Unsupervised Multimodal Learning for Fault Detection of Fan Coils Units Using Building Automation System Data

Chin Choy Chai, and J. J. McArthur*

Department of Architectural Science, Toronto Metropolitan University, Canada.
chaicc@torontomu.ca; jennifer.mcarthur@torontomu.ca**corresponding author*

Abstract

An unsupervised multimodal learning scheme based on autoencoder (AE) is proposed and developed for fault detection of building heating, ventilation, and air conditioning (HVAC) systems, applicable across multiple types of fan coil units (FCUs), i.e., cooling, heating, cooling & heating FCUs. Building automation system data of these three types of FCUs are used in developing, training and evaluating the proposed models and algorithms. The main contributions of this work include: (i) a methodology for developing a generalized AE model through multimodal and transfer learning; (ii) generalizing the AE model across three specific types of FCUs through cross-testing and transfer learning, which enables fault detection at the device-level (a specific FCU type), and system-level (across different FCU types).

Keywords–Unsupervised learning; multimodal learning; autoencoder; fan coil units; fault detection, building HVAC

1 Introduction

According to Intergovernmental Panel on Climate Change (IPCC) report [1], buildings account for 37% of global carbon emissions and 34% of energy demand worldwide. In particular, a building's heating, ventilation, and air conditioning (HVAC) system accounts for half of its total energy use, making them a primary focus for energy optimization [2]. Due to existing design flaws, operational variances, and system faults [3] [4], any undetected faults can cause an additional 103 to 500 TWh of energy consumption annually in the U.S. [5]. To mitigate these issues, fault detection and diagnosis (FDD) has been employed as a cost-effective solution for identifying potential system failure or deficiencies, and addressing these faults [6].

A fan coil unit (FCU) is the "last mile" in many building HVAC systems, providing zone-level heating and cooling. FCUs typically consist of a heating and/or cooling coil, fan, and filter, and are controlled with local thermostats to achieve desired space conditions. Large buildings may contain dozens – even hundreds – of FCUs installed in a building, thus ineffective or faulty FCUs operation can result in occupant discomfort and energy waste. Limited research has explored FDD for FCUs and the majority rely on simulation-based methods using simulated data [7, 8], or unsupervised [9] [10] methods. Further, existing FDD solutions are mostly developed specifically for particular, limited fault categories (e.g. valve stuck, fan failure) [10] or a particular FCU type. There is still a lack of generalized solutions for characterizing FCU operation and FDD for the FCUs, in lieu of designing different FDD solutions for different types of FCUs for cooling or heating. Even if machine learning or neural network-based solutions were to be implemented, they require retraining for at least three different types of FCUs (one for each type), which incurs significant additional resources.

In this paper, we propose an unsupervised multimodal learning approach for developing and designing novel algorithms for fault detection of the FCUs. A methodology for developing a generalized autoencoder (AE) model through multimodal and transfer learning is first presented. This methodology facilitates unsupervised learning for fault detection at the equipment-level (for a specific FCU type), and system-level (across different FCU types). Then, a global model and specific models are developed for fault detection of multiple FCU types. Fault detection performance of the global model and specific models are evaluated to demonstrate the generalizability of the AE across different FCU types. These are achieved through cross-testing and transfer learning based on real building automation system data. In a nutshell, based on unsupervised multimodal learning, only a single global AE model is required for fault detection through unsupervised learning from historical data of cooling-only, heating-only, cooling & heating FCUs.

2 Background

Figure 1 shows a schematic of an FCU, consisting of a heating coil (H/C), a cooling coil (C/C), and a fan. It is part of an HVAC system and is widely used in multi-unit residential, commercial, and industrial buildings.

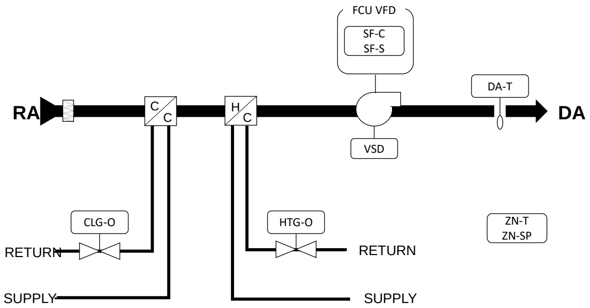

Figure 1: Schematic diagram of a FCU; VSD denotes variable speed drive.

Each component of a FCU and their functionalities are described in Table 1.

Table 1. FCU components and functionalities.

Component	Description
Cooling coil command (CLG-O)	The cooling coil is controlled by a valve controlled by a command CLG-O (of continuous value between 0 and 1), which indicates the percentage the valve is open. The valve is for controlling the flow of chilled water to regulate the coil's cooling capacity as needed.
Heating coil command (HTG-O)	The heating coil is controlled by a valve controlled by a command HTG-O (of continuous value between 0 and 1), which indicates the percentage the valve is open for supplying hot water from the building's central heating plant. When heating is needed, this coil introduces warmth to the air drawing into it, effectively raising the temperature. The flow of hot water through the heating coil is regulated by a valve for controlling the amount of heat to the air.
Supply air (DA)	Conditioned air to the space. The temperature of the supply air is indicated by the DA-T sensor.
Zone sensors (ZN-T, ZN-SP)	Sensors placed within the conditioned zone. ZN-T is the zone temperature; ZN-SP is the set-point of the temperature.
Fan sensors (SF-C, SF-S)	The fan draws air through the coils and then discharges it into the room. Supply Fan Command (SF-C) is a control output which commands the operation of the supply fan. Supply Fan Status (SF-S) indicates the supply fan's on/off status.
Occupancy schedule (OCC-SCHEDULE)	The system will operate based on a time of day schedule adjustable by the user.
Unit enable signal (UNITEN-MODE)	A network unit enable signal (UNITEN-MODE) will control the mode of the unit.

In this work, real building automation system datasets of cooling-only, heating-only, and cooling & heating FCUs in a building HVAC system are adopted. The data information covers the room temperature (ZN-T), supply air temperature (DA-T), as well as command data such as cooling valve output (CLG-O), heating valve output (HTG-O), and temperature set point (ZN-SP), etc., as summarized in Table 1.

2.1 Principle of an Autoencoder

AEs are a type of neural network model that enable unsupervised features extraction [11], due to their strong representation learning capabilities, simple structure, and ease of training. As illustrated in Figure 2, the architecture is formed by an encoder and a decoder.

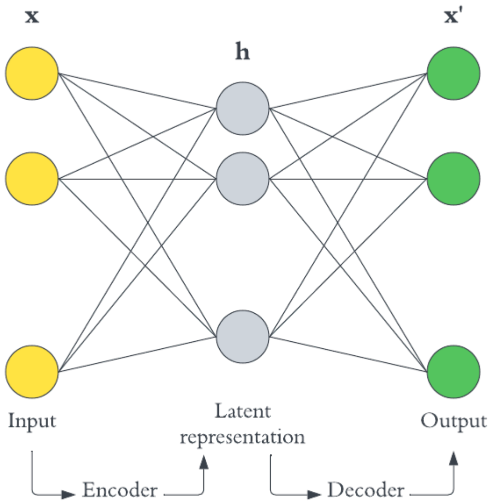

Figure 2: Neural network architecture of an AE.

The mathematical formulation of the AE architecture is as follows:

$$\boldsymbol{h} = f(\boldsymbol{W}_e\boldsymbol{x} + \boldsymbol{b}_e) \tag{1}$$

$$\boldsymbol{x}' = g(\boldsymbol{W}_d\boldsymbol{h} + \boldsymbol{b}_d) \tag{2}$$

where $\boldsymbol{x}$ and $\boldsymbol{x'}$ denote the input and reconstructed output, respectively, $\boldsymbol{h}$ denotes the latent space; $\boldsymbol{W_e}$ and $\boldsymbol{b}_e$ denotes the encoder's weights and bias respectively, while $\boldsymbol{W}_d$ and $\boldsymbol{b}_d$ denote the decoder's weights and bias respectively; $f()$ and $g()$ denote each of their respective activation functions.

In an AE, the encoder section first compresses high-dimensional input data into a compact latent representation. This compact latent representation is then used by its decoder section to reconstruct the original data.

During training phase, the AE's objective is to minimize the error between the original input data and its reconstructed data. Given a training data set of size N, the neural network weights of an AE are optimized by minimizing a loss function $\mathcal{L}$ using gradient descent algorithm. The loss function $\mathcal{L}$ quantifies the reconstruction error as,

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left|x^{(i)} - x'^{(i)}\right|^2 \quad (3)$$

After successful training, the latent space $\boldsymbol{h}$ can be adopted as the feature representation for any given input $\boldsymbol{x}$.

This training process facilitates the neural network to learn and capture essential relationships among input variables, while filtering out any anomalies and noise, enabling the latent space to characterize any hidden patterns in the original input data. In general, the above encoder-decoder framework also facilitates integration of other deep learning networks like recurrent neural networks and convolutional neural networks for more complex representation learning [12].

A larger latent space that is close to the input size could lead to overfitting, as the AE model merely copies the input data without extracting essential patterns or features from the original input data. Conversely, a latent space that is too small, such as one dimension could cause underfitting and poor reconstruction of input data, as the AE model fails to extract any features or features from the original input data.

3 Multimodal AE for Fault Detection

The primary goal of this work is to design a generalized AE that captures the dynamics and interaction of unimodal (cooling-only, heating-only) and bimodal (heating & cooling) operations of three different FCU types, while also models each of them separately. Previous works in multimodal research has proposed feature concatenation and tensor fusion approaches for multimodal learning. In this work, the tensor fusion approach is adopted for multimodal modeling and learning of three types FCUs, by first defining the following vector field using a 2-fold Cartesian product:

$$\{(\boldsymbol{z}_c, \boldsymbol{z}_h) | \boldsymbol{z}_c \in \begin{bmatrix} CLG\_0 \\ 1 \end{bmatrix}, \boldsymbol{z}_h \in \begin{bmatrix} HTG\_0 \\ 1 \end{bmatrix}\} \quad (4)$$

where the subscripts $_c$ and $_h$ denote, respectively, cooling and heating; CLG-O denotes the cooling valve output command to a FCU for cooling; HTG-O denotes the heating valve output command to a FCU for heating. Note that the extra constant dimension with value 1 is for generating the three distinct input feature variables that capture, respectively, the dynamics and interaction unimodal (cooling-only, heating-only) and bimodal (heating & cooling) operations. As a result, for the designed AE neural network, each coordinate $(\boldsymbol{z}_c;\, \boldsymbol{z}_h)$ can be seen as a 2-D point in the 2-fold Cartesian space defined by the embedding vector $\begin{bmatrix} CLG\_0 \\ 1 \end{bmatrix}$ for cooling, and the embedding vector $\begin{bmatrix} HTG\_0 \\ 1 \end{bmatrix}$ for heating. Mathematically, this is equivalent to an outer product between the representation for cooling, $\boldsymbol{z}_c$, and the representation for heating, $\boldsymbol{z}_h$,

$$\boldsymbol{z}_m = \begin{bmatrix} CLG\_0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} HTG\_0 \\ 1 \end{bmatrix} \quad (5)$$

where $\otimes$ denotes the outer product between two vectors.

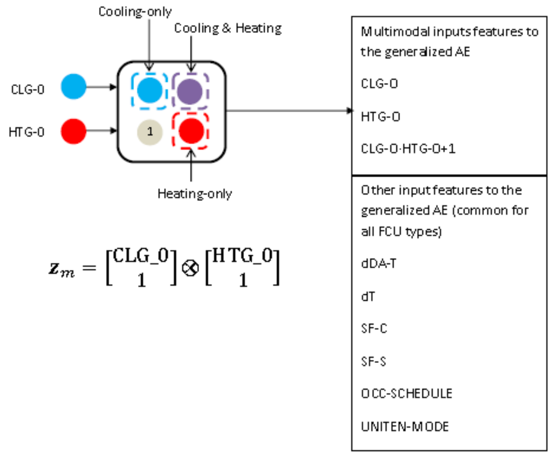

Figure 3: Proposed fusion approach with unimodal (cooling-only, heating-only) and bimodal (heating & cooling) features as additional inputs to the multimodal AE

As illustrated in Figure 3, the resulting matrix $\boldsymbol{z}_m$ can be viewed as a 2D surface that can be divided into 3 sub-regions, corresponding to three combinations of the data variables CLG-O, HTG-O and 1. In a nutshell, these three sub-regions characterize unimodal and bimodal dynamics, as well as interactions between three operational modes (cooling-only, heating-only, heating & cooling) of these three different FCU types.

3.1 Multimodal Learning for Fault Detection

Based on the design in Figure 3, the new input feature variable, G = CLG-O·HTG-O+1, is created, where CLG-O·HTG-O denotes the element-wise product of command points of cooling valve output command (CLG-O) and heating valve output command (HTG-O). Essentially, this variable is created for modeling the multimodal AE's behavior in learning and capturing the overall dynamics and interaction across the cooling-only, heating-only, heating & cooling FCUs as follows:

- For cooling-only FCUs, CLG-O ≠ 0 and HTG-O = 0; G = 1. HTG-O does not contribute to the error for fault detection.
- For heating-only FCUs, CLG-O = 0 and HTG-O ≠ 0; G = 1. CLG-O does not contribute to the error

for fault detection.

- For heating & cooling FCUs, CLG-O ≠ 0 and HTG-O ≠ 0; G > 1. Both CLG-O and HTG-O contribute to the error for fault detection.

The value of G also provides an indication to the multimodal AE that the FCU is operating for any of the cooling-only, heating-only, or heating & cooling purpose. In particular, the G enables the AE to capture the instances with hot water flowing through the heating coil, while chilled water flowing through the cooling coil (simultaneously), which is likely a fault event. On top of the above, all input variables are standardized using a min/max scaler to ensure uniform scaling across the dataset.

In addition, at the HVAC controllers, the temperature difference between the observed temperature and the set point is used for regulating the heating and cooling. Due to the temperature set-points are adjusted by different users and according to different cooling/heating requirements at different rooms in the building, there is a lack of consistency (rule) across different FCUs for set-point setting. It is challenging for a small neural network to capture this dynamic due to such inconsistency in setting set-points. To address this limitation, new feature variables are created [13]: (i) dT (room temperature ZN-T minus its set-point ZN-SP); (ii) dDA-T (discharge air temperature DA-T minus the room temperature set-point ZN-SP) and used as input to the AE models. These new variables emulate the controller's behavior, by taking account only the difference relative to the temperature set-point.

The other input feature variables to the AE, including CLG-O, HTG-O, SF-C, SF-S, etc., are points on the building automation system sensors, which are evaluated to identify unusual behavior, which could indicate an equipment fault. Based on the encoder and decoder outputs of each of the AE model for FCUs, the mean absolute error (MAE), which quantifies the error between each model's predictions and the actual observations, can be computed. The MAE is used for evaluation of both global model and specific models for FCUs, as to be explained in Section 3.3.

3.1.1 Effects of Faults on AE Input Variables

Potential fault examples of the FCUs are summarized in Figure 4. In Table 2, we present potential fault examples and their effects on the input variables to the AE. Note that these cover a subset of potential faults that can be inferred from existing sensors and not all possible fault combinations.

Note how these potential faults affect the input feature variables to the AE models. As an example, for fault CV1 (cooling valve stuck open), the valve command signal CLG-O could be at 0% (or any low value), while the room temperature (ZN-T) and supply air temperature (DA-T) are low due to full cooling conditions

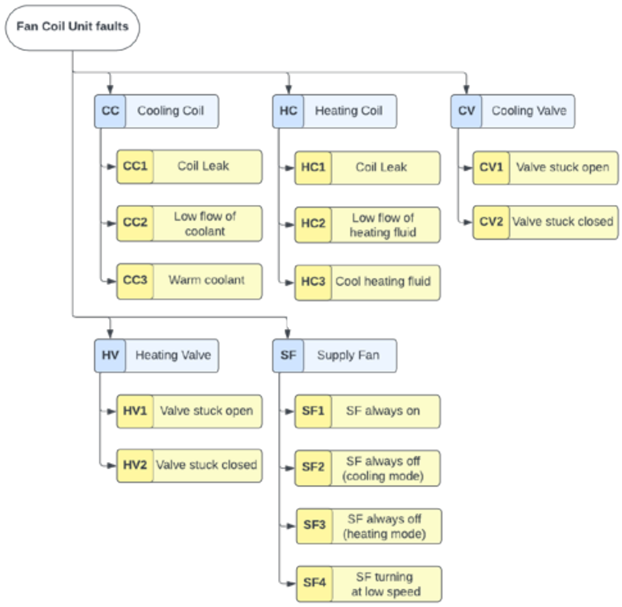

Figure 4: Classification of FCU faults [13].

Table 2: Potential faults examples [13] and their effects on the input feature variables to the AE.

Fault codes	Condition	Effects of fault
CC1, CC2, CC3, CV2	CLG-O = 100%	High DA-T
HC2, HV2, HC1, HC3	HTG-O = 100%	Low ZN-T & DA-T
CV1	CLG-O = 0%	Low ZN-T & DA-T
HV1	HTG-O = 0%	High ZN-T & DA-T
SF2, SF4	CLG-O = 100%	High ZN-T

3.2 Development of Generalized AE Models

Two generalized AE models are developed and proposed: (i) Specific models (cooling-only, or heating-only, or heating & cooling); (ii) A global model that characterizes all three types of FCUs.

Proposed unsupervised multimodal AE for FCUs: The created feature G is adopted as a new input to the multimodal AE model. The proposed multimodal AE model that covers all cooling-only, heating-only, heating & cooling FCUs is illustrated in Figure 5. Following the abovementioned fusion design approach, the multimodal AE characterizes dynamics of these three FCU-types for fault detection.

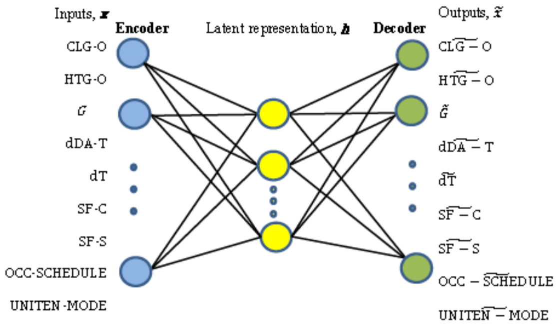

Figure 5: Proposed multimodal AE model

3.2.1 Global Model and Specific Models for FCUs

In this subsection, training steps for the multimodal AE to develop a global model and specific models for fault detection of FCUs are presented.

Specific models: The multimodal AE model in Figure 5 is trained to be ('specifically') applicable for fault detection for a specific FCU type (e.g., cooling-specific FCU). To that end, each specific model for a given FCU is trained with two-third of its own (or type) data. For example, the specific model of cooling-only FCU-C1 can be trained by its own historical data. By extension, we can also develop models for each type of FCUs, for e.g., cooling-only FCU, using data from multiple cooling-only FCUs. The specific model of any FCU is trained and used for fault detection of that particular FCU. Consequently, a specific model is expected to learn and characterize specific operational patterns which are limited to each type of FCU (as it is trained for), with a tradeoff for limited generalization capability.

Global model: The multimodal AE model in Figure 5 is trained to be (globally) applicable for fault detection across all three types of FCUs (cooling-only, heating-only, or heating & cooling). To that end, the (single) global model is trained with two-third of the data from a combination of all three types of FCUs. This global model is developed through a combination of training data from cooling-only FCUs (labeled as FCU-C1, FCU-C2, FCU-C3), heating-only FCUs (FCU-H1, FCU-H2, FCU-H3) as well as heating & cooling FCUs (FCU-HC1, FCU-HC2, FCU-HC3).

The single global model is trained for fault detection of three types of FCUs. Consequently, the global model is expected to learn and characterize more general operational patterns across all three types of FCUs due to its generalization capability, with a tradeoff for less accurate prediction of finer details.

As an example, consider a cooling-only FCU conditioning a server room. Due to constant heat emission from servers, this FCU is likely to operate according to more extreme cooling requirements. The specific (AE) model that is trained (using its cooling-only FCU data) is better at detecting anomalies during the FCU's normal or usual operational setting, but risks false alarms in new operational setting, such as during server maintenance shutdowns.

Alternatively, the above specific model can be substituted with the global AE model that has been trained with data from all three types of FCUs. Due to the global (AE) model's generalization capability, we expect that the global model's is less likely to give false alarms in new operational setting, while it may miss certain finer pattern shifts in the data information.

3.3 Quantifying Reconstruction Error

After learning a latent representation, the AE can be used to predict expected values for all variables at the output of its decoder section. To quantify the reconstruction error, individual MAEs are first evaluated by calculating the difference between the input values and outputs values of each feature variable. The average of all individual MAEs is called the global MAE. Then, fault detection is performed by comparing the global MAE against a predetermined threshold. This threshold can be prescribed through observing deviations of data, for example, by setting it as 10% variation in normalized data. During normal operation of any FCU, the MAE (or reconstruction error) is low. In the presence of any faults or anomalies, a significant deviation of any inputs from their normal values results in a larger MAE for the affected variables. A global MAE exceeding the prescribed threshold indicates that a potential fault occurs, as the FCUs operation is likely deviating from its normal operation. The input feature variable(s) associated with the highest individual MAE(s) are typically implicated as the cause of the fault.

Moreover, the MAE can also be used as a relative performance score for: (i) monitoring and comparing performance across different or same types of FCUs; (ii) tracking the operational health of a single FCU over time. The latter is due to the increasing MAE values could provide indication for potential deteriorating operation and system health of any FCUs.

4 Results and Discussion

In Figure 6 and Figure 7, fault detection results are presented for a heating & cooling FCU using the global model and its specific model with both real and predicted data. As shown in Fig. 6, both the global model and specific model detect similar anomalies patterns of CLG-O·HTG-O > 0, while hot water was flowing through the heating coil, and chilled water was flowing through the cooling coil simultaneously, which are likely fault events.

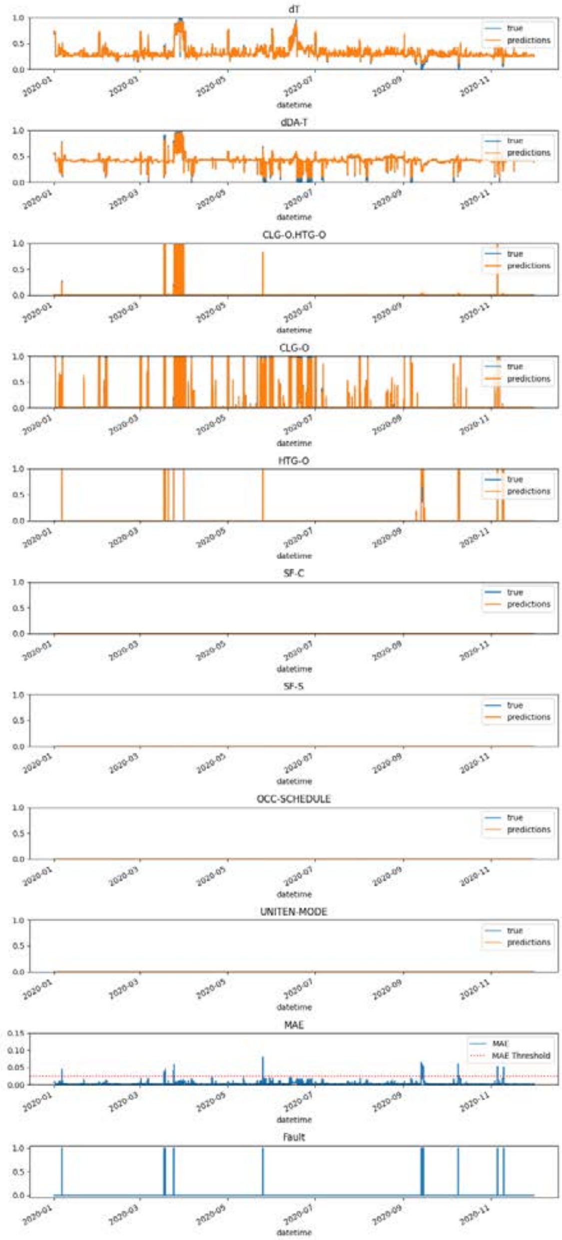

Figure 6: Prediction by the developed specific model for the heating & cooling FCU-HC1.

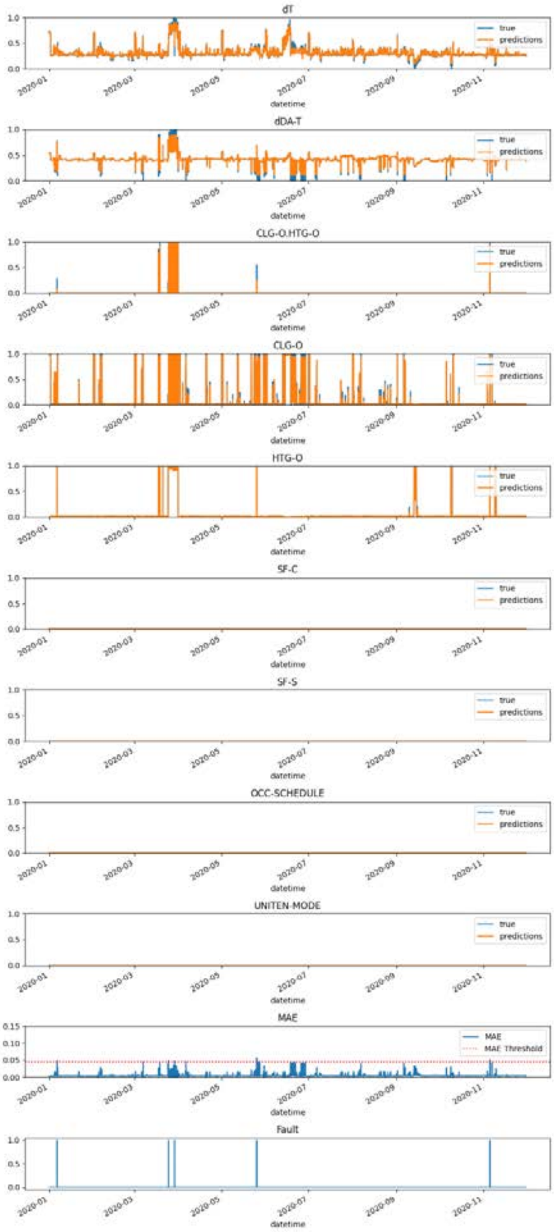

Figure 7: Prediction for the heating & cooling FCU-HC1 by the developed global model.

In particular, in April 2020, as both the supply air temperature (DA-T) and the room temperature (ZN-T) were increasing and causing spikes in dDA-T (= DA-T minus ZN-SP) and dT (= ZN-T minus ZN-SP), the FCU continued to supply hot water through the heating coil (HTG-O $\neq$ 0). Similarly, in Fig. 7, the global and specific models each detect the instance of CLG-O$\cdot$HTG-O > 0, which contribute to the peak of the overall MAE for fault detection. The threshold for MAE is calibrated such that that 0.5% of faults of data points is faulty. As expected, the global model predicts finer fluctuations less accurately than the specific model.

In Figure 8, fault detection results of the global model are presented for a heating-only FCU using real data and prediction. During the winter period of January to April 2020 in Canada, both the supply air temperature (DA-T) and the room temperature (ZN-T) regularly dropped below the setpoint, likely due to envelope heat loss. The heating valve output command (HTG-O) was regularly activated to supply hot water to the FCU's heating coil. Both the global model and the specific model (not shown) detect similar anomalies patterns of the FCU operation as faulty events, which lead to increasing total MAE above the detection threshold.

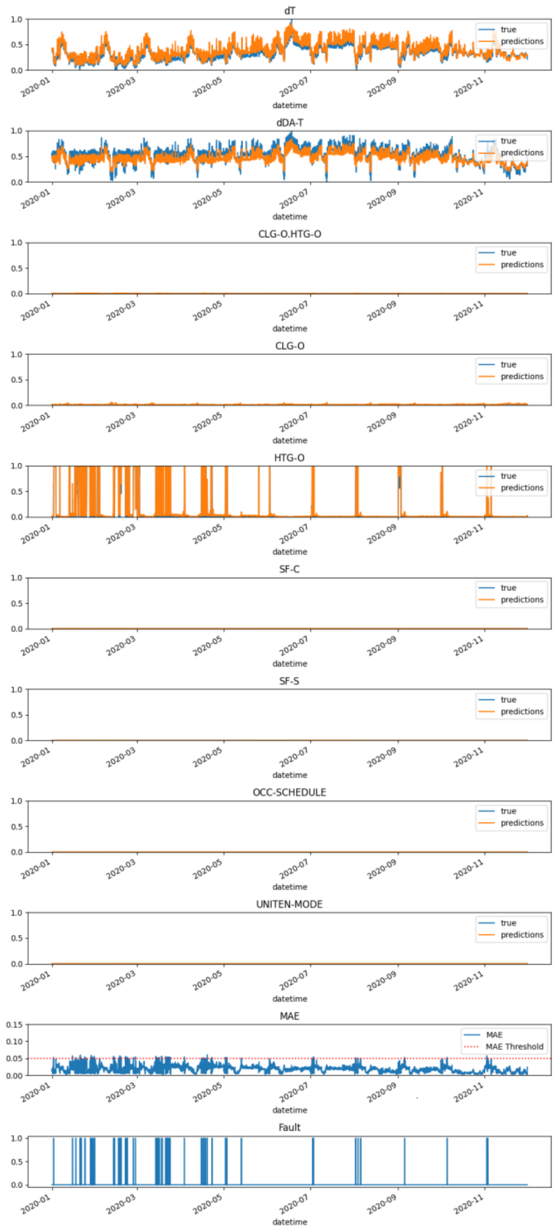

Figure 8: Prediction by the global model for the heating-only FCU-H2.

Data instances of normal operation as well as faulty ones are required for testing the developed models and evaluating the MAE for fault detection. Due to their scarcity, faulty data have been emulated by incorporating expected changes in appropriate sensor measurement data points for each FCU type [13].

In the left column of Figure 9, test results of the global model are presented for the heating-only FCU-H1 in detecting an emulated fault. The test dataset consists of 200 samples of normal data followed by 200 samples of emulated fault data of the heating valve being stuck open, despite not being commanded on (HTG-O = 0).

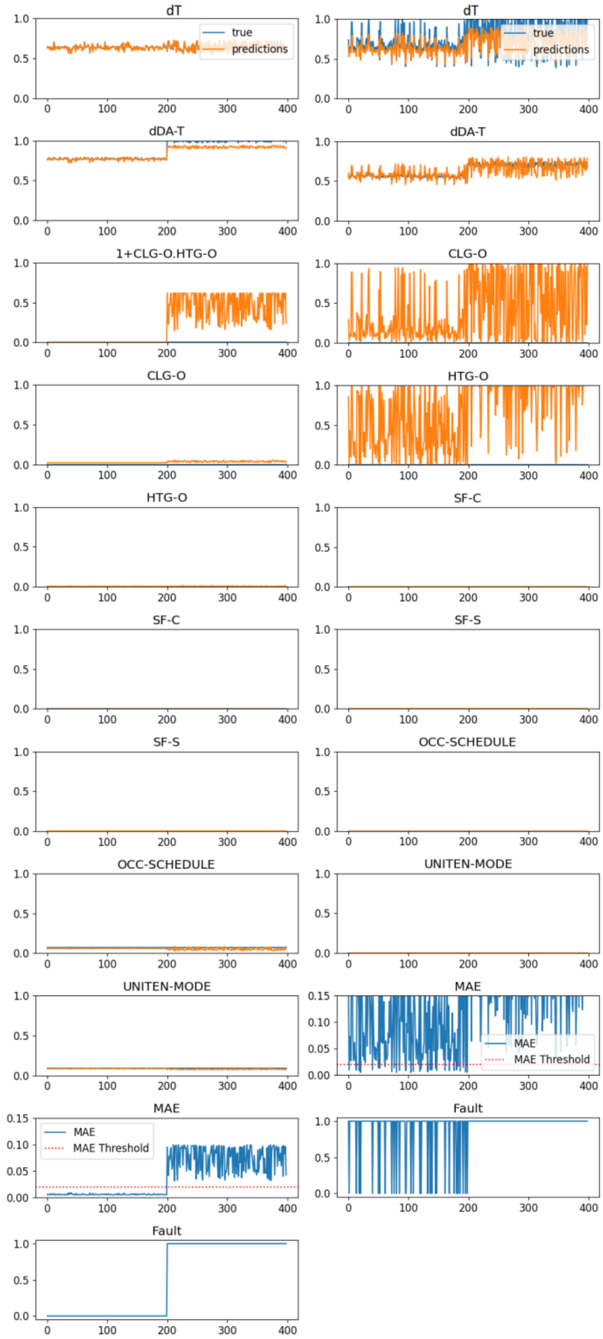

Figure 9: Detection of emulated fault by the: (left) global model with unsupervised multimodal learning, and (right) AE model without multimodal learning [13] for the heating-only FCU-H1.

The room temperature (ZN-T) and supply air temperature (DA-T) were modified accordingly to emulated full heating conditions, exhibiting higher supply air temperature (DA-T) values for the 200th to 400th samples, which the global model correctly identifies as faulty events, with *G* (= CLG-O·HTG-O+1) and dDA-T (= DA-T minus ZN-SP) being the variables that lead to increasing total MAE above the detection threshold.

In the right column of Figure 9, fault detection results are presented for a recently published AE model in [13] without multimodal learning capability, after training it using historical data of 6 FCUs of the heating & cooling type. Without loss of generality, in-slab heating is omitted. In stark contrast, the AE model in [13] predicts that both the heating and cooling valve output commands (HTG-O and CLG-O) are to be activated throughout the normal and emulated fault intervals, in spite of FCU-H1 being a heating-only FCU. Due to lacking of multimodal learning capability, it fails to identify the actual emulated fault event, as the error between its predicted values and true values of dT, CLG-O and HTG-O significantly increase the total MAE above the detection threshold for all samples.

5 Conclusion

A multimodal AE is proposed for developing the global model and specific models for different types of FCUs through unsupervised multimodal learning: (i) the specific models are trained using data of a specific (or type) of FCU(s); (ii) the global model is trained using a combination of data from three types of FCUs. The applicability of unsupervised multimodal AE for fault detection across different types of FCUs is achieved through multimodal learning and transfer learning.

Acknowledgments

This research was funded by the Natural Science and Engineering Council of Canada (ALLRP 597645-24).

References

[1] H. Lee and J. Romero, “Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change,” IPCC, Geneva, 2023.

[2] L. Perez-Lombard, J. Ortiz, R. Gonzalez and I. R. Maestre, “A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes,” *Energy and Buildings,* 41, 272-278, 2008.

[3] A. Mahdavi, C. Berger, H. Amin, E. Ampatzi, R. Andersen, E. Azar, V. Barthelmes, M. Favero, J. Hahn, D. Khovalyg and H. Knudsen, “The role of occupants in buildings’ energy performance gap: myth or reality?” *Sustainability,* 13(6), 3146, 2021.

[4] K. H. Andersen, S. B. Holøs, A. Yang, K. Thunshelle, Ø. Fjellheim and R. Jensen, “Impact of typical faults occurring in demand-controlled ventilation on energy and indoor environment in a Nordic climate,” in *E3S Web of Conferences*, 2020.

[5] K. W. Roth, D. Westphalen, Llana and M. Feng, “The energy impact of faults in US commercial buildings,” in *International Refrigeration and Air Conditioning Conference*, Purdue, 2004.

[6] A. Isazadeh, R. Kamal, C. Yagua, S. Eluvathingal and D. Claridge, “Detecting deficiencies using building performance data in healthcare facilities: Improving operational efficiency with Continuous Commissioning®,” *Energy and Buildings,* 241, 110953, 2021.

[7] Y. Chen, G. Lin, Z. Chen, J. Wen and J. Granderson, “A simulation-based evaluation of fan coil unit fault effects,” *Energy and Buildings,* 263, 112041, 2022.

[8] S. Pourarian, J. Wen, D. Veronica, A. Pertzborn, X. Zhou and R. Liu, “A tool for evaluating fault detection and diagnostic methods for fan coil units,” *Energy and Buildings,* 136, 151-160, 2017.

[9] F. Rajabi and J. McArthur, “Applying OPTICS with and without PCA for fault detection of fan coil units using building automation system data,” *Energy and Buildings,*114368, 2024.

[10] S. Dudley, M. Dey and S. Rana, “Smart building creation in large scale HVAC environments through automated fault detection and diagnosis,” *Future Generation Computer Systems*, 108, 950-966, 2018.

[11] Y. Bengio, A. Courville and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Analysis and Machine Intelligence,* 35(8), 1798-1828, 2013.

[12] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai and T. Chen, “Recent advances in convolutional neural networks,”*Pattern Recognition,* 77, 354-377, 2018.

[13] K. El Mokhtari and J. McArthur, “Autoencoder-based fault detection using building automation system data,” *Advanced Engineering Informatics,* 62, 102810, 2024.