

EARLY DETECTION AND RECONSTRUCTION OF ABNORMAL DATA USING HYBRID VAE-LSTM FRAMEWORK

Fangli HOU

The Hong Kong University of Science and Technology

Jun MA

The University of Hong Kong

Jack C P CHENG & Helen H L KWOK*

The Hong Kong University of Science and Technology

ABSTRACT: *Early failure detection and abnormal data reconstruction in sensor data provided by building ventilation control systems are critical for public health. Early detection of abnormal data can help prevent failures in crucial components of ventilation systems, which can result in a variety of issues, from energy wastage to catastrophic outcomes. However, conventional fault detection models ignore valuable features of dynamic fluctuations in indoor air quality (IAQ) measurements and early warning signals of faulty sensor data. This study introduces a hybrid framework for early failure detection and abnormal data reconstruction applying variance analysis and variational autoencoders (VAE) coupled with the long short-term memory network (VAE-LSTM). The periodicity and stable fluctuation of IAQ data are exploited by variance analysis to detect unusual variations before failure occurs. The IAQ dataset which is corrupted by introducing complete failure, bias failure and precision degradation fault is then used to verify the feasibility of the VAE-LSTM model. The results of variance analysis reveal that unusual behavior of the data can be detected as early as 12 hours before failure occurs. The reconstruction performance of the developed method is shown to be superior to other methods under different abnormal data scenarios.*

KEYWORDS: *Early failure detection, Abnormal data reconstruction, Variational autoencoder (VAE), Long short-term memory network (LSTM), Sustainable IAQ management*

1. INTRODUCTION

Indoor air quality (IAQ) in public buildings is regarded as a hot study topic as it has a big impact on human health. Recent research has shown a connection between indoor air pollutants, including CO₂, with health effects and academic performance (Szabados et al., 2022). According to the EPA's IAQ tools for schools (EPA, 2009), CO₂ concentrations in schools should adhere to the ASHRAE standard 62-2001 limit of 700 ppm over the outdoor concentration (just above 1000 ppm overall) for CO₂ concentrations. Besides various laws and regulations, there is a need for continuous monitoring of IAQ, which includes the installation of sensors to detect anomalous events that may have a detrimental impact on the IAQ. Sensors are generally placed on walls or ceilings to collect hourly levels of pollutants, such as CO₂, NO₂, and particulate matter (PM), which are small and aerodynamic. In addition, sensors can also collect relative humidity and temperature data. These monitoring sensors are important in the management of ventilation systems. Unfortunately, hardware sensors can encounter various issues, such as bias and precision degradation. In addition, they may experience data loss due to environmental or operability issues, which results in their measurements being unrealistic (Kim, Liu, Kim, & Yoo, 2014). When air quality is not monitored properly, it can lead to a decrease in IAQ levels. On the other hand, overestimation of the levels of pollutants can cause energy wastage. For these reasons, an effective method for early failure detection and reconstruction of faulty IAQ sensors can help increase the uptime of ventilation management.

Some investigations have used statistical methods for abnormal data reconstruction (Kasam, Lee, & Paredis, 2014; Ouyang, Zha, & Qin, 2017). Although statistical methods are easier to implement and work well when there are few abnormal data, their performance is constrained as the data complexity increases. Additionally, the majority of statistical techniques rely on linear assumptions, which are incompatible with nonlinear real-world situations. Traditional machine learning approaches can use the whole data set to understand the patterns of failure performance in order to solve this problem. Unfortunately, they require a lot of manually classified anomalies to learn a predictor from given observations (Wang, Feng, & Liu, 2021), and due to the failure of unanticipated patterns of learning, such methods have poor performance (Bu et al., 2018). The emergence of neural methods without labelled information that is capable of handling non-linear data is a major factor that has led to the increasing number of applications of deep learning in process monitoring.

Time series prediction using deep learning methods, especially the long short-term memory neural network (LSTM), has achieved significant achievements in recent years (X. Li et al., 2017; Qing & Niu, 2018; Su & Kuo, 2019). A hybrid convolutional neural network and long short-term memory model (CNN-LSTM) was used to impute missing values in time-series datasets for air-conditioning appliances (Hussain et al., 2022). The hybrid technique outperformed the CNN and LSTM variants in terms of performance. Ma et al. suggested a hybrid Bi-directional Imputation method using an LSTM model and Transfer Learning to fill the gaps in the energy consumption data (Ma et al., 2020). Transfer learning was utilized to prevent network saturation problems while the basic model was pre-trained on data from a comparable building. The performance demonstrated that the suggested architecture could successfully handle various scenarios with missing data, including continuous and random missing data. The developed strategy, however, was predicated on the prior assumption of source and target data collected on sufficiently similar buildings. In general, LSTMs are commonly used in a wide range of applications due to their ability to model non-linear dependencies. However, the prediction performance of LSTM can also be sensitive to the anomalies in the input due to its non-linear nature. Our proposed approach is to make sure that the input of the LSTM prediction network contains as little abnormal data as possible. Therefore, a method for abnormal detection and reconstruction in time series is necessary. However, existing abnormal detection methods require a lot of manually labelled abnormal observations. Some methods, which are based on unsupervised learning, are used for time-series abnormal detection to address these concerns by concentrating on normal patterns rather than anomalies (Breunig, Kriegel, Ng, & Sander, 2000; Cao, Nicolau, & McDermott, 2016; Erfani, Rajasegarar, Karunasekera, & Leckie, 2016). Unfortunately, due to the failure of unanticipated pattern learning, such discriminative modeling-based techniques still necessitate a significant number of normal observations and have low accuracy (Bu et al., 2018).

Traditional fault detection methods based on supervised learning require sufficient training data (D. Li, Zhou, Hu, & Spanos, 2016; Zhao, Li, Zhang, & Zhang, 2019). However, the amount of data is usually insufficient in reality, because it is difficult to get high-quality training data sets for each type of failure. Yan et al. proposed a semi-supervised fault detection method, which only uses a small amount of data to detect the failure of the air-conditioning unit (Yan, Zhong, Ji, & Huang, 2018). However, it is limited only when the same failure occurs again.

Recently, there has been a rise in deep generative modeling techniques that can be used for detecting anomalies. Autoencoder (AE) is a powerful deep learning technique that is appropriate for failure diagnosis with limited fault data since it can learn data features, avoiding the dependence on failure data (Zhang, Jiang, Zhan, & Yang, 2019). In addition, AE is a crucial tool of non-linear process monitoring as it can handle the encoding of input data and the extraction of features to provide meaningful representations of data in various applications, such as failure detection and data reconstruction. Variational autoencoder (VAE) technology has been shown to have benefits over conventional AE architecture. Both VAE and AE architectures can compress data from high-dimensional space to low-dimensional space (also known as latent space) and reconstruct complicated data. The main difference between VAE and regular AE architectures is that the former has a continuous latent space, allowing it to learn the distribution of data and reconstruct new information, which is crucial for process monitoring. However, as VAE is not a sequential model and cannot handle long-term dependencies in time series, it is possible to combine a sequential modeling approach such as LSTM models with VAE to solve this issue. Lin et al. proposed a hybrid VAE-LSTM model which can detect anomalies on multiple time scales (Lin et al., 2020). The VAE module forms local features on brief windows, while the LSTM module estimates the sequence's long-term correlation. However, if there is no abnormal data in the dataset, the hybrid VAE-LSTM model is not suitable as a means of prediction, as it will increase the computational complexity and cost. Thus, it is possible to make both the VAE-LSTM and the LSTM alone train independently and be exchanged if needed.

To effectively address the issue of sensor faults, a comprehensive framework with early fault detection and reconstruction techniques is required. However, the combination of fault data reconstruction and early failure detection is rarely reported. Previous publications have demonstrated that early failure detection has a variety of applications, including analysis of climate pattern change (Drake & Griffen, 2010; Rogers et al., 2018), credit risk diagnosis (Ali & Dağtekin, 2008; Lu, Shen, & Wei, 2013), and early failure detection of key system components (Lee, House, Park, & Kelly, 1996; Yu, Woradechjumboon, & Yu, 2014). As introduced earlier, the increasing popularity of VAE in fault detection also makes it a new approach in early fault detection. The ball screw degradation assessment method used in (Wen & Gao, 2018) is similar to the one used in the manufacturing industry. The assessment shows that the deterioration of a ball screw can be evaluated using the Variational Autoencoder Reconstruction Error (VAERE). Malfunctions in an air handling unit (AHU) were studied in (Mesa-Jiménez, Stokes, Yang, & Livina, 2021), in which the VAERE was used to reproduce the sudden change of temperature before the fault occurred. This is because VAE can model the underlying probability distribution of the input, especially when processing a time sequence with a typical periodic pattern. In the case of failure, the periodic

characteristics of the time sequence will be destroyed. Therefore, the reconstruction error of the VAE can be used to observe the unusual behavior in the time series data. However, there are still issues that the existing literature does not address. Firstly, some studies adopt a reactive approach rather than a proactive approach. When a system fails, it often leads to service interruptions and necessitates engineers to temporarily shut down certain pieces of equipment in order to remedy the problem. Secondly, the probability of misjudgment of the failure diagnosis by a single indicator is high, and it is more scientifically correct to use multiple indicators for early warning. Therefore, it is necessary to develop an active multi-index method for early failure detection.

To solve the problem of sensor faults including fault detection and reconstruction, a sustainable and real-time IAQ monitoring framework is proposed, which mainly focuses on early failure detection, failure data reconstruction, and assessment of the impact of failure data on ventilation performance. Early fault detection mainly utilizes the periodic characteristics of IAQ time series data and conducts variance analysis on reconstruction errors to monitor early warning failure signals. The reconstruction model combines the VAE architecture and Long Short-Term Memory neural network (LSTM). The purpose of integrating the two structures in this study is to extract data features according to the dynamic characteristics and nonlinear dependence of IAQ data, so as to reconstruct abnormal data. The contributions of this study are described in detail as follows:

- A proactive early failure detection method is proposed for IAQ time series data. Taking advantage of the periodic and stationary fluctuation characteristics of IAQ data under normal operating conditions, the unstable behavior of the raw data before the failure is reproduced using variance analysis. The variance analysis is applied to the reconstruction error of VAE to check the fluctuation of IAQ data indicating where the failure has already occurred. Therefore, the engineers can find the potential failure and carry out maintenance, when necessary, before these failures actually happen.
- When an anomaly is detected, the reconstructed data using VAE-LSTM replaces the abnormal data. The restored data is then fed into the LSTM neural network to forecast the time series. Thus, both the hybrid VAE-LSTM and the LSTM may be learnt independently and replaced as needed. The VAE-LSTM is developed by using the normal IAQ measurement data. Given that IAQ often exhibits changing patterns over time, the time variable, Hour, is translated into one-hot encoders as conditional information. For example, the IAQ in a restaurant typically present dramatic differences during meal hours and non-meal hours, and the time variable Hour can be used to provide additional conditional information. Therefore, Hour, which can be written by one-hot encoding vectors, is supplied as an input to both the encoder and decoder to provide additional controls over the process of data generation.
- To verify the superiority of the proposed method over other neural approaches, different types of abnormal data are presented in the test dataset: the IAQ dataset is corrupted by introducing complete failure, bias failure and a precision degradation fault.

The rest of the work is organized as follows: Section 2 provides the dataset used, the method description, steps in network training and explanations of the validation performance analysis. Section 3 compares the performance of the proposed method to other methods. Section 4 discusses the conclusions and limitations.

2. MATERIAL AND METHODS

In this section, a framework for early failure detection and fault data reconstruction is designed based on VAE, as illustrated in Fig.1. Firstly, the variance analysis is applied to find the abnormal fluctuations of IAQ data before the failure occurs. Once the abnormal signal is detected, the proposed VAE-LSTM hybrid model is applied to reconstruct the abnormal data. To verify the superiority of the proposed reconstruction model, various scenarios of abnormal data are introduced into the test dataset. The remainder of this section illustrates the detailed procedures.

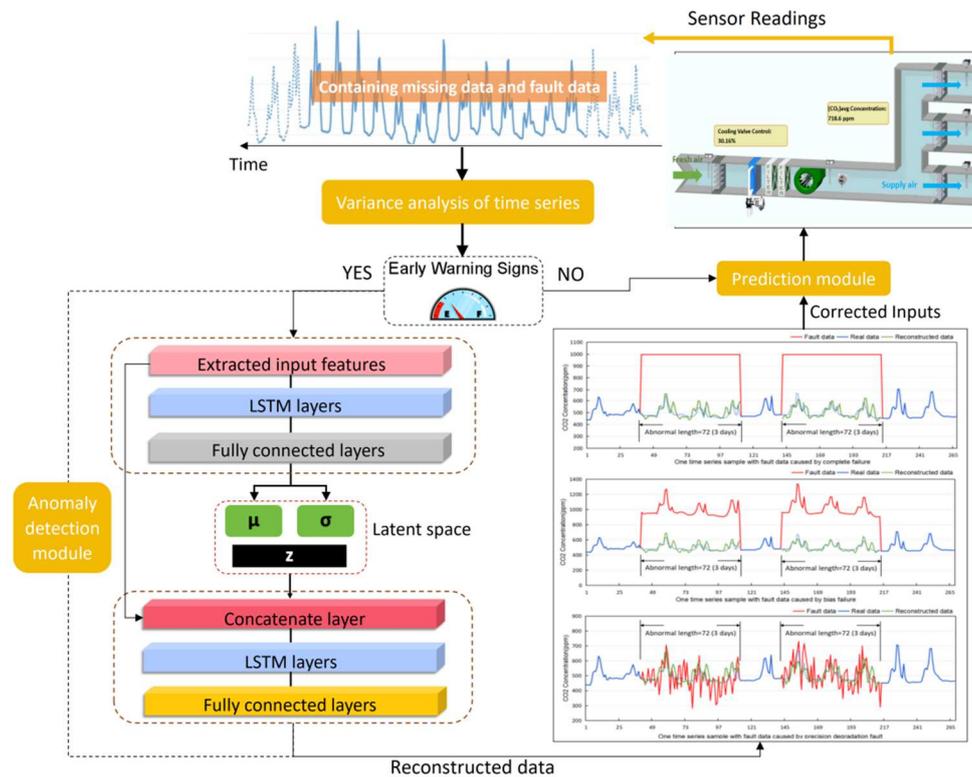


Fig. 1: Framework of this study

2.1 Data collection

The Facilities and Management Office of the Hong Kong University of Science and Technology (HKUST) provided the data used in this study. The dataset recorded IAQ data of various types of campus buildings, including canteens, library buildings, lab buildings, etc. Among them, the canteen exhibits large IAQ oscillations caused by obvious variations of pedestrian flow. Furthermore, during the peak dining hours of the canteen, pedestrian flow increases significantly, and indoor pollutants like CO₂ concentration can sometimes exceed the standard indoor concentration of 1000 ppm. This necessitates more precise ventilation management system control, and high-quality CO₂ sensor data are required for achieving this control. Therefore, we chose the CO₂ concentration of the canteen as an example to test the proposed methodology. The chosen time period included holidays, non-holidays, and the final exam period, which brings certain challenges for data analysis. In holidays, the behavior of occupants will be different from regular days, and the number of people during peak hours will be significantly reduced. These variations will affect the data patterns of indoor pollutants such as CO₂. Typical temporal models are hard to adapt, resulting in error-prone predictions. Therefore, external features need to be added to provide additional clues to the temporal model to maintain high prediction accuracy when dealing with changes in holidays and examination periods.

2.2 Early failure detection

In order to analyze the early warning signals of the sensor data generated by the ventilation management system and give the engineering maintenance personnel sufficient time to repair the failure, a fault detection technique, i.e., variance analysis, is applied to the reconstruction error of the VAE-LSTM model. The early warning indicator is applied to the time series with failures through a selected sliding window. The choice of sliding window length is a compromise between the time resolution and the clarity of transitional signal changes.

The variational autoencoder (VAE) is an algorithm for stochastic variational inference and learning using neural networks as the recognition model (Kingma & Welling, 2013). The reconstruction error of VAE can be calculated for abnormal detection. The idea underlying abnormal detection is that the VAE is not able to reconstruct unpredictable patterns or noise as well as it can regular data. Therefore, when x_i in a given time series i is reconstructed by VAE, the error between the output \hat{x}_i and input of abnormal data is significantly larger. Variance is used to measure the degree of fluctuation of a set of data. Variance analysis is very straightforward to use and does not require specialized knowledge because it is a simple failure detection approach. The goal of this study is to integrate the reconstruction error based on VAE model with variance analysis to identify out-of-law abnormal

fluctuations in IAQ data in advance, described as:

$$e_i = x_i - \hat{x}_i \quad (1)$$

$$\sigma_v^2 = \frac{\sum(e_i - \bar{e})^2}{n - 1} \quad (2)$$

where n is the number of observations in a sample, σ_v^2 is the sample variance, e_i is the reconstruction error for each input, \bar{e} is the mean value of all observations in the sample, and x_i and \hat{x}_i are the actual and the reconstructed output, respectively. Therefore, the VAE indicator is derived from the reconstruction error, which is referred to as the variational autoencoder reconstruction error (VAERE).

2.3 Model development

For faulty data reconstruction and missing data imputation in ventilation control systems, a technique that can effectively handle complicated and failure data is required. This work benefits from combining the representation learning capabilities of deep generative models—in the form of variational autoencoders (VAEs)—with the temporal modeling capabilities of long short-term memory (LSTMs) to manage long-term time sequence data and generate accurate data based on intrinsic distributions. To train the proposed VAE-LSTM model without supervision, the dataset needs to be divided into a training set and a test set, with a continuous segment containing no anomalies serving as the training data and the remaining time series containing anomalies used for evaluation in the test set. Fig. 2 illustrates the architecture of the VAE-LSTM model.

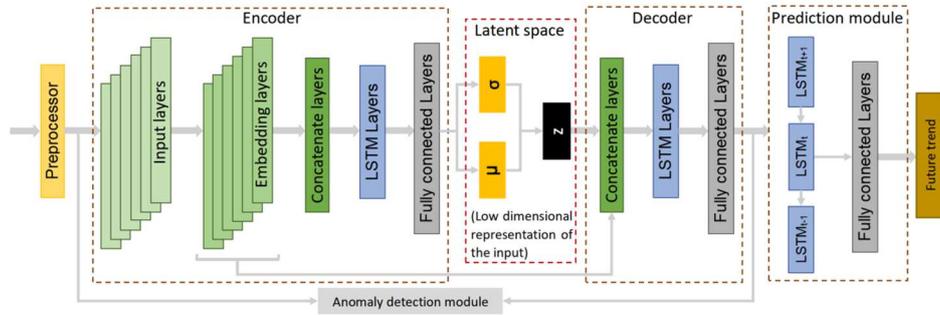


Fig. 2: The proposed VAE-LSTM architecture

The design process of the proposed model is as follows. Firstly, collect the IAQ training dataset without faulty and missing readings, and we introduce some fault data and missing intervals (with a fixed length and proportion) on the dataset to validate the model performance. In addition to sequences with faulty or missing data as the main input of the VAE-LSTM model, sequences of integers that encode the additional information provided as categorical features, such as month, weekday, hour, and holiday, serve as the second meaningful inputs. Due to the characteristics of buildings with significantly varying occupancy patterns at different hours, and to make the predictive model more concise, the variable Hour is taken as additional information for IAQ sequence reconstruction. The original IAQ data and the categorical feature which is transformed by the embedding operation are concatenated into the LSTM layer to capture the relationship between temporal features. The ReLU is selected as the activation function of the LSTM layer. The output of the LSTM goes through a dense layer with a non-linear activation function. It then generates a 2D output, just like every other encoder in a VAE architecture, which is used to approximate the mean and variance of the latent distribution. The decoder takes samples from the 2D latent distribution upsampling and then concatenates the generated sequence with the original categorical embedding sequence to provide more control over reconstructing the original IAQ sequence. LSTMs and dense layers with ReLU activations constitute the rest of the decoder structure. The training of VAE-LSTM adopts the early stopping training mechanism to minimize the combination of reconstruction loss and distribution loss. The patience was set as 10. Specifically, the training process will end if the model loss does not decrease after 10 iterations. Adam was chosen as the optimizer as it provides the best convergence (Kingma & Ba, 2014). The hyperparameters were chosen by fine-tuning the VAE-LSTM structure. The best hyperparameters were selected based on their performance in fault data reconstruction and missing data imputation.

The reconstructed sequence is utilised for time series prediction by the LSTM neural network after the abnormal data is replaced by the VAE-LSTM output. The prediction module consists of one layer of LSTM and one dense layer. Grid search is used to optimise the model architecture and hyperparameters. Input and output temporal dimensions are the same. Mean squared error (MSE) is used as the loss function throughout the training process, which was performed with 500 epochs, and the Adam optimizer with a learning rate of 0.001.

2.4 Validation scenarios

Since anomalous events are rare, it is usually not feasible to collect sufficient abnormal data for detailed characterization. Therefore, to assess the effectiveness of the proposed framework, we designed different types of abnormal scenarios with fixed lengths and proportions on the test set, with the key benefit of creating an arbitrary amount of abnormal data while using the original data as a ground truth. The downside of this procedure is that it may overfit the abnormal data or provide worse results for real anomalies. As a result, only real data is utilized to train the model, whereas anomalous data is solely used to evaluate it.

A detailed explanation of the abnormal scenarios is as follows. Anomalies such as gain and offset of sensor signals may arise due to incorrect calibration or mechanical wear over a period of time. We attempt to simulate three types of typical sensor faults: 1. Complete failure: the size is assumed to be twice the average concentration of the original data; 2. Bias failure: the size is assumed to be twice the original faulty data segment; and 3. Precision degradation fault: the size in the temporal dimension is taken as the average and standard deviation of the original data.

3. RESULTS AND DISCUSSION

We now employ the methodology described in Section 2 for the collected sensor data for analysis. Table 1 presents a statistical summary of the data in this study. Real faulty data are utilised to evaluate the effectiveness of variance analysis in early failure detection. Abnormal data scenarios are then introduced to evaluate the reconstruction and imputation performance of the proposed method against other approaches.

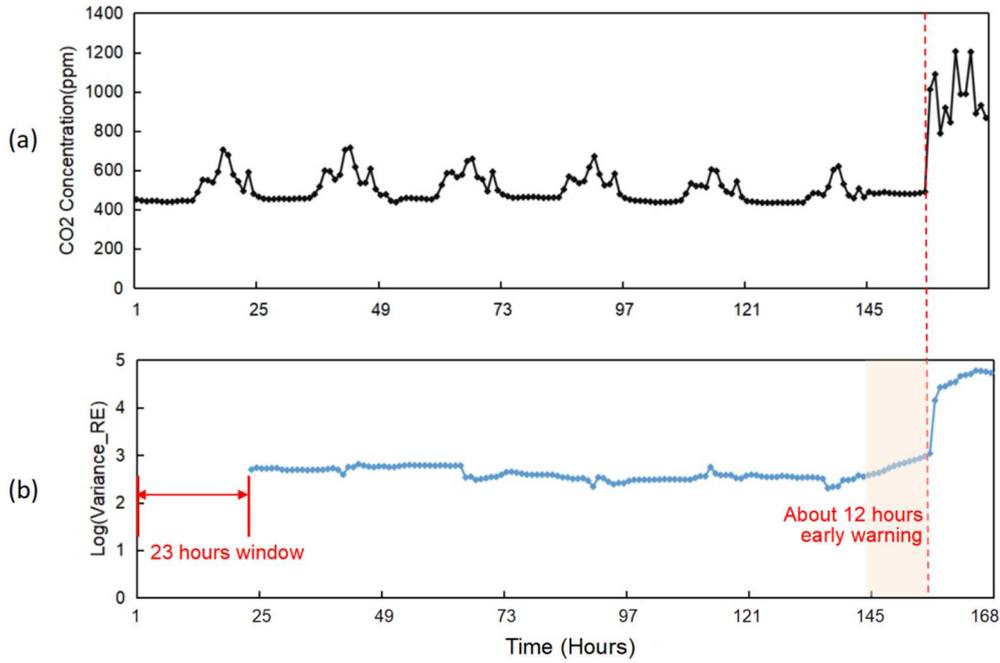
Table 1: The basic statistics of variables.

Attribute	Content
Variable	CO ₂ concentration
Time period	From 2021/11/08 13:00 to 2022/02/19 22:00
Unit	ppm
Resolution	Hour
Mean	508.56
Minimum	412
Maximum	844.7
Standard Deviation	77.67

3.1 Early failure detection analysis in IAQ measurements

We applied variance analysis to analyze the failure of the CO₂ sensor in the indoor ventilation control system, which resulted in abnormal changes in CO₂ concentration up to 1000 ppm instantaneously. Therefore, the purpose of applying variance analysis in this study is to detect this anomaly before it occurs. Figure 3(a) shows one week of CO₂ data containing the failures, with an abnormally high CO₂ concentration.

The analysis results are shown in Figure 3, where the collected CO₂ data is presented together with the analysis results. For convenience, the variance on the Y-axis is represented on a logarithmic scale. We used different windows to obtain early failure signals, with a 14-hour window when applied to CO₂ data and a 23-hour window when applied to the reconstruction error. The choice of window size is based on the clarity of the provided signal. It is evident from the figure that the variance results for the normal data are periodic, while unexpected fluctuation patterns appear before the failure. When the variance is applied to the reconstruction error as shown in Figure 3(b), the failure signal is generated about 12 hours prior to the failure, and the reconstruction error gradually increases, which shows the unexpected fluctuation pattern before the failure. Therefore, early failure signals give time for the maintenance engineers to make the necessary adjustments and repairs before the failure actually occurs.

Fig. 3: Variance analysis of the CO₂ sensor failure.

3.2 Reconstruction performance of the IAQ measurements

As discussed in previous sections, three types of abnormal data scenarios are used to compare the performance of the proposed method to that of other approaches. Due to sensor ageing, damage, poor working environmental conditions, etc., a number of sensor failures may occur, with complete failure and bias failure being the most common. To evaluate the reconstruction performance of AE-based sensor faults, we introduced different kinds of faulty data in the test dataset. The magnitude size of each sensor failure is described in Section 2.4. The rate of fault data was set to 0.5, and the fault data lasting 3 days were randomly inserted into the test set. Figure 4 shows a section of CO₂ data containing fault data segments and the reconstructed results. Table 2 demonstrates the fault data reconstruction results of the AE-based model for different faulty data. Root mean square error (RMSE) and mean absolute error (MAE) are used as metrics to measure the capability of the AE-based model of reconstructing the fault data. The largest value of RMSE is 43.197 ppm calculated by the standard AE model. When the encoder and decoder structures are designed using LSTM, the reconstruction performance improves by up to 17%, which proves that LSTM can capture the nonlinear and autocorrelated relationships of CO₂ data. In addition, the reconstruction model based on VAE provides better capability for fault data reconstruction. This is because VAE can solve the problem of non-regularized latent space in the encoder and provide generation capability for the whole space. The encoder of AE produces the vectors in the latent space, while VAE outputs the distribution in the latent space for each input, adding a constraint on that distribution to convert it to a normal distribution, and this constraint guarantees that the latent space is regularized. As a result, the VAE-LSTM reconstruction accurately forces the faulty data to normality.

Table 2: Reconstruction performance of different approaches under different types of fault data.

Fault data reconstruction performance						
Reconstruction methods	Complete failure		Bias failure		Precision degradation fault	
	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)	RMSE (ppm)	MAE (ppm)
AE	39.441	28.105	43.197	30.126	39.789	28.696
AE-LSTM	37.343	26.008	36.882	25.867	36.800	25.814
VAE-MLP	34.370	24.111	33.228	26.104	32.712	25.087
VAE-LSTM	32.153	24.508	31.133	18.212	27.133	17.081

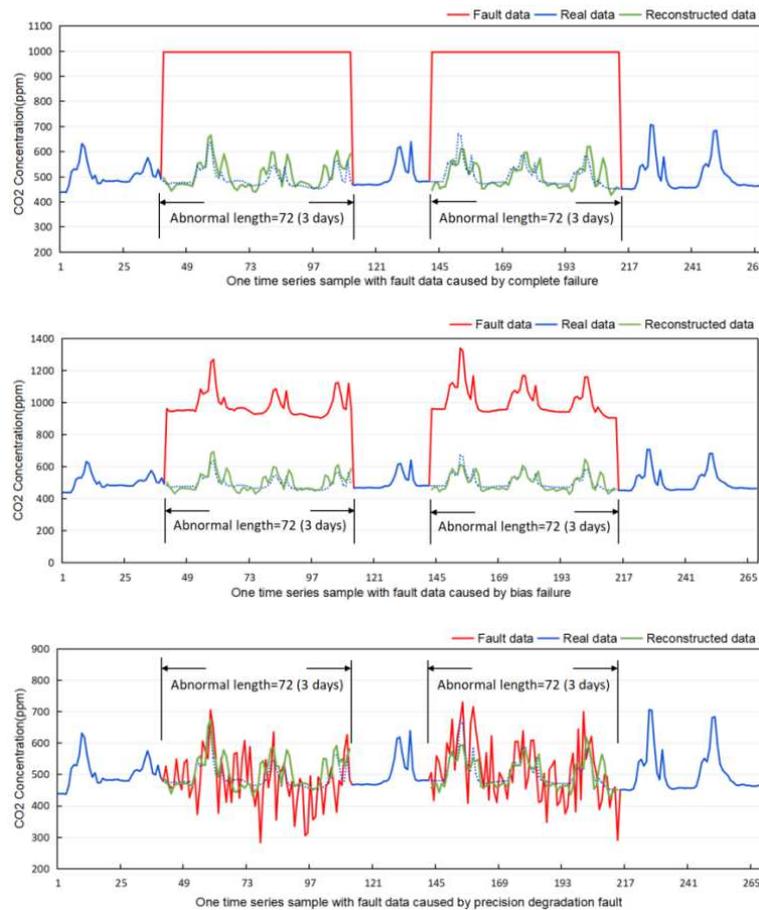


Fig. 4: Reconstruction performance of the VAE-LSTM model under interval-based fault scenarios

3.3 Discussion

This study presents a method to reconstruct IAQ data since abnormal data such as bias failure, complete failure, precision degradation fault often occur due to sensor malfunctions. One of the main contributions of this paper is the development of the VAE-based model for reconstructing various types of abnormal data, including LSTM configurations that properly depict indoor environmental patterns. The time series of CO₂ concentration, in particular, has periodic peaks and extremes, which is a complexity to consider when developing models employing LSTM. Furthermore, model training should be done offline to guarantee that the network has sufficiently learned the basic parameters in order to provide optimal learning performance for IAQ data and achieve accurate abnormal data reconstruction.

In addition to the reconstruction of indoor CO₂ data, the VAE-LSTM model developed can be generally applied to other tasks involving periodic abnormal data processing, such as indoor crowd and energy consumption. In fact, there is a link between indoor CO₂ concentration, indoor crowd, and energy consumption. CO₂ concentration can be generally used as a proxy indicator to assess whether indoor space is occupied and whether indoor crowd affects energy consumption. Our proposed VAE-LSTM approach encodes categorical features, such as months, weekdays, hours, and holidays into integer sequences as auxiliary information to capture the periodic patterns of time series data, and the original categorical sequences are connected to the generated sequences of the decoder to provide more control over the process of reconstructing and imputing the sequences. The flow pattern of a human crowd and fluctuations of energy demand have similarities with fluctuations in indoor CO₂, and both follow a cyclic pattern, so our proposed VAE-LSTM method can also be used to constitute a model for processing crowd and energy demand from abnormal data.

In contrast to other studies that only utilize VAE-based models for abnormal detection, this study incorporates variance analysis to detect non-periodic abnormal signals in IAQ data in advance, and early failure detection can prevent problems in critical parts of the Heating Ventilation and Air Conditioning (HVAC) system. For example, the case in this study is the indoor CO₂ concentrations at a university restaurant. When abnormal signals are detected using our proposed approach, the restaurant manager can contact engineers to check the system in time.

Even if part of the system facilities is shut down, the restaurant manager can prepare backup ventilating equipment in advance to ensure that the customers can enjoy their meal in a good indoor environment, especially during peak hours.

4. CONCLUSIONS

A neural approach, consisting of a variational autoencoder and long-short-term memory network (VAE-LSTM), was developed for early detection and reconstruction of malfunctioning sensors in HVAC systems in order to improve the reliability of the sensors in indoor environment control. Taking advantage of the periodicity and stable fluctuation characteristics of IAQ data, the results of variance analysis on reconstruction errors reveal that unusual behavior of the data can be detected as early as 12 hours before failure occurs. The abnormal data are then reconstructed using the developed VAE-LSTM model. The validation is carried out by introducing different types of abnormal data on the CO₂ sensor. The superiority of the VAE-LSTM was then illustrated by comparing the developed approach to other methods.

However, for an approach dealing with faulty sensors, an explanatory function of fault locations and causes should be provided to the on-site engineer in order to avoid time-consuming proactive repairs, and the knowledge-based method or expert rules can meet this requirement and capability. Therefore, in our future work, we will provide rational explanations for system failures by combining analytical-based, knowledge-based and data-driven approaches and apply them to fault detection and diagnosis of ventilation control systems, especially for large-scale building systems.

REFERENCES

- Ali, A., & Dağtekin, R. (2008). Early warning signals of the 2000/2001 Turkish financial crisis. *International Journal of Emerging and Transition Economies*, 1(2), 191-218.
- Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). *LOF: identifying density-based local outliers*. Paper presented at the Proceedings of the 2000 ACM SIGMOD international conference on Management of data.
- Bu, J., Liu, Y., Zhang, S., Meng, W., Liu, Q., Zhu, X., & Pei, D. (2018). *Rapid deployment of anomaly detection models for large number of emerging kpi streams*. Paper presented at the 2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC).
- Cao, V. L., Nicolau, M., & McDermott, J. (2016). *One-class classification for anomaly detection with kernel density estimation and genetic programming*. Paper presented at the Genetic Programming: 19th European Conference, EuroGP 2016, Porto, Portugal, March 30-April 1, 2016, Proceedings 19.
- Drake, J. M., & Griffen, B. D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, 467(7314), 456-459.
- EPA. (2009). Indoor Air Quality Tools for Schools. Retrieved from <https://www.epa.gov/iaq-schools/indoor-air-quality-tools-schools-action-kit>
- Erfani, S. M., Rajasegarar, S., Karunasekera, S., & Leckie, C. (2016). High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. *Pattern Recognition*, 58, 121-134.
- Hussain, S. N., Abd Aziz, A., Hossen, M. J., Ab Aziz, N. A., Murthy, G. R., & Mustakim, F. B. (2022). A novel framework based on cnn-lstm neural network for prediction of missing values in electricity consumption time-series datasets. *Journal of Information Processing Systems*, 18(1), 115-129.
- Kasam, A. A., Lee, B. D., & Paredis, C. J. (2014). *Statistical methods for interpolating missing meteorological data for use in building simulation*. Paper presented at the Building Simulation.
- Kim, M., Liu, H., Kim, J. T., & Yoo, C. (2014). Evaluation of passenger health risk assessment of sustainable indoor air quality monitoring in metro systems based on a non-Gaussian dynamic sensor validation method. *Journal of hazardous materials*, 278, 124-133.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
- Lee, W.-Y., House, J. M., Park, C., & Kelly, G. E. (1996). Fault diagnosis of an air-handling unit using artificial neural networks. *Transactions-American society of heating refrigerating and air conditioning engineers*, 102, 540-

549.

Li, D., Zhou, Y., Hu, G., & Spanos, C. J. (2016). Fault detection and diagnosis for building cooling system with a tree-structured learning method. *Energy and buildings*, *127*, 540-551.

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental pollution*, *231*, 997-1004.

Lu, Y.-C., Shen, C.-H., & Wei, Y.-C. (2013). Revisiting early warning signals of corporate credit default using linguistic analysis. *Pacific-Basin Finance Journal*, *24*, 1-21.

Ma, J., Cheng, J. C., Jiang, F., Chen, W., Wang, M., & Zhai, C. (2020). A bi-directional missing data imputation scheme based on LSTM and transfer learning for building energy data. *Energy and buildings*, *216*, 109941.

Mesa-Jiménez, J. J., Stokes, L., Yang, Q., & Livina, V. (2021). Early warning signals of failures in building management systems. *International Journal of Metrology and Quality Engineering*, *12*, 11.

Ouyang, T., Zha, X., & Qin, L. (2017). A combined multivariate model for wind power prediction. *Energy conversion and management*, *144*, 361-373.

Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, *148*, 461-468.

Rogers, B. M., Solvik, K., Hogg, E. H., Ju, J., Masek, J. G., Michaelian, M., . . . Goetz, S. J. (2018). Detecting early warning signals of tree mortality in boreal North America using multiscale satellite data. *Global change biology*, *24*(6), 2284-2304.

Su, Y., & Kuo, C.-C. J. (2019). On extended long short-term memory and dependent bidirectional recurrent neural network. *Neurocomputing*, *356*, 151-161.

Szabados, M., Kakucs, R., Páldy, A., Kotlík, B., Kazmarová, H., Dongiovanni, A., . . . Kukec, A. (2022). Association of parent-reported health symptoms with indoor air quality in primary school buildings—the InAirQ study. *Building and Environment*, *221*, 109339.

Wang, H., Feng, D., & Liu, K. (2021). Fault detection and diagnosis for multiple faults of VAV terminals using self-adaptive model and layered random forest. *Building and Environment*, *193*, 107667.

Wen, J., & Gao, H. (2018). Degradation assessment for the ball screw with variational autoencoder and kernel density estimation. *Advances in Mechanical Engineering*, *10*(9), 1687814018797261.

Yan, K., Zhong, C., Ji, Z., & Huang, J. (2018). Semi-supervised learning for early detection and diagnosis of various air handling unit faults. *Energy and buildings*, *181*, 75-83. doi:10.1016/j.enbuild.2018.10.016

Yu, Y., Woradechjumroen, D., & Yu, D. (2014). A review of fault detection and diagnosis methodologies on air-handling units. *Energy and buildings*, *82*, 550-562.

Zhao, Y., Li, T., Zhang, X., & Zhang, C. (2019). Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renewable and Sustainable Energy Reviews*, *109*, 85-101.