

2022

Fault Detection and Diagnostic Method Based on Evolving Data-driven Model for Radiant Heating and Cooling Systems

Sujit Dahal

Liping Wang

James Braun

Follow this and additional works at: <https://docs.lib.purdue.edu/ihpbc>

Dahal, Sujit; Wang, Liping; and Braun, James, "Fault Detection and Diagnostic Method Based on Evolving Data-driven Model for Radiant Heating and Cooling Systems" (2022). *International High Performance Buildings Conference*. Paper 421.
<https://docs.lib.purdue.edu/ihpbc/421>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries.
Please contact epubs@purdue.edu for additional information.
Complete proceedings may be acquired in print and on CD-ROM directly from the Ray W. Herrick Laboratories at
<https://engineering.purdue.edu/Herrick/Events/orderlit.html>

Fault Detection and Diagnostic Method Based on Evolving Data-driven Model for Radiant Floor Heating and Cooling Systems

Sujit DAHAL^{1*}, Liping WANG¹, James BRAUN²

¹University of Wyoming, Department of Civil and Architectural Engineering,
Laramie, WY, USA

²Purdue University, School of Mechanical Engineering,
West Lafayette, IN, USA

* Corresponding Author: email: sdahal@uwyo.edu, Tel: (307) 223-6025

ABSTRACT

Faults in components (valves, sensors, etc.) of radiant floor heating and cooling systems affect the efficiency, cooling and heating capacity as well as the reliability of the system. While various fault detection and diagnostic (FDD) methods have been developed and tested for building systems, FDD algorithms for radiant heating and cooling systems have previously not been available. This paper presents an evolving learning-based FDD approach for a radiant floor heating and cooling system based on growing Gaussian mixture regression (GGMR). The experimental space was controlled with a building automation system (BAS) in which the operating conditions can be monitored, and control parameters can be overridden to desired values. Trend data for normal operation and faulty operation were collected. A total of six fault types with different severities in a radiant floor system were emulated through overriding control parameters.

An FDD model based on the GGMR approach was developed with training data and the performance of the model was tested for "known" faults that were including in the training and new "unknown" faults that were implemented in the fault testing. The prediction accuracy for each known fault was extremely high with the lowest prediction accuracy of 98% for one of the faults. The algorithm was successful in detecting the new fault as an unknown state before evolving the model and in diagnosing it as a new fault after evolving the model.

1. INTRODUCTION

Hydronic radiant systems have been adopted as an alternative to conventional HVAC systems for achieving higher thermal comfort and energy savings (Fabrizio et al., 2012; Joe & Karava, 2019). Heating or cooling in a radiant system is provided by circulating water through pipes installed within the structure (floors, ceilings, beams, etc.). Hydronic radiant heating and cooling systems consist of several components, including hot and cold-water sources, valves controlling water flow, pumps, temperature sensors, flow sensors, pressure sensors, thermostats, etc. Faults in any of these components, errors in the sensors and instruments, and degradation due to normal wear or lack of maintenance deteriorates the performance of the system. Identifying aberrations from normal operation plays an important role in reducing energy use, maintaining a healthy indoor environment, and prolonging the equipment service life (Chakraborty & Elzarka, 2019). Although a significant amount of data is available from sensors installed in buildings which contains information about faults, it is difficult and time-consuming to analyze and detect faults manually (West et al., 2011). Automated fault detection and diagnostics (AFDD) procedures present the potential to reduce energy waste (Wall et al., 2011).

Methods for fault detection can be categorized into data-driven or process history-based methods, quantitative physical model-based, and qualitative rule and physics-based methods (Katipamula & Brambley, 2005). Qualitative rule and physics-based methods use certain if-then-else rules which can be obtained from first principles or expert knowledge. Rule-based methods are inefficient in real-time computation and in detecting and diagnosing sudden faults if the system requires a long time to reach a steady-state (Yu et al., 2014). Quantitative physical model-based methods are based on physical laws that regulate the system's behavior. The system behavior is predicted for a given set of measured input and model parameters which is compared to measured outputs. The key advantage of a quantitative

(data-driven) model over other FDD methods is that it makes predictions using historical data, does not require detailed building information, and has lower complexity (Katipamula & Brambley, 2005).

The limitation of existing data-driven models is that the models are static. Although data-driven models have been extensively used in the field of fault detection and diagnosis of the HVAC system, the FDD method cannot properly handle the unknown faults in the system. Most of the existing FDD methods can diagnose the faults which are used during the training. When new faults occur in the system, identifying this as an unknown state and integrating its information into the model is crucial to prevent the fault from being misdiagnosed. This limitation can be handled by using adaptive learning which is capable of processing new measurements and performing runtime updates of model parameters in response to the changes in the building system (Karami & Wang, 2018).

Although extensive research is available on the FDD of HVAC systems and components, FDD for radiant heating and cooling system is still lacking. Also, the methods of evolving the FDD algorithm and model in response to the detection of an unknown fault in the system are yet to be explored in the radiant heating and cooling systems and such studies are very rare in the HVAC system as well. In this study, learning based Gaussian mixture model (GMR) model is used for fault detection and diagnosis for the radiant floor heating and cooling systems.

2. METHODOLOGY

To develop the proposed FDD model, experiments were conducted to collect normal operation data that was used to train a baseline model and FDD model under fault-free conditions and faulty data through the implementation of individual faults in the system. A GMR model was developed as a baseline model for predicting radiant slab (RS) flow and heating and cooling capacity. Next, using a feature selection algorithm the features which showed promising characteristics in detecting and diagnosing faults were selected from a list of features measured during the experiments. The features were used to develop an FDD model based on GGMR trained with the normal operation and faulty data collected during the experiments. The GGMR model was tested with known and unknown fault data and if any unknown state was detected and was verified as a new fault, the FDD model was evolved to incorporate the information of the new fault.

2.1. System description

The Living Lab 1 at the Herrick Laboratories, Purdue University was used as the case study to evaluate the proposed learning based FDD method. Living Lab 1 is equipped with a hydronic radiant floor system that provides both cooling and heating. The system diagram of the radiant slab system is shown in Figure 1. The concrete slabs and heat exchangers have RTD sensors to measure the slab temperature and the supply and return water temperatures. The main supply water pipe is divided into 10 channels with individual control valves which supply water to the 10 slabs (A-K) in parallel. The 10 valve positions are determined using a PI feedback loop with space temperature as the controlled variable with a setpoint temperature of 73°F (22.78 °C). The maximum valve position of all the valves is set at 100% and the minimum valve position is set at 35%. The water flow is maintained in the loops by a VFD pump. The supply water temperature is controlled through a heat exchanger connected to a district cool water supply and hot water supply. The cool water flow and hot water flow are controlled using cool water and hot water valves on the supply side of the heat exchanger. Both the cool water valve position and hot water valve position are determined using individual PID loops with supply water temperature as the controlled variable and the supply water temperature setpoint. If the space temperature is above 73 °F (22.78 °C), the system will be in the cooling mode and if it is lower than 73 °F (22.78 °C), the system will be in heating mode. During the cooling mode, the maximum of either the space dew point temperature plus 3 °F (1.7 °C) or 58 °F (14.4 °C) is used as the supply water setpoint temperature to calculate the cool water valve position. During the heating mode, 85 °F (29.44 °C) is used as the supply water setpoint temperature to calculate the hot water valve position. The supply water pump is enabled if either the cool water loop or hot water loop is enabled. The supply water pump position is also controlled based on the output of a PI loop with differential pressure in the supply pipe as the controlled variable with a setpoint pressure of 10 psi (68947.6 pa). The maximum pump output is set at 100% and the minimum is set at 0%.

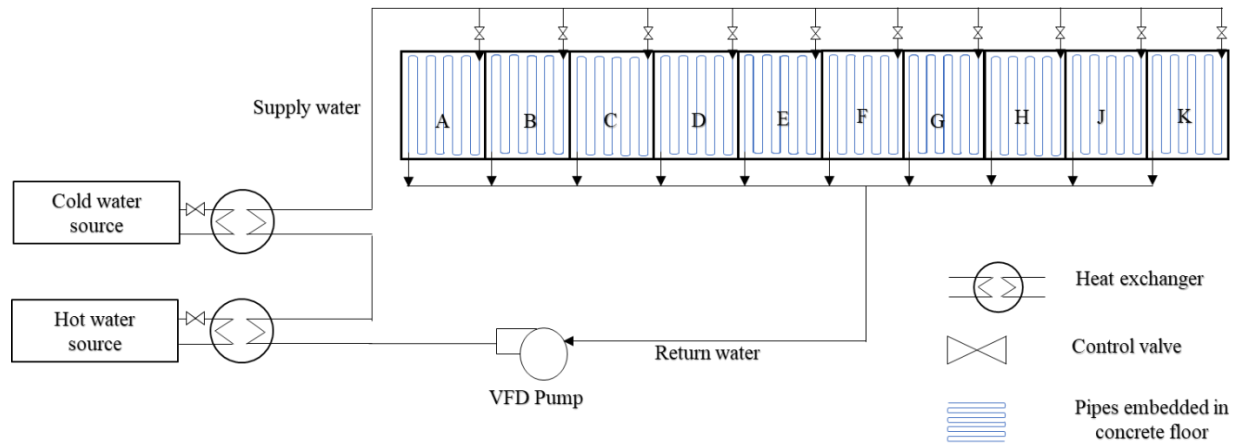


Figure 1: System diagram of the Living Lab 1 radiant floor

2.2. Gaussian Mixture Regression (GMR)

Gaussian mixture regression (GMR) was used to create a baseline model for flow through the radiant slabs, a baseline model heating capacity and cooling capacity of the radiant slab system, and an FDD model. Gaussian Mixture Model (GMM) is a model-based clustering method where the data is identified as a population with K different components. Each component of the GMM is approximated by a Gaussian distribution characterized by a mean as the center of the cluster and a variance. The Gaussian mixture probability density function is a weighted sum of Gaussian probability density functions (pdfs) and is given by Equation 1.

$$f(x, y | \pi, \mu, \delta) = \sum_{j=1}^k \pi_j N(x, y | \mu_j, \delta_j) \quad (1)$$

where $N(x, y | \mu_j, \delta_j)$ is a multivariate Gaussian density function, k is the number of Gaussian mixtures, μ_j is the mean vector, δ_j is the covariance matrix of j -th Gaussian component and π_j is the weight of the j -th component.

GMM is trained using a training dataset to find the unknown parameters of the mixture model (π_j, μ_j, δ_j) using the Expectation-Maximization (EM) algorithm in which the maximum likelihood estimates of the GMM parameters are predicted. A k -means clustering algorithm is used in the EM process to get the initial values for the parameters. When the new input (x) is provided, Gaussian Mixture Regression (GMR) is used for the prediction of the dependent variable Y which can be obtained by computing expectation over $[Y|X = x]$ as equation 2. $\hat{y}(x)$ is the prediction of the dependent variable Y .

$$\hat{y}(x) = E[Y|X = x] = \sum_{j=1}^k w_j(x) m_j(x) \quad (2)$$

where m_j is the regression function derived from the joint mixture Gaussian density and w_j is the weight of each regression model which is extracted using the Bayes rule. The detailed equations and working mechanism of the Gaussian mixture model and Gaussian mixture regression are presented in (Sung, 2004).

2.3. Evolving Mechanism

The GMR model is trained with the baseline data and when there is new data available, the existing GMR is updated. The evolving mechanism can be described using three steps.

Step 1: Calculate the probability density of the new sample data for each Gaussian component, apply the distance criterion and determine the best-matched component using equation 3.

$$P_j = \begin{cases} \pi_j N(\mathbf{d}(t) | \boldsymbol{\mu}_j(t), \delta_j(t)), & \frac{|\mathbf{d}(t) - \boldsymbol{\mu}_j(t)|}{|\delta_j(t)|} < \frac{1}{k} \sum_{j=1}^k \frac{|\mathbf{d}(t) - \boldsymbol{\mu}_j(t)|}{|\delta_j(t)|} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Step 2: If a match is found ($P_j \neq 0$), the parameters (π_j, μ_j, δ_j) of the matching Gaussian are updated.

Step 3: If no existing Gaussian mixture component matches the new sample data ($P_j = 0$) and the number of existing components is less than a maximum threshold provided, a new Gaussian component is generated and the parameters prior, mean, and covariance matrix of the newly generated component are initialized using equation 4.

$$(\pi, c, \mu, \delta) = (\alpha, 1, x_i, \delta_0) \quad (4)$$

As shown in equation (5), after implementing all the steps of the evolving learning approach, the priors of the new Gaussian mixture model are normalized to satisfy the condition of forming a standard Gaussian mixture model which requires the sum of components' priors to be one.

$$\pi_j = \frac{\pi_j}{\sum_{j=1}^{k'} \pi_j} \quad (\text{for } j = 1 \dots k') \quad (5)$$

k' is the number of components in the updated Gaussian mixture model. The evolving process is described briefly in section 2.6 and in detail in the research work by (Bouchachia & Vanaret, 2011).

2.4. Selection of number of Gaussian components

In GMM it is crucial to determine the number of Gaussian components (k) to group the data into an optimal number of clusters. In this study, Akaike Information Criterion (AIC) was used as a model selection criterion. The data was fitted to a Gaussian mixture model and the model was tested for several k values. The model with the smallest AIC and its corresponding k value was chosen as the number of parameters in the Gaussian model. The equation for computing AIC for parameters in the fitted model (k) is given by equation 6 (Akaike, 1998). Here, L is the maximum value of the likelihood function. This criterion was used to select the number of Gaussian components for the flow prediction, heating and cooling capacity prediction, and the FDD model.

$$AIC = -2\ln L + 2k \quad (6)$$

2.5. Fault selection and implementation

The experimental space is controlled with a building automation system (BAS) in which the operating conditions can be monitored, and control parameters can be overridden to desired values. Normal operation and faulty data were collected through experiments performed using an RS system in a real open-plan office setting. A total of 6 fault types with different severities were emulated by overriding control parameters which were selected considering the likeliness of their occurrence in a radiant floor system and the feasibility of implementation using the BAS. The fault types include stuck water flow valves for the radiant floor, hot water supply, cool water supply; sensor offsets for the space thermostat and supply water temperature sensors; and a VFD pump that is stuck at a fixed speed. The implemented faults with varying severity levels along with the dates of implementation are listed in Table 1.

Table 1: Implemented faults

Fault ID	Description (Implementation dates)	Fault ID	Description (Implementation dates)
F1	Space temperature sensor offset of +5 °F (2.78 °C) (12/10-12/2021)	F8	Hot water valve stuck at 0% open (12/6-8/2021)
F2	Space temperature sensor offset of -5 °F (2.78 °C) (12/4-6/2021)	F9	Hot water valve stuck at 50% open (12/8-10/2021)
F3	RS valves stuck at 0% open (12/14-16/2021)	F10	Hot water valve stuck at 100% open (12/12-14/2021)
F4	RS valves stuck at 50% open (12/16-18/2021)	F11	Cold water valve stuck at 0% open (12/25-27/2021)
F5	RS valves stuck at 100% open (12/18-20/2021)	F12	Cold water valve stuck at 50% open (1/7-9/2022)
F6	Supply water temperature sensor offset of +9 °F (5 °C) (12/20-22/2021)	F13	Cold water valve stuck at 100% open (1/9-12/2022)
F7	Supply water temperature sensor offset of -9 °F (5 °C) (12/22-25/2021)	F14	VFD pump stuck at 50% of its maximum speed 1/12-14/2022)

2.6. Fault detection and diagnosis method

The fault detection and diagnosis process along with evolving mechanism is shown in Figure 2 and proceeds through steps 1 to 8.

1. The baseline GMR model for energy use is trained with baseline data.
2. A GGMR FDD model is trained with faulty and normal operation training data.
3. The testing data containing known faults, normal operation, and the unknown fault is run through the Baseline GMR model, and the difference between measured and predicted RS heating and cooling capacity use is compared with the mean baseline difference.
4. If the error is less than the threshold, it is regarded as normal operation and the existing baseline GMR model is evolved with normal operation data. If the error is larger than the threshold, it is deemed as faulty data.
5. The testing data is then run through the GGMR FDD model to get the prediction of the fault label.
6. If the testing data has a predicted fault id that belongs to one of the labeled data, the fault id is reported. If the testing data is predicted as 0 or other labels not included in the training data, it is deemed as an unknown state.
7. The unknown state data is analyzed by an expert and if it is evaluated as a new fault, a new label for the new fault is provided and the GGMR FDD model is evolved to incorporate the information of the new fault.
8. The performance of the GGMR FDD model is tested both before and after evolving the model with new fault data.

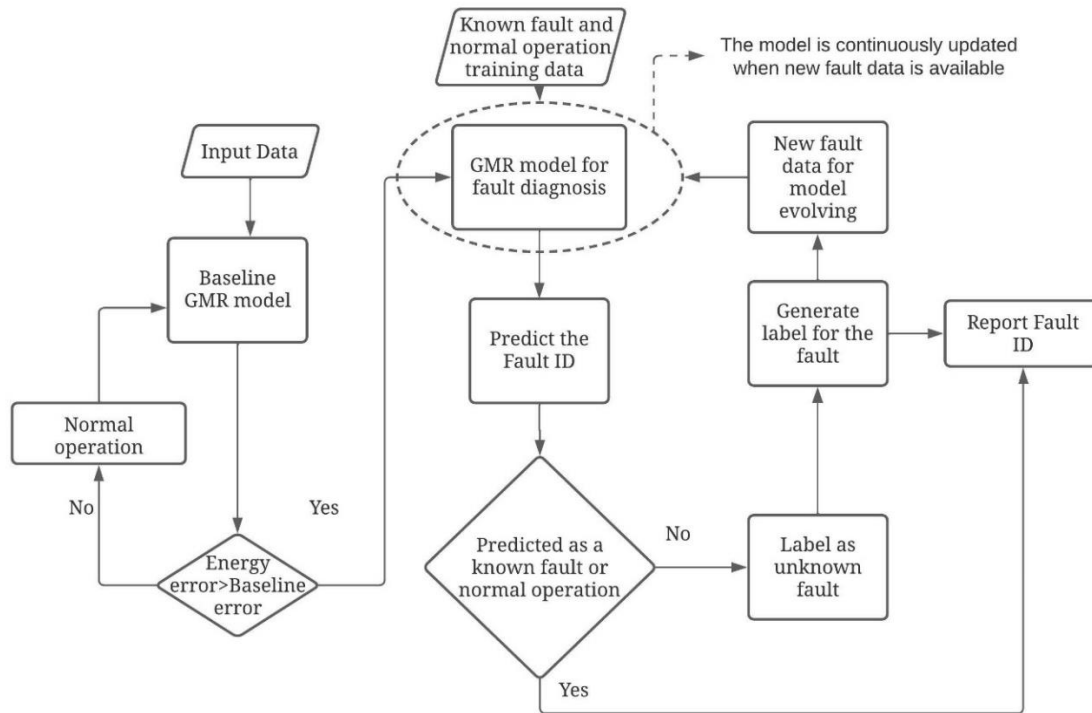


Figure 2: GGMR model generation, evolving, and Fault detection and diagnosis process diagram.

3. RADIANT SLAB SYSTEM HEATING AND COOLING CAPACITY PREDICTION MODEL

The RS heating and cooling capacity was calculated based on the water flow rate, supply water, and return water temperature using equation 7.

$$q = 0.063 \times \frac{5}{9} \times M \times C_p \times \Delta T \quad (7)$$

where q is the calculated heating or cooling capacity (KW), M is the mass flow rate of supply water (gal/min), C_p is the specific heat capacity of water = $4.2 \frac{KJ}{Kg.K}$, and ΔT is the difference between supply and return water temperature. 0.063 is the factor to convert water flow rate (gal/min) to mass flow rate (kg/s), and $\frac{5}{9}$ is the factor to convert °F to °C. Negative values of q represent cooling capacity and positive values represent heating capacity. If the calculated q was positive but the supply water temperature was less than the average slab temperature, the RS heating capacity was set as 0. Similarly, if the calculated q was negative but the supply water temperature was greater than the average slab temperature, the RS cooling capacity was set as 0.

RS heating and cooling capacity are predicted using the Gaussian mixture regression model. The input data used in the heating and cooling capacity prediction model were identified using a correlation coefficient. GMR with evolving was used to develop the model. Normal operation data were divided into training data (11/30 to 12/2 2021) which was used to train the GMR model and testing data (1/14 to 2/3 2022) which was first used to test the model and later evolve the model. The evolved model is then used as the baseline heating and cooling capacity prediction model. When faulty data was collected, this model was used to get the heating and cooling capacity of the RS system under normal operating conditions. Along with the heating and cooling capacity, the water flow rate through the RS system was also predicted using GMR. The predicted water flow rate is one of the inputs for predicting the RS heating and cooling capacity. The input parameters selected for the prediction of water flow rate are the space temperature, RS valve

position, and differential pressure. The input parameters selected for the prediction of RS heating and cooling capacity are the space temperature, RS valve position, average slab temperature, cold water valve position, hot water valve position, and the GMR predicted water flow rate. Ten Gaussian components were selected for both the RS water flow rate prediction model and the heating and cooling capacity prediction model. Both of those values were selected based on the AIC discussed in the previous section. The selected values for learning rate and closeness threshold are 0.001 and 10, respectively. Comparisons between measured and predicted RS flows and measured and predicted heating and cooling capacities during model testing are shown in Figure 3. The coefficient of determination R^2 between predicted and measured RS flow for the training data was 0.95 and between predicted and measured RS heating and cooling capacity was 0.98, which shows excellent prediction accuracies. For model testing data, the CV-RMSE for the predicted RS flow was 6% and was 15.6% for the predicted heating and cooling capacities, which is within the ASHRAE recommended threshold of 30% for the baseline model.

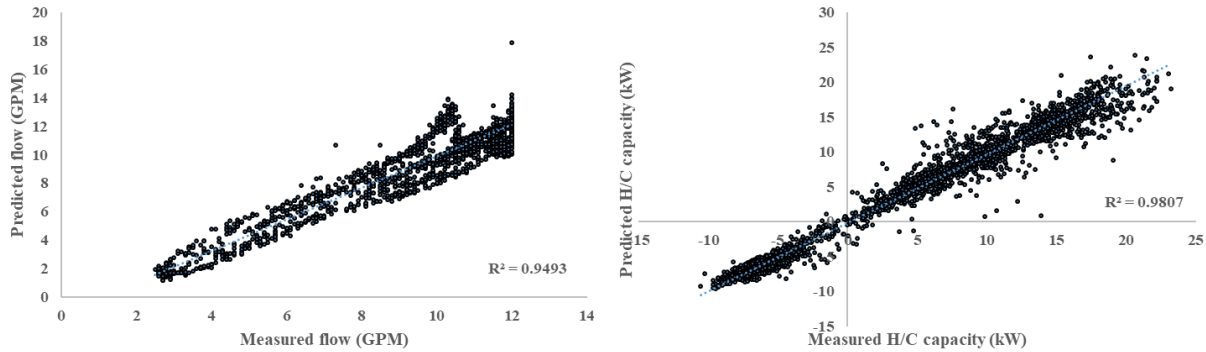


Figure 3: Measured vs Predicted RS flow (left) and Measured vs Predicted RS heating and cooling capacity (right) during model testing

4. FEATURE SELECTION FOR FDD MODEL

The features (input parameters) required for fault detection by GGMR were selected using the feature selection method presented in the research work of (Fijany & Vatan, 2006). The features selected contain the required ability to distinguish between all faults. In the first step, training data is used to recognize the symptoms associated with each fault. In the second step, an optimization algorithm is applied to find the minimum number of features required to diagnose the faults. The occurrence of each fault will have some effect on features that are correlated to the fault. This effect on features could be used as a symptom to identify the faults by comparing the feature values during faulty conditions with feature values during the fault-free condition. Table 2 lists the features used for the case study.

Table 2: Initial set of features

Initial features	
1. Space temperature difference from the setpoint (ΔT)	10. 12-hour cumulative sum of ΔT
2. Differential pressure difference from setpoint (ΔP)	11. 12-hour cumulative sum of ΔP
3. Measured-predicted flow (ΔF)	12. 12-hour cumulative sum of ΔF
4. Measured – predicted Heating and Cooling capacity (ΔC)	13. 12-hour cumulative sum of ΔC
5. Slab temperature difference from setpoint ($T_{\text{slab-sp}}$)	14. 12-hour cumulative sum of $T_{\text{slab-sp}}$
6. Supply water temperature-return water temperature ($T_{\text{sw-rw}}$)	15. 12-hour cumulative sum of $T_{\text{sw-rw}}$
7. Change of RS valve signal (ΔV)	16. 12-hour cumulative sum of ΔV
8. The ratio of pump speed to maximum speed (P_s)	17. 12-hour cumulative sum of P_s
9. Flow change rate for the 24 hours (Δf)	18. 12-hour cumulative sum of Δf

Each data point for each feature was converted into a feature matrix by assigning 2 for symptoms resembling normal operation with the consideration of uncertainties, 1 for symptoms of a faulty condition where the measured values exceed the baseline upper limit, and -1 for symptoms of a faulty condition where the measured values are lower than the baseline lower limit to form a signature matrix. The feature matrix for one of the faults (space temperature positive sensor offset fault) is shown in Table 3 in which each column is an individual feature, and the rows represent each data point obtained with the faulty condition. Only a few features are shown in the table but this process was followed for all the features and all the faults and normal operation data.

Table 3: Feature matrix for space temperature positive sensor offset fault

ΔT	ΔP	ΔF	ΔC	$\Delta T(\text{slab-sp})$	$\Delta T(\text{sw-rw})$	ΔV	P_s	Δf
2	2	2	2	-1	2	2	2	2
2	2	2	2	-1	2	2	2	2
2	2	2	2	-1	2	2	2	2
2	2	2	2	-1	2	2	2	2
2	2	2	2	-1	2	2	2	2

In the next step, key signatures for each fault are identified. If a feature column has the same number in each row for a fault, the symptom is retained (2, 1, or -1), or else the symptom is represented by 0 meaning the feature does not provide a clear symptom required to diagnose the fault. Following this process for all the faults, a sensor matrix (S) is developed which has a key signature for each fault. Now, assuming $S(x)$ is a vector containing the number of features in each row of the sensor matrix S, the optimal sensor placement problem can be formulated as a constrained optimization problem illustrated in equation 8 (Fijany & Vatan, 2006).

$$\begin{aligned} & \text{minimize } \sum_{i=1}^m S(x_i) \\ & \text{Constraint: } \tilde{M}x \geq 1, x_i = 0 \text{ or } 1 \end{aligned} \quad (8)$$

The function in equation 8 minimizes the number of the sensors while the constraint assures that each row of sensor matrix is non-zero and unique. A linear integer programming routine in MATLAB was used to obtain the solution to the optimization problem.

Five different baseline scenarios as listed in Table 4 were developed with varying upper and lower limits to find the best performing baseline threshold. Here, Max and Min represent the maximum and minimum value of the feature respectively and SD is the standard deviation of the feature during the baseline period. The feature selection was done for all five cases with 18 initial features. The selected features for FDD are features 9, 11, 12, 13, 14, 15, and 17 from Table 2.

Table 4: Cases for baseline establishment and features selected for each case

Cases	Upper limit	Lower limit
1	Max	Min
2	Max+0.5*SD	Min-0.5*SD
3	Max+1*SD	Min-1*SD
4	Max+2*SD	Min-2*SD
5	Max+3*SD	Min-3*SD

5. FDD RESULTS

A total of 14 faults were implemented among which 13 known faults along with normal operation period were used to train the FDD model. One of the faults (VFD pump stuck at 50% of its maximum speed) was treated as a new fault. Five different GGMR FDD models with corresponding features and the number of Gaussian components were trained using the criteria for the five cases discussed in

Table 4. The selection of the best case was done by evaluating the model training accuracy, false alarm rate, and fault misdiagnosis. All five cases had excellent model training accuracy as all the faults were trained with 100% accuracy.

All the faults except the ones presented in Table 5 did not have any misdiagnosed faults for all the cases. There were no false alarms when normal operation data was presented to the FDD model for any of the five cases. The space temperature negative sensor offset fault was misdiagnosed as supply water positive sensor offset fault (0.7%) and cool water valve stuck at 100% open fault (3.5%) in Case 1. It was misdiagnosed as the space temperature positive sensor offset fault (7.7%) in Case 5. The cool water valve stuck at 0% open was misdiagnosed as the hot water valve stuck at 100% open fault (5.6%) in Case 2 and Case 3. The cool water valve stuck at 50% open fault was also misdiagnosed in Case 2 and Case 3 as the cool water valve stuck at 0% open fault (20%). The cool water valve stuck at 100% open fault was misdiagnosed as a different fault in all 5 cases where the lowest misdiagnosis percent was for Case 1 and Case 4 (1.8%). Overall, Case 4 performed the best with the lowest misdiagnosis rates of the known faults. Hence, this baseline case was adopted to test the model in the presence of an unknown fault.

Table 5: Misdiagnosis of known faults during model testing

Faults	Case 1	Case 2	Case 3	Case 4	Case 5
F2	4.22	0	0	0	7.74
F11	0	5.6	5.6	0	0
F12	0	20	20	0	0
F13	1.796407	55.68862	55.68862	1.796407	16.16766

The performance of the FDD model determined using the criteria for Case 4 from

Table 4 was tested in presence of an unknown fault. All the unknown fault data was predicted as 0 which was not a label for any fault. As there was no information regarding the data presented to the model, the prediction was 0 for all data points. All the other faults had a testing accuracy of 100% except the cool water valve stuck at 100% open fault, which was predicted as normal operation (1.8%). Upon recognition of the type of the new fault by an expert, the data associated with the new fault was labeled (15) and fed to the evolving engine to update the Gaussian mixture regression model in the diagnosis section. Values for learning rate (0.01) and the closeness threshold (10) that resulted in the lowest prediction error were selected while evolving the GMR model. In the evolved diagnosis model, two new components were created to learn the information about the new fault. Before evolving the model, the new fault data fell into one of the 16 existing Gaussian components (Gaussian component 8) which was allocated to Fault 7 and Fault 8. After evolving the model, two new Gaussian components (17 and 18) were added to the model which had the information about the new fault.

The performance of the evolved FDD model was evaluated using the same dataset used to test the model before evolving. The results showed that the FDD model remains stable after being evolved since the accuracy of the true diagnosis lies in the range of 98% to 100% for known faults. Only 1.8% of the cool water valve stuck at 100% open fault was misdiagnosed as a normal operation. The evolved model also exhibits excellent performance in the diagnosis of the new fault as the accuracy of true diagnosis for the new fault was 100%.

6. CONCLUSIONS

The focus of this study was to develop a learning-based FDD model for radiant floor heating and cooling systems. An experimental study was conducted to investigate the performance of the learning-based GGMR model for fault detection and diagnostics of faults in an radiant slab system. A total of 6 common faults of different categories and severities (14 total faults) were implemented and were divided into known and new faults. The performance of the AFDD algorithm in the diagnosis of experimental fault data significantly improved after implementing the evolving process. A GGMR model was developed and trained with normal operation and 13 known faults data and was evolved using a new fault data. The performance of the model was tested using both known and new faults before and after evolving the model. The prediction accuracy for each fault using test data was extremely high with the lowest prediction accuracy of 98%. The algorithm was successful in detecting the new fault as an unknown state before evolving the model and in diagnosing it as a new fault after evolving the model. Further research can be conducted to evolve the model with the data from more than one new fault in the system. Another addition to this study would be to modify and test the model if simultaneous faults occur in the system. Further study can be conducted to test the performance of the model in presence of the data from the real buildings. This will also help to know if the model can be transferred to a real building setting without compromising the fault detection accuracy.

ACKNOWLEDGEMENT

This study was supported by the National Science Foundation EPSCoR Research Infrastructure program under Grant No. 1929209. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of* 199-213). Springer.
- Bouchachia, A., & Vanaret, C. (2011). Incremental learning based on growing gaussian mixture models. 2011 10th International Conference on Machine Learning and Applications and Workshops,
- Chakraborty, D., & Elzarka, H. (2019). Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold. *Energy and Buildings*, 185, 326-344.
- Fabrizio, E., Corgnati, S. P., Causone, F., & Filippi, M. (2012). Numerical comparison between energy and comfort performances of radiant heating and cooling systems versus air systems. *HVAC&R Research*, 18(4), 692-708.
- Fijany, A., & Vatan, F. (2006). A new efficient algorithm for analyzing and optimizing the system of sensors. 2006 IEEE Aerospace Conference,
- Joe, J., & Karava, P. (2019). A model predictive control strategy to optimize the performance of radiant floor heating and cooling systems in office buildings. *Applied Energy*, 245, 65-77.
- Karami, M., & Wang, L. (2018). Fault detection and diagnosis for nonlinear systems: A new adaptive Gaussian mixture modeling approach. *Energy and Buildings*, 166, 477-488.
- Katipamula, S., & Brambley, M. R. (2005). Methods for fault detection, diagnostics, and prognostics for building systems—a review, part I. *Hvac&R Research*, 11(1), 3-25.
- Sung, H. G. (2004). *Gaussian mixture regression and classification*. Rice University.
- West, S. R., Guo, Y., Wang, X. R., & Wall, J. (2011). Automated fault detection and diagnosis of HVAC subsystems using statistical machine learning. 12th International Conference of the International Building Performance Simulation Association.
- Wall, J., Guo, Y., Li, J., & West, S. (2011). A Dynamic Machine Learning based Technique for Automated Fault Detection in HVAC Systems. *Ashrae Transactions*, 117(2).
- Yu, Y., Woradechjumnroen, D., & Yu, D. (2014). A review of fault detection and diagnosis methodologies on air-handling units. *Energy and Buildings*, 82, 550-562.