Fault detection and diagnosis of air temperature sensors in an air handling unit using machine learning techniques

Behrad Bezyan, Radu Zmeureanu
Centre for Net-Zero Energy Buildings Studies, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada

Abstract

This paper presents the development of machine learning models for multiple fault detection and diagnosis of air temperature sensors of an air handling unit (AHU). For this purpose, air temperatures in critical points of an AHU such as mixed air temperature and temperature after the heating coil, are predicted. A compound artificial neural network (ANN) model is proposed for prediction of air temperature sensors, then the fault is detected if the differences between measured and expected values exceed the defined threshold. For fault diagnosis aspect, the Recurrent Neural network (RNN) as a deep learning model, and shallow Feedforward Neural network are developed for prediction of the air temperature value at the current time step (t) using the previous measurements of that target sensor; the faults are diagnosed if the residuals exceed the threshold. This paper uses the synthetic hourly data from the simulation of an institutional building with eQuest program as a proxy for real measurements. Models developed in this paper will be used in future work, and will be tested with real measurements for the multiple fault detection and diagnosis (MFDD) purposes.

Introduction

In the United States, almost 40% of the total annual energy use belongs to the building sector (Beiter, Elchinger, Tian, 2017). Therefore, it is so important to design and operate energy-efficient buildings. The building performance should be monitored in order to detect any potential fault in the system and diagnose the sources of malfunction in the heating, ventilation, and air conditioning (HVAC) systems.

About 15 to 30% of the energy in the commercial buildings would be wasted if HVAC systems are not maintained regularly or they are inappropriately controlled, and if the system degradation has not been detected at early stages (Katipamula and Brambley, 2005). In order to monitor the building energy performance and record HVAC system operation in real-time, the building energy management system (BEMS) should be used, and databases for the operation of each equipment should be maintained and used.

The malfunction in the system can be detected with the application of two fault detection and diagnosis approaches by using the historical databases of building performance (Katipamula and Brambley, 2005): 1) process history-based, and 2) qualitative model-based. History-based models are the data-driven models classified into supervised and unsupervised learning-based models that use historical measured data of the system operation. Qualitative models are the knowledge-based models of fundamental physics-based operation of the system (Zhao, Li, Zhang, Zhang, 2019).

Machine learning, which is a data-driven based method in the artificial intelligence field, gives the computer the ability to learn without being explicitly programmed (Samuel, 1959). An artificial neural network (ANN) is a machine learning technique that is used for prediction and forecasting purposes.

Many researchers used ANN models for the prediction of variables or indices of performance in building HVAC systems. Such models can be used to predict the reference or typical operating performance that is useful for fault detection. If the difference between the measurement in an HVAC system and ANN prediction exceeds the defined threshold, a fault is detected.

Some researchers have focused on the development of different ANN algorithms, such as single, auxiliary, and backpropagation ANN using MATLAB and Python. They considered the input time lags and the number of hidden neurons for prediction or forecasting sensor values of the HVAC system or building energy use, for instance, forecasting the electrical use in commercial buildings (Chae, Horesh, Hwang, Lee, 2016). Another example is the prediction of the sensor values such as air temperature in the AHU with a variable air volume valve (VAV) system for the fault detection and diagnosis purposes, and comparison with measurements (Lee, House, Park, Kelly, 1996). They also combined the model with the subtractive clustering method and concluded that the fault detection performance is acceptable (Du, Fan, Jin, Chi, 2014).

A sensitivity analysis was conducted in order to select the variables as the inputs which have the most impact on the accuracy of predictions (Fan, Du, Jin, Yang, Guo, 2010).

Adaptive ANN is another model that was used for the estimation the solar hot water system with application of the

generated data set from a TRNSYS simulation (He, Menicucci, Caudell, Mammoli, 2011), (He, Caudell, Menicucci, Mammoli, 2012).

Levenberg Marquart algorithm for ANN was applied to predict the values in a vapor compression refrigeration system (Koçyigit, 2015). The combination of ANN and fuzzy inference system found eight faults in the system.

Two hyperparameters, such as the number of hidden layers and number of hidden layer neurons, were considered in a deep neural network model for fault detection and diagnosis (FDD) by (Heo and Lee, 2018). They concluded that increasing the network size does not improve the fault detection accuracy at more than about 97.26%.

Different ANN algorithms, such as feedforward and cascade, were used for the prediction of outdoor air temperature using the measured data set from four European cities. With the comparison of predicted and measured outdoor air temperature, they concluded that the developed models have accurate prediction performance over the prediction time horizon of 4 to 24 hours, with RMSE less than 2°C and R^2 higher than 90% (Papantoniou and Kolokotsa, 2016).

Auto-Associative ANN (AANN) was developed by (Elnour, Meskin, Al-Naemi, 2020) for sensor prediction and then compared with PCA-based algorithm for fault detection and diagnosis. They concluded that AANN would improve multiple and simultaneous fault detection accuracy.

Geyer et al. (2018) proposed a component-based approach using ANN for prediction of energy performance in heating and cooling modes. They used the physical properties of building components as the inputs in each ANN of components in order to predict the conduction heat transfer rate. Furthermore, they used the summation of the outputs along with other input parameters as the input into the next ANN for the prediction of heating and cooling loads.

Recurrent neural networks (RNN) were developed for simultaneous fault detection and diagnosis of sensors in the HVAC system using simulation and real measured databases (Shahnazari, Mhaskar, House, Salsbury, 2019).

ANN models work very well for the prediction of air temperature and FDD in the AHU (Hou, Lian, Yao, Yuan, 2006). A shallow feedforward ANN is proposed for fault detection, as a robust machine-learning algorithm, for prediction of the sensor values in the AHU, as compared with other models represented in (Chae, Horesh, Hwang, Lee, 2016).

In this paper, first, a feedforward compound ANN model is used for fault detection in an air handling unit. A fault is detected if the difference between measurements of target values (e.g., the mixing air temperature) and the ANN predictions exceeds the pre-defined threshold. Second, a Recurrent Neural Network (RNN) model, using long short-term memory (LSTM) as a deep learning approach, is used for fault diagnosis. The RNN is developed for predictions of each regressor (e.g., the outdoor air temperature, the return air temperature) in order to compare the predicted values with the measurements, and if the difference exceeds a pre-defined threshold, then the corresponding sensor is defined as the faulty sensor.

A shallow feedforward neural network (FFNN) is also used for the fault diagnosis, and compared with the RNN model. Other machine learning models are also used such as support vector regression (SVR), decision tree, and random forest. Their prediction performance for air temperature is compared with ANN model. The proposed machine learning models are developed using Scikit learn (Pedregosa, Varoquaux, Gramfort, Michel, Thirion, Grisel, Blondel, Prettenhofer, Weiss, Dubourg, Vanderplas, Passos, Cournapeau, Brucher, Perrot, 2011), Keras (Chollet, 2015) and Tensorflow (Abadi, Barham, Chen, Chen, Davis, Dean, Devin, Ghemawat, Irving, Isard, 2015) which are the open source libraries in Spyder Python 3.8 (Python, 3.8.1).

Case study

In this study, synthetic hourly data from eQuest simulation program (Mihai, 2014) of an institutional building located in Concordia University, Montreal, over the heating season from January 2nd to March 31st are extracted and used for the models` development and validation. The dataset is available online on Github repository (Bezyan and Zmeureanu, 2020). Figure 1 shows the schematic of the AHU, and Table 1 lists the available variables for this study. The use of synthetic data replaces the measurements as the first step in the development and testing of prediction models as they are free of errors. The model prediction performance should normally be higher when using the synthetic data compared with the measurements in existing buildings. Therefore, these results should be considered as the upper threshold of performance.

The target values used for the fault detection are the mixed air temperature (*T_ma*), the air temperature after heating coil (*T_ahc*) and the supply air temperature (*T_sa*) from the AHU.

Table 1: List of AHU variables.

No.	Variable	Description	Unit
1	*T_oa_db*	Outdoor air dry-bub temperature	C
2	*V_oa*	Outdoor air volumetric flow rate	m^3/s
3	*T_ra*	Return air temperature	C
4	*V_ra*	Return air volumetric flow rate	m^3/s
5	*T_ma*	Mixed air temperature	C
6	*V_ma*	Mixed air volumetric flow rate	m^3/s
7	*T_ahc*	Air temperature after heating coil	C
8	*Q_HC*	Electric input for heating coil ($\dot{Q}$)	kW
9	$\Delta T\_s,fan$	Air temperature rise over supply fan	C
10	*T_sa*	Supply air temperature	C

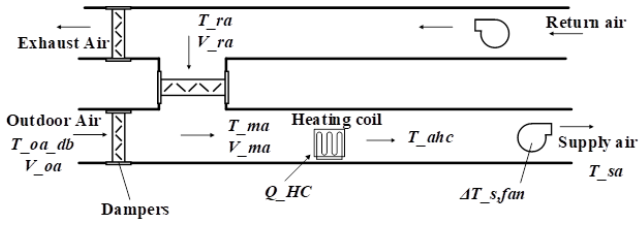

Figure 1: AHU schematic.

Method

Fault detection

A compound ANN architecture of the air handling unit (AHU) is proposed (Figure 2) that is composed of two ANNs. The ANN#1 predicts the mixed air temperature (*T_ma*) at time (t) using the related input variables at time (t). The output of ANN#1 is input in the ANN#2 that predicts the air temperature after heating coil (*T_ahc*) at the same time (t). Then the supply air temperature (*T_sa*) at that time *(t)* is predicted.

The measurements of regressors at current time *(t)* ($x_{1,m}^t$, $x_{2,m}^t$, …) are the inputs to ANN, and the target value (Y_p^t) at the same time *(t)* is the output (Equation 1).

$$Y_p^t = f(x_{1,m}^t, x_{2,m}^t, \dots, x_{n,m}^t) \quad (1)$$

Where, x_m is the measured values of the regressors at time *(t)*, and Y_p is the predicted target variable at time *(t).*

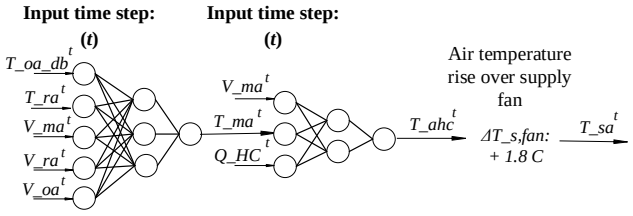

Figure 2: Compound ANN for prediction of T_ma and T_ahc in heating mode of AHU.

Feedforward ANN is used for both ANN#1 and ANN#2, using Python 3.8.

The following steps are taken:

Several ANN architectures are compared with the number of training days from 1 to 30, and the number of hidden layer neurons from 1 to 10; in all cases, there is only one hidden layer. The Sigmoid activation function is used for the hidden layer neurons, and the output node.

1. Each architecture is evaluated ten times; the average RMSE (between the predictions and measurements) of 10 times ANN architecture evaluations are recorded. The optimum ANN architecture, which requires the lowest number of training days with RMSE smaller than the overall uncertainty of the corresponding temperature sensor, is selected as the optimum ANN#1 model. In this study, the overall uncertainty of the air temperature sensor is 0.49°C (Cotrufo and Zmeureanu, 2018) based on previous measurements of the same AHU.

   The number of hidden layer neurons (*N*) is selected based on (Heaton, 2008).

$$N = \frac{N_{Input}+N_{Output}}{2} \quad (2)$$

   Where, N_{Input} is the number of inputs, and N_{output} is the number of outputs.
2. The optimum ANN#1 architecture for the prediction of *T_ma* over the time horizon of three days uses 72 hours of training data, and has three hidden layer neurons (Figure 3).
3. The mixed air temperature *T_ma* is predicted at the time *(t)* by using the regressors that are measured at the same time *(t)*. If the difference between the measurement of *T_ma* and the prediction of *T_ma* obtained with proposed ANN#1 does not exceed the threshold value, there is no fault with the *T_ma* sensor. Then, the measured *T_ma* is used as an input to ANN#2. While, if the difference exceeds the threshold value, there is a fault with the *T_ma* sensor, and the predicted *T_ma* is used as an input for ANN#2.
4. The same approach is used for the ANN#2, which is trained by using the measurements over the past 72 hours. The optimum ANN#2 for the prediction of *T_ahc* at the current time *(t)* requires 72 hours for training and two hidden layer neurons (Figure 3).
5. The air temperature *T_ahc* after the heating coil is predicted at the time *(t)* by using regressors *V_ma*, *Q_HC* that are measured at the same time *(t)*, and *T_ma* predicted or measured (as discussed at item 2 above). If the difference between the measured *T_ahc* and the ANN#2 prediction of *T_ahc* does not exceed the threshold value, there is no fault of *T_ahc* sensor. If the difference exceeds the threshold value, the predicted *T_ahc* is used as an input to predict *T_sa*.
6. The supply air temperature (*T_sa*) is predicted at the same time *(t)* by adding the air temperature rise in the supply fan (*ΔT_s,fan*) to the *T_ahc value*. In this study, *ΔT_s,fan*=1.8°C, based on previous measurements of the same AHU (Zibin, 2014).

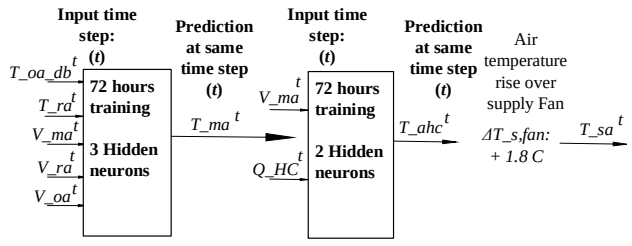

Figure 3: Compound ANN for prediction of T_ma and T_ahc in heating mode of AHU.

When a fault is detected with the sensors of *T_ma* and *T_ahc*, the fault diagnosis model is applied.

Fault diagnosis

Two different approaches are developed and compared for the multi-fault diagnosis: 1) Recurrent Neural Network (RNN) with use of long short-term memory (LSTM), and 2) Feedforward Neural Networks (FFNN). Both methods were developed in Python (3.8.1).

The application of fault diagnosis method is explained by assuming a case when an abnormal measurement was detected from the *T_ma* sensor at 2 P.M. One can suppose that the *T_ma* sensor is faulty, or one or more readings of the regressors (i.e., *T_oa, T_ra, V_ma, V_ra, V_oa*) (Figure 3) are faulty. The predicted values of each regressor at 2 P.M, which are obtained under normal operation conditions based on past monitored values, are compared with the measured values. If the differences between predicted and expected values for one or several sensors values exceed the defined threshold, then the faulty regressor sensor(s) is diagnosed. If none of the regressors are faulty, then it can be assumed that the *T_ma* sensor is faulty.

For fault diagnosis, a single network for prediction of each variable at time (t) is developed for each of FFNN and RNN models (Figure 4 and 5). The past monitored values of the target variable at (t-1, t-2, …, t-n) are used as the input values. For example, the FFNN model uses the recorded values of *T_ra* over the past five hours (t-1, t-2, …, t-5) to predict T_ra at time (t) (Figure 5).

Recurrent Neural Network (RNN)

The Recurrent Neural Network (RNN) that is a deep learning model with long-short-term memory (LSTM) architecture is applied to the normalized data set (Equation 3). The RNN model is developed and used for the prediction of X_p^t at time *(t)* by using the inputs of previous 10 hours measurements (i.e., *t-1, t-2…, t-10*) (Equation 4).

RNN uses the previous sequential information to learn and predict the present values based on the trained model. LSTM architecture has the chain like structure of the neural networks and able to learn the long-term dependencies. The LSTM is capable of adding, storing and removing the information which are helpful for the prediction (Hochreiter and Schmidhuber, 1997).

$$X_{normalized} = \frac{X - min(X)}{max(X) - min\,(X)} \quad (3)$$

Where X is the value at time *(t)*; $min(X)$ *and* $max(X)$ are the minimum and maximum values of the selected data set.

$$X_p^t = f(X_m^{t-1}, X_m^{t-2}, \dots, X_m^{t-10}) \quad (4)$$

where, X_m is the measured value at the previous time steps *(t-n)*.

The RNN model architecture, which was developed by trial and error, consists of one input layer, four hidden layers with 50 hidden neurons on each hidden layer, and one output layer. The Sigmoid activation function is used on the hidden layer neurons and the output layer node.

The RNN model is trained using a sliding window technique over 62 hours, to predict *T_ra* at time *(t)*. The RNN prediction model architecture of return air temperature *T_ra* is presented as an example in Figure 4.

The results of RNN model are compared with those from a shallow feedforward neural network (FFNN), which is presented in the following section.

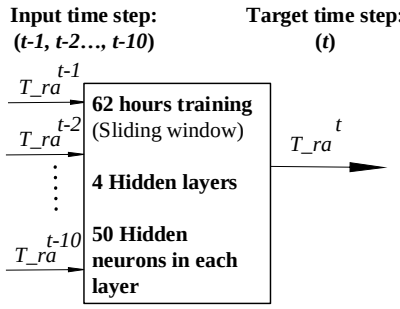

Figure 4: Recurrent Neural Network model for prediction of T_ra at time (t) in heating mode of AHU.

Feedforward Neural Network (FFNN)

The FFNN predicts X_p^t at time *(t)* by using the inputs of previous 5 hours measurements (i.e., *t-1, t-2…, t-5*) (Equation 5). The schematic of the FFNN for *T_ra* is illustrated in Figure 5.

$$X_p^t = f(X_m^{t-1}, X_m^{t-2}, \dots, X_m^{t-5}) \quad (5)$$

Where, X_m is the measured value of the regressor at previous time steps *(t-n)*.

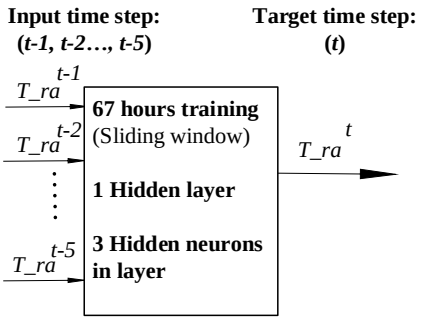

Figure 5: Feedforward Neural Network for prediction of T_ra at time (t) in heating mode of AHU.

The FFNN architecture, which was obtained by trial and error, consists of one input layer with five inputs, one hidden layer with 4 to 6 hidden layer neurons (depending on the regressor), and one output neuron.

Prediction performance

The performances of the proposed prediction models are evaluated with statistical indices (Equations 6 to 10): Coefficient of determination (R^2), Root-Mean-Squared-Error ($RMSE$), Mean Absolute Percentage Error ($MAPE$), Mean Bias Error (MBE), and Maximum Absolute Error (ME_{max}).

$$R^2 = \left[1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right] . 100 \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (7)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| . 100 \quad (8)$$

$$MBE = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \quad (9)$$

$$ME_{max} = |\hat{y}_i - y_i| \quad (10)$$

where, $\hat{y}_i$ is the predicted value, y_i is the measured value, and $\bar{y}$ is the average measured values of the variable within the selected period.

Fault detection and diagnosis performance

The performance of the MFDD models are evaluated by using the accuracy, precision, and sensitivity measures (Equations 11 to 13) (Olson and Delen, 2008). Accuracy measures the number of correct predictions of faulty and normal readings, respectively, over the total number of predictions. Precision measures the number of correct predictions of faults out of all predicted faulty readings. Sensitivity measures the number of faults predictions out of all actual faults.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (11)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

where, TP is number of true positives, e.g., the number of correctly predicted faults; FP is number of false positives, i.e., the incorrectly predicted faults; FN is number of false negatives, e.g., the incorrectly labelled data as no-fault; and TN is number of true negatives, e.g., the correctly labelled readings as no-fault. as represented in a confusion matrix (Table 2).

Table 2: Confusion matrix for fault detection

		Actual faults	
		Negative (0)	Positive (1)
Predicted faults	Negative (0)	*TN*	*FP*
	Positive (1)	*FN*	*TP*

Results and discussions

The selected compound ANN architecture (Figure 3) for the prediction of *T_ma* and *T_ahc* is applied over 24 hours, using sliding window technique, and the statistical indices of prediction performance are presented in Table 3.

Table 3: Predictions of the compound ANN model.

Mode	Target	ANN Architecture with one hidden layer and 72 hours of training		Prediction over validation data set for 24 hours				
		Input Variables	Hidden layer neurons	R^2 (%)	*RMSE* (C)	*MAPE* (%)	*MBE* (C)	ME_{max} (C)
Heating	*T_ma*	*T_oa_db, T_ra, V_ma, V_ra, V_oa*	3	99.06	0.053	0.28	0.04	0.10
	T_ahc	*Q_HC, T_ma, V_ma*	2	97.52	0.130	0.43	0.11	0.24

In addition to the ANN, three different machine learning models, e.g. support vector regression (SVR), decision tree, and random forest are developed for prediction of air temperature sensors (*T_ma* and *T_ahc*) over 24 hours. Their results are compared with those ANN (Table 4). The proposed ANN model has less RMSE compared to other models.

Table 4: Predictions performance of the compound ML models.

Target	Input Variables	Machine learning model using 72 hours training set	Prediction performance over validation data set for 24 hours				
			R^2 (%)	*RMSE* (C)	*MAPE* (%)	*MBE* (C)	ME_{max} (C)
T_ma	*T_oa_db, T_ra, V_ma, V_ra, V_oa*	ANN	99.06	0.053	0.28	0.04	0.10
		SVR	98.50	0.075	0.41	0.06	0.13
		Decision tree	98.13	0.077	0.32	0.05	0.16
		Random forest	94.85	0.135	0.60	0.09	0.39
T_ahc	*Q_HC, T_ma, V_ma*	ANN	97.52	0.130	0.43	0.11	0.24
		SVR	95.26	0.151	0.51	0.13	0.28
		Random forest	33.22	0.416	1.26	0.31	1.04
		Decision tree	75.90	0.473	1.50	0.37	1.07

Table 5: Feedforward Neural Networks prediction results.

Target		FFNN model		Prediction performance over validation data set for 24 hours					
Variable	Time step	Input	Time lag	Unit	R^2 (%)	*RMSE* (C)	*MAPE* (%)	*MBE* (C)	ME_{max} (C)
T_{oa}	t	T_{oa}	t-1, t-2…, t-5	C	67.80	1.36083	52.1779	0.9448	4.9319
T_{ra}		T_{ra}		C	85.75	0.020239	0.0742	0.0162	0.0436
V_{ma}		V_{ma}		m^3/s	100	0.000527	0.0195	0.0005	0.0005
V_{ra}		V_{ra}		m^3/s	100	0.001476	0.2194	0.0015	0.0015
V_{oa}		V_{oa}		m^3/s	100	0.000908	0.0447	0.0009	0.0009
Q_{HC}		Q_{HC}		kW	84.13	1.700485	3.8161	1.2275	4.8029
T_{sa}		T_{sa}		C	83.76	0.307907	0.8015	0.2178	0.7697

The performance indices of FFNN and RNN models are presented in tables 5 and 6, respectively. In the FFNN and RNN models, the target prediction variables are the air temperature sensors, volumetric air flow rate, and electric

input for heating at time (t) and the inputs for each model are the past monitored values of that target variable (t-1, t-2, …, t-n).

Both FFNN and RNN models give acceptable prediction performance in order to be deployed for the fault diagnosis approach. For example, for prediction of *T_ra* the RMSE and MAPE are approximately 0.0202°C, 0.074% using FFNN model, and 0.0207°C, 0.069% using the RNN model.

Table 6: Predictions of the Recurrent Neural Network.

Target		RNN model		Prediction performance over validation data set for 24 hours					
Variabl	Time step	Input	Time lag	Unit	R^2 (%)	*RMSE* (C)	*MAPE* (%)	*MBE* (C)	ME_{max} (C)
T_{oa}	t	T_{oa}	t-1, t-2…, t-10	C	49.42	1.39944	65.4873	0.9666	4.2991
T_{ra}		T_{ra}		C	84.90	0.02076	0.0692	0.0151	0.0557
V_{ma}		V_{ma}		m^3/s	100.00	0.00025	0.0091	0.0002	0.0002
V_{ra}		V_{ra}		m^3/s	100.00	0.00129	0.1913	0.0013	0.0013
V_{oa}		V_{oa}		m^3/s	100.00	0.00190	0.0935	0.0019	0.0019
Q_{HC}		Q_{HC}		kW	80.84	1.68295	3.6108	1.1756	4.7031
T_{sa}		T_{sa}		C	71.97	0.56523	1.5654	0.4175	1.2236

Artificial fault generation using grey-box models

In order to evaluate the MFDD methodology performance in the air temperature sensors of an AHU, artificial faults are generated in T_oa and T_ra sensors. A linear regression grey-box model is assumed for prediction of *T_ma* in terms of T_ra, and T_oa (Equation 14).

Another grey-box model is developed for prediction of air temperature after the heating coil (*T_ahc*) (Equation 15).

$$T\_ma = a.T\_oa + b.T\_ra \quad (14)$$

$$T\_ahc = \frac{c.Q}{d.V\_ma} + e.T\_ma \quad (15)$$

where, *T_ma* is the mixed air temperate (C), *T_oa* is the outdoor air temperature (C), *T_ra* is the return air temperature (C), *Q* is the electric input for heating (kW), *and V_ma* is the mixed air volumetric flow rate (m^3/s). Coefficients *a, b, c, d, and e* are obtained using the least square error method (LSEM).

Table 7 presents the identified coefficients of the regression models (Equations 14 and 15) along with the statistical indices for training and testing data sets.

Artificial faults have been introduced manually into the synthetic data of *T_ma* and *T_ra* from 11:00 to 24:00.

From 1:00 to 10:00, the readings are normal as the differences are below the threshold of 0.49°C. Faults start at 11:00 with negative drifting bias for *T_ra* and frozen reading in *T_oa* at the same time steps. Values for *T_ma* are generated using grey-box model 1 (Equation 14).

Table 7. Statistical indices of the developed models.

No.	Model	Training data set (24 hours in Jan.4)					Test data set (24 hours in Jan.5)	
		Parameter	Value	Unit	R^2 (%)	RMSE (C)	R^2 (%)	RMSE (C)
1	$T\_ma = a.T\_oa + b.T\_ra$	*a*	0.232	-	99.69	0.031	99.68	0.075
		b	0.747	-				
2	$T\_ahc = \frac{c.Q}{d.V\_ma} + e.T\_ma$	*c*	0.038	C/kW	99.87	0.091	98.65	0.124
		d	0.345	s/m^3				
		e	1.037	-				

In Figure 6, the differences between normal and faulty values for *T_oa*, *T_ra*, and generated *T_ma* are illustrated. The sensor uncertainty for the air temperature is defined as the threshold (0.49 C). If the difference between measured and expected (predicted) values exceeds the threshold a fault is detected.

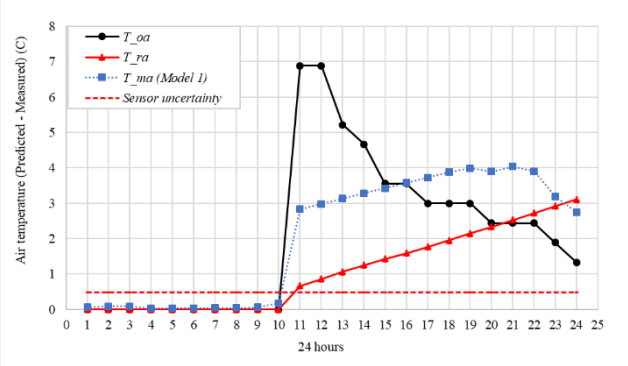

Figure 6: Difference between prediction and measurement of faulty T_oa, T_ra, and T_ma over validation data set.

The artificial faults with faulty values are generated for electric input of heating (*Q*), and then the new faulty data sets including the faulty values in addition to the generated values for *T_ma* using model 1 (Equation 14) are input into the model 2 (Equation 15) for generation of data set for *T_ahc*. Figure 7 depicts the differences between measured and predicted values with normal and faulty values of *T_ahc*, in addition to the sensor uncertainty which is a measure for fault detection threshold.

The predicted values of air temperature sensors using the proposed compound ANN models are applied for fault detection purposes. If the residuals exceed the threshold the fault is detected. As can be seen in Figure 7, the residuals for the *T_ma* and *T_ahc* is higher than the threshold starting from 11 to 24 which shows the detection of faults existence. The performance of fault detection and diagnosis are represented in Table 8. In addition, the confusion matrices for fault detection in *T_ma* and *T_ahc* are represented in tables 9 and 10.

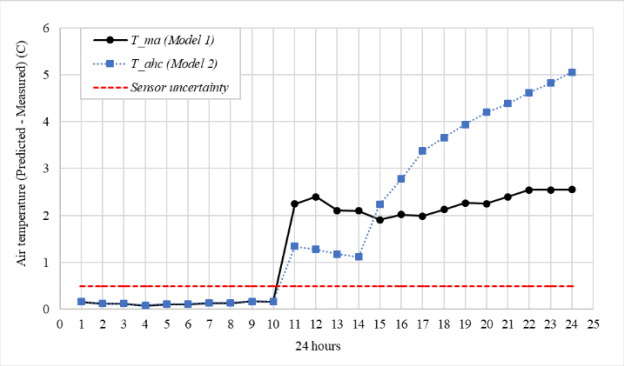

Figure 7: Difference between prediction and measurement of faulty T_ma, and T_ahc over validation data set.

Table 8. Accuracy, precision and sensitivity of the ANN models for MFDD.

Mode	Target	ANN Architecture with one hidden layer and 72 hours of training		MFDD over validation data set for 24 hours		
		Input Variables	Hidden layer neurons	Accuracy (%)	Precision (%)	Sensitivity (%)
Heating	*T_ma*	*T_oa_db, T_ra V_ma, V_ra, V_oa*	3	100	100	100
	T_ahc	*Q_HC, T_ma, V_ma*	2	100	100	100

Table 9: Confusion matrix for fault detection of T_ma

		Actual faults	
		Normal (0)	Faulty (1)
Predicted faults	Normal (0)	*10*	*0*
	Faulty (1)	*0*	*14*

Table 10: Confusion matrix for fault detection of T_ahc

		Actual faults	
		Normal (0)	Faulty (1)
Predicted faults	Normal (0)	*10*	*0*
	Faulty (1)	*0*	*14*

The proposed RNN algorithm using the LSTM and FFNN model are applied for fault diagnosis. Thereby, at the time steps when the faults were detected, the values of each regressor (*T_oa*, *T_ra*, *V_ma*, *V_ra*, *V_oa*) are predicted using the RNN and FFNN models with the sliding window technique as the prediction performances were represented in Tables 5 and 6. The abnormal readings of *T_ma* are detected after 11:00 when the differences between the predicted and measured values exceed the threshold. Then, the residual between the expected (predicted) and measured values of regressors, such as return air temperature (*T_ra*) are calculated, and where the residual is higher than the defined threshold, the faulty performance is diagnosed (Figure 6). Hence, it is concluded that the *T_ra* sensor is diagnosed as a faulty sensor which results in the abnormal readings of *T_ma* sensor.

Conclusions

In this paper, the development and application of multiple fault detection and diagnosis models for the air temperature sensors of an AHU were presented. For the fault detection, the compound ANN architecture was proposed for the prediction of mixed air temperature (*T_ma*) and air temperature after heating coils (*T_ahc*) at current time *(t)*. The differences between prediction and real measured values were calculated, and once the differences exceed the defined threshold (sensor uncertainty), the fault was detected.

The results reveal a good compound prediction model that can be used as a tool for fault detection of air temperature sensors of AHU in heating mode. The performance of the ANN was compared with three machine learning models; SVR, decision tree, and random forest. The results revealed ANN has better prediction performance and lower RMSE compared with other models.

For the fault diagnosis, the recurrent neural network (RNN) using the LSTM model was proposed for the prediction of every single regressor at the current time (t), using previous time steps observations. If the difference between measurement and prediction of any regressor at the time (t) is higher than the defined threshold, then the corresponding regressor sensor is diagnosed as faulty.

In addition to the RNN, the feedforward neural network (FFNN) model is used for the prediction of regressors at the time (t). The results proved that both proposed fault diagnosis models (RNN and FFNN) have good prediction performance for the scope of this study.

The use of real measurements for the development and validation of the proposed models for multiple fault detection and diagnosis applications is planned for near future studies.

Acknowledgement

The authors acknowledge the financial support from Natural Sciences and Engineering Research Council of Canada, and Gina Cody School of Engineering and Computer Science of Concordia University.

References

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard. (2015). *Large-Scale Machine Learning on Heterogeneous Systems*. Retrieved from TensorFlow: http://tensorflow.org/

Beiter, P., M. Elchinger, T. Tian. (2017). *2016 Renewable Energy Data Book.* United States: National Renewable Energy Lab. (NREL), Golden, CO (United States).

Bezyan, B., R. Zmeureanu. (2020). *Dataset of the Air handling unit*. Retrieved from Github: https://github.com/bbezian/Air-Handling-Unit.git

Chae, Y. T., R. Horesh, Y. Hwang, Y. M. Lee. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings, 111*, 184-194.

Chollet, F. (2015). Retrieved from Keras: https://github.com/fchollet/keras

Cotrufo, N., R. Zmeureanu. (2018). Virtual outdoor air flow meter for an existing HVAC system in heating mode. *Automation in Construction, 92*, 166–172.

Du, Z., B. Fan, X. Jin, J. Chi. (2014). Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis. *Building and Environment, 73*, 1-11.

Elnour, M., N. Meskin, M. Al-Naemi. (2020). Sensor data validation and fault diagnosis using Auto-Associative Neural Network for HVAC systems. *Journal of Building Engineering, 27*.

Fan, B., Z. Du, X. Jin, X. Yang, Y. Guo. (2010). A hybrid FDD strategy for local system of AHU based on artificial neural network and wavelet analysis. *Building and Environment, 45*, 2698-2708.

Geyer, P., S. Singaravel. (2018). Component-based machine learning for performance prediction in building design. *Applied Energy, 228*, 1439-1453.

He , H., T. P. Caudell, D. Menicucci, A. A. Mammoli. (2012). Application of adaptive resonance theory neural networks to monitor solar hot water systems and detect existing or developing faults. *Solar Energy, 86*, 2318-2333.

He, H., D. Menicucci, T. Caudell, A. Mammoli. (2011). Real-time fault detection for solar hot water system using adaptive resonance theory neural network. *ME 2011 5th International Conference on Energy Sustainability & 9th Fuel Cell Science, Wahington, DC, USA*, 1059-1065.

Heaton, J. . (2008). *Introduction to neural networks with Java. Heaton Research.*

Heo, S., J. H. Lee. (2018). Fault detection and classification using artificial neural networks. *IFAC PapersOnLine, 51*(18), 470-475.

Hochreiter, S., J. Schmidhuber. (1997). Long Short-Term Memory. *Neural Computation* , 1735-1780.

Hou, Z., Z. Lian, Y. Yao, X. Yuan. (2006). Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Conversion and Management, 47*(15-16), 2479-2490.

Katipamula, S., M. R. Brambley. (2005). Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I. *HVAC&R Research, 11*(1), 3-25.

Koçyigit, N. (2015). Fault and sensor error diagnostic strategies for a vapor compression refrigeration system by using fuzzy inference systems and artificial neural network. *International Journal of Refrigeration, 50*, 69-79.

Lee, W.Y., J.M. House, C. Park, G.E. Kelly. (1996). Fault diagnosis of an air-handling unit using artificial neural networks. *ASHRAE Transact, 102*, 540-549.

Mihai, A. . (2014). *Calibration of a Building Energy Model Using Measured Data for a Research Center.* Montreal: MASc thesis- Concordia University.

Olson, L.D., D. Delen. (2008). *Advanced Data Mining Techniques.* Heidelberg: Springer.

Papantoniou, S., D. D. Kolokotsa. (2016). Prediction of outdoor air temperature using neural networks:Application in 4 European cities. *Energy and Buildings, 114*, 72-79.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12* , 2825-2830.

Python . (3.8.1). Retrieved from https://www.python.org/

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 210-229.

Shahnazari, H., P. Mhaskar, J. M. House, T. I. Salsbury. (2019). Modeling and fault diagnosis design for HVAC systems using recurrent neural networks. *Computers and Chemical Engineering*, 189-203.

Zhao, Y., T. Li, X. Zhang, C. Zhang. (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.

Zibin, N. . (2014). *A Bottom-Up Method to Calibrate Building Energy Models Using Building Automation System (BAS) Trend Data.* Montreal: MASc thesis - Concordia Unversity.