/10.1016/j.applthermaleng.2022.119955.
[26] Zhong F, Calautit JK, Wu Y. Fault data seasonal imbalance and insufficiency impacts on data-driven heating, ventilation and air-conditioning fault detection and diagnosis performances for energy-efficient building operations. Energy 2023; 282:128180. https://doi.org/10.1016/j.energy.2023.128180.
[27] Bode G, Thul S, Baranski M, Müller D. Real-world application of machine-learning-based fault detection trained with experimental data. Energy 2020;198:117323. https://doi.org/10.1016/j.energy.2020.117323.
[28] Rogers AP, Guo F, Rasmussen BP. A review of fault detection and diagnosis methods for residential air conditioning systems. Build Environ 2019;161:106236. https://doi.org/10.1016/j.buildenv.2019.106236.
[29] Guo F, Chen Z, Xiao F, Li A, Shi J. Real-time energy performance benchmarking of electric vehicle air conditioning systems using adaptive neural network and Gaussian process regression. Appl Therm Eng 2023;222:119931. https://doi.org/10.1016/j.applthermaleng.2022.119931.
[30] Chen Z, Guo F, Xiao F, Jin X, Shi J, He W. Development of data-driven performance benchmarking methodology for a large number of bus air conditioners. Int J Refrig 2022. https://doi.org/10.1016/j.ijrefrig.2022.12.027. S014070072200487X.
[31] Chen B, Braun JE. Simple rule-based methods for fault detection and diagnostics applied to packaged air conditioners/Discussion. ASHRAE Trans 2001;107:847.
[32] Rossi T, Braun J. A statistical, rule-based fault detection and diagnostic method for vapor compression air conditioners. HVAC&R Res 1997;3:19–37. https://doi.org/10.1080/10789669.1997.10391359.
[33] Breuker M, Braun J. Common faults and their impacts for rooftop air conditioners. HVAC&R Res 1998;4:303–18. https://doi.org/10.1080/10789669.1998.10391406.
[34] Hu Y, Yuill DP, Ebrahimifakhar A, Rooholghodos A. An experimental study of the behavior of a high efficiency residential heat pump in cooling mode with common installation faults imposed. Appl Therm Eng 2021;184:116116. https://doi.org/10.1016/j.applthermaleng.2020.116116.