Modeling of Chilled Water VAV Systems Using a Machine Learning Approach

Rand Talib[1], Priyan Rai[2], Nabil Nassif[3], and Mostafa Tahmasebi[1]

[1]Graduate student, Department of Civil and Architectural Engineering and Construction Management, University of Cincinnati, Cincinnati, OH, USA
[2]Undergraduate student, Department of Electrical Engineering and Computer Science, University of Cincinnati, Cincinnati, OH, USA
[3]Associate Professor, Department of Civil and Architectural Engineering and Construction Management, University of Cincinnati, Cincinnati, OH, USA

Abstract: *In 2017, about 39% of total U.S. energy consumption was consumed by the residential and commercial sectors according to EIA. Thus, developing methods to lower the energy consumption of buildings becomes more crucial. This paper introduces a predictive model based on the Machine Learning algorithm of aggregated bootstrapping and compares it to a more commonly used Vector Machine Model along with an Artificial Neural Network Model. The models predict the supply air temperature as a function of chilled water temperature, chilled water valve position, mixed air temperature, and supply air flow. the models that were created were trained and developed using a real data collected from a real building. The results of this study have legalized the use of bootstrap aggregation as a powerful tool in predicting the performance of a chilled water AHU. The accuracy of the testing and training results were within 95-99.9% depending on the model type.*

Keywords: Bagged trees, ANN, Bootstrap aggregation, AHU, Cooling coil.

Introduction

With increasing of the awareness of the effect of global warming and energy cost, the need for more efficient buildings has become urgent (Talib et al, 2018). Moreover, quotes for the most recent Residential Energy Consumption Survey (RECS) demonstrate that 48% of the energy consumption in 2009 in the U.S homes was consumed for heating and cooling. However, it was stated that this percentage was drooped down from 58% in 1993. Buildings in the US. Consume around 48% of the nation total energy. Around 55% of it is dedicated to operating the HAVC systems (Pérez-Lombard et al, 2008). Heating ventilation and air conditioning systems are the mechanical systems that aim to heat or cool the air and maintain the human comfort of the building. HVAC systems are a complex nonlinear- systems that has different variables as the parameters of that system. The chilled water system is on the most common HVAC systems that is widely used in the commercial, industrial, and multi-story buildings. The chilled water system is currently used in different type of buildings due to its high efficiency and large available different applications that will allow the designer to easily control the different building zones. Moreover, today most of the buildings are equipped with BAS (Building Automation System). BAS can record a large amount of the building performance data that can be later used for the energy performance assessment (Nassif 2014). However, even with these latest technologies, most of the buildings still does not operate efficiently due to the shortage of the computational modeling technique that can be used to better predict the energy performance in the building and help in reducing the energy consumption of it. With increasing the awareness of the consumption of the HVAC systems and the cost that is associated with it, it is crucial to implement the intelligent models that can help save on the cost and energy of the building. The intelligent machine learning models that have the ability to predict the performance of those mechanical systems are proven to be reliable in helping to design a more energy efficient and a sustainable building. Thus, to reduce the energy consumed by buildings many strategies have been developed (ASHRAE, 2011). In this paper a bootstrap aggregation that is also called bagging, which is a machine learning technique, will be used to create several models that will be using the same dataset to train. This technique combines the output prediction of several models and produce an output that is more accurate than any output individually (Breiman, 1996). Some of the models that can be used for the bagging technique are artificial neural networks, classification and regression trees, and subset selection in linear regression. Finally, models that are derived based on collected real data have proven to be a powerful tool in predicting the performance of HVAC systems and explain the inner relationship of the system components. Also, models are used to achieve the indoor air quality (IAQ) especially as it is related to the health and comfort of the inhabitant (EPA, 2018).

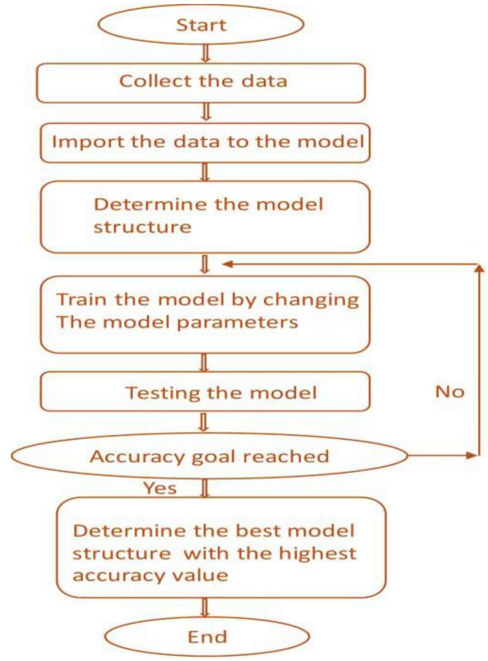


Figure 1 Workflow diagram

Methodology and Data Collection

The goal of this research is to create different predictive models using the same dataset for training and then comparing the output of the models and decide on the optimal model. Figure 1 shows the steps that were followed in creating the models. Moreover, the data that where collected in order to conduct this research are real data from an existing building located in Greensboro, North Carolina covering 88000 sf^2.It is a three story multi uses building. The building is equipped with a building automation system to record the performance of the building. The layout of the BAS is showing in figure 2. The building mechanical system is furnished with VAV systems. The layout of that mechanical system consists of -6- Air Handling Units (AHUs) and chilled water central plant with two chillers. The layout of each air handling unit is showing in Figure 3. The arrangement of each AHU consists of supply and return fan, exhaust, outside air dampers, heating and cooling coils, VAV boxes and multiple zones layout. The temperature range that was covered is at a

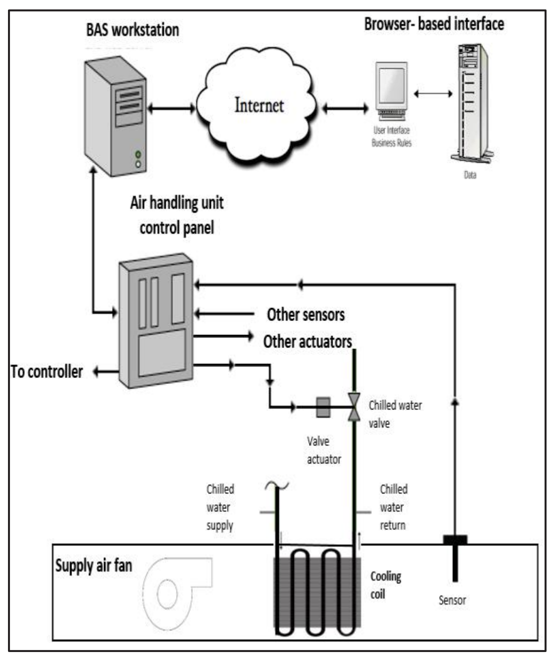


Figure 2 Building Automation System (BAS) component.

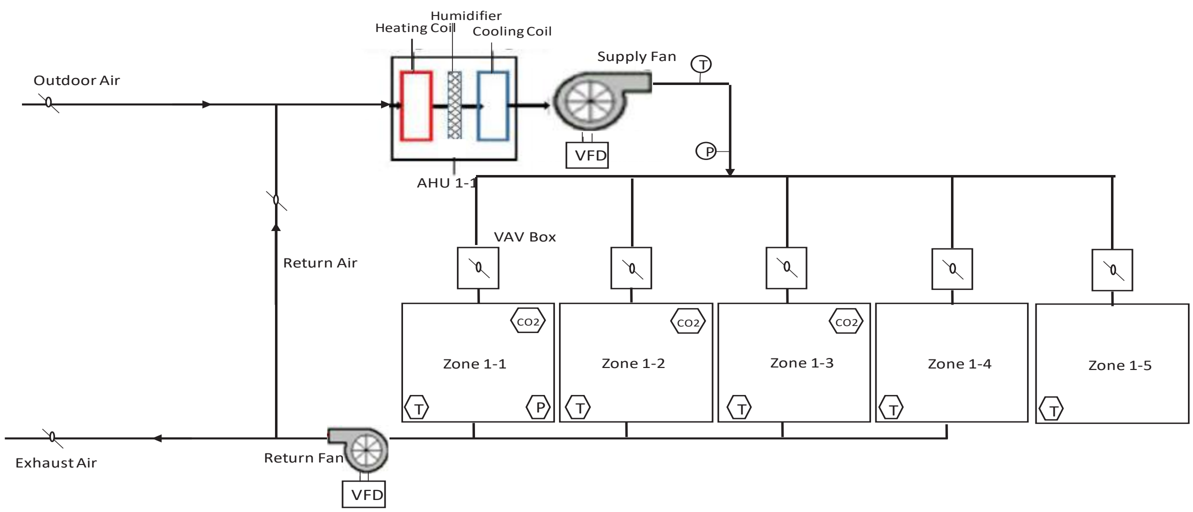


Figure 3 A schematic of the components inside of a CHW AHU.

minimum of 50˚F and maximum of 65˚F for the supply air temperature. The building is designed to have a Supply Air Temperature (SAT) of 55˚F. However, when recording data by the building automation system one thing must be noted that the BAS is recording data even when the system is turned off during the weekends. These data will show as a zero in the measurement and it will affect the accuracy of the study. Thus, those recorded data have to be removed so it will not affect the learning abilities of the machine learning models that will be later created. Finally, the collected data were transformed into an Excel sheet and were divided into two sets. (1) is the training set, from November 1st to December 31st and (2) testing set was conducted from January 1st to February 1st.

Comparison of Estimation Models

Table 1 shows three different predictive models that were trained on four of the following inputs:

Table 1 Models discussed in paper

Name	Number
Support Vector Machine	Model 1
Artificial Neural Network	Model 2
Aggregated Bootstrapping	Model 3

(1) Chilled water temperature, (2) Chilled water valve position, (3) Mixed air temperature & (4) Supply air flow. Supply Vector Machine (SVM) is one of the methods that use supervised learning used for classification and clustering purposes. In general, SVM is also extended to solve regression problems, and thus support vector regression. Model 1 utilizes a cubic SVM kernel function with principal component analysis (PCA) where all four training components are kept by PCA to explain variance. The main idea of principal component analysis is to reduce the dimensionality of a data set consisting of many variables correlated with each other, either heavily or lightly, while retaining the variation present in the dataset, up to the maximum extent. Model 2 utilizes an Artificial Neural network that used the inputs for training as a formatted matrix. An infinite variety of network architectures can be used for this purpose, but the simplest structure was needed to be considered that could be efficient but could still keep a high accuracy. Figure 4 shows the model that was finally designed for optimal accuracy. Four input layers, 4 hidden layer neurons and 1 output layer neuron.

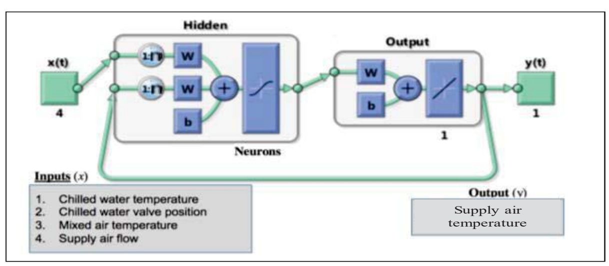


Figure 4 Neural Network Design.

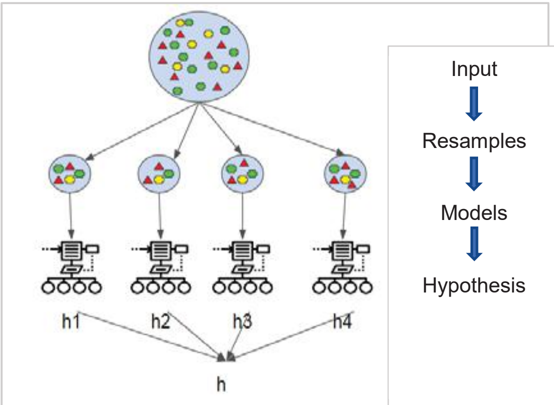


Figure 5 Working of Bootstrap Aggregation.

Machine learning provides algorithms that utilize more advanced and reliable techniques to produce more accurate results (Rai, 2019). Model 3 is trained using a machine learning algorithm known as bootstrapping, which simply put is the method of random sampling with replacement. Such a sample is referred to as a resample. This allows the model or algorithm to get a better understanding of the various biases, variances and features that exist in the resample. Since the data was measured on very short intervals of time, the data set was large. When this occurs, the regression coefficients represent the noise rather than the genuine relationships in the population making predictions inaccurate. To prevent overfitting, therefor this is prevented by bootstrap aggregation or Bagging. The proposed Model 3 even after training, can be supplemented by further data points without disturbing the pre-existing training samples. This allows the model to continuously improve its accuracy over time. The models are compared in terms of classification accuracies. Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions that the model got right. Formally, accuracy has the definition as shown in Eq.1. For binary classification, accuracy can also be calculated in terms of positives and negatives as follows where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives as shown in Eq.2.

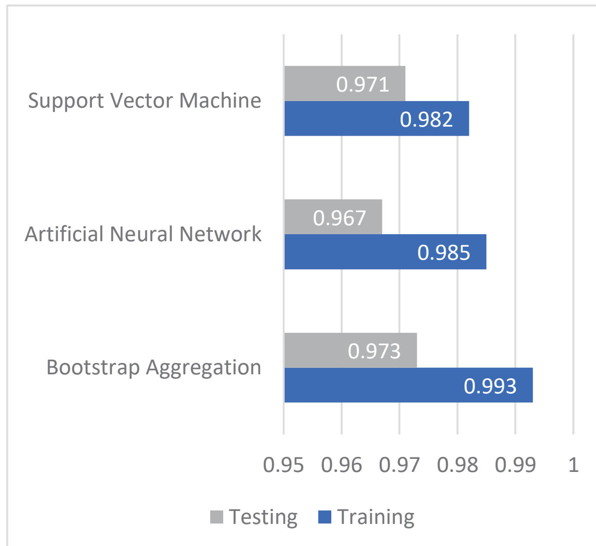
$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Predictions} \quad \text{Eq.1}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Eq.2}$$

Results for Model Training & Testing

It can be seen in Table 2 that all the models provide a highly accurate prediction of the Supply Air Temperature based on its inputs. Model 3 based on Bootstrap Aggregation achieved the highest training as well as testing accuracy at 99.3% and 97.3% respectively.

Table 2 Comparison of Model Accuracies.



While, table 3 shows that Model 2 based on Artificial Neural Network had the lowest training time at 341.3 seconds (less than 6 minutes). While Model 3 had the highest accuracy, its training time is considerably higher than the rest. Also, the training time have increased with increase in training set size. Model 1 has a balance of ~97% percent accuracy along with a moderate training time.

Table 3 Comparison of Model Training Times.

Number	Training Time (s)
Model 1	1349.3
Model 2	341.3
Model 3	2335.1

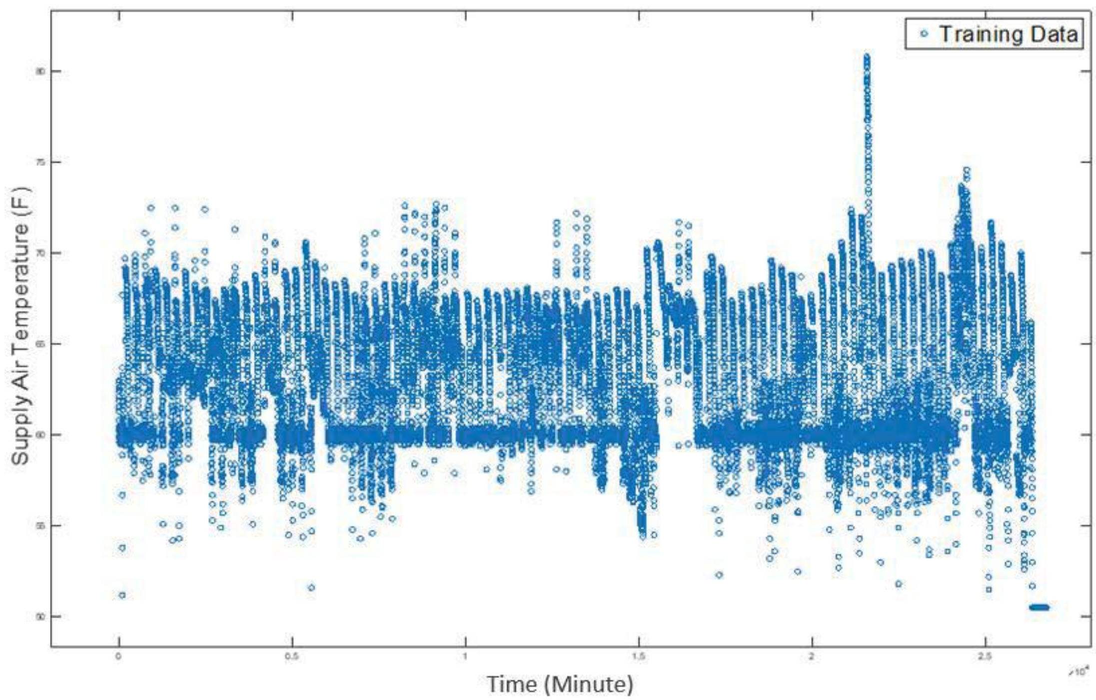


Figure 6 Training Set with all 26793 SAT Output value.

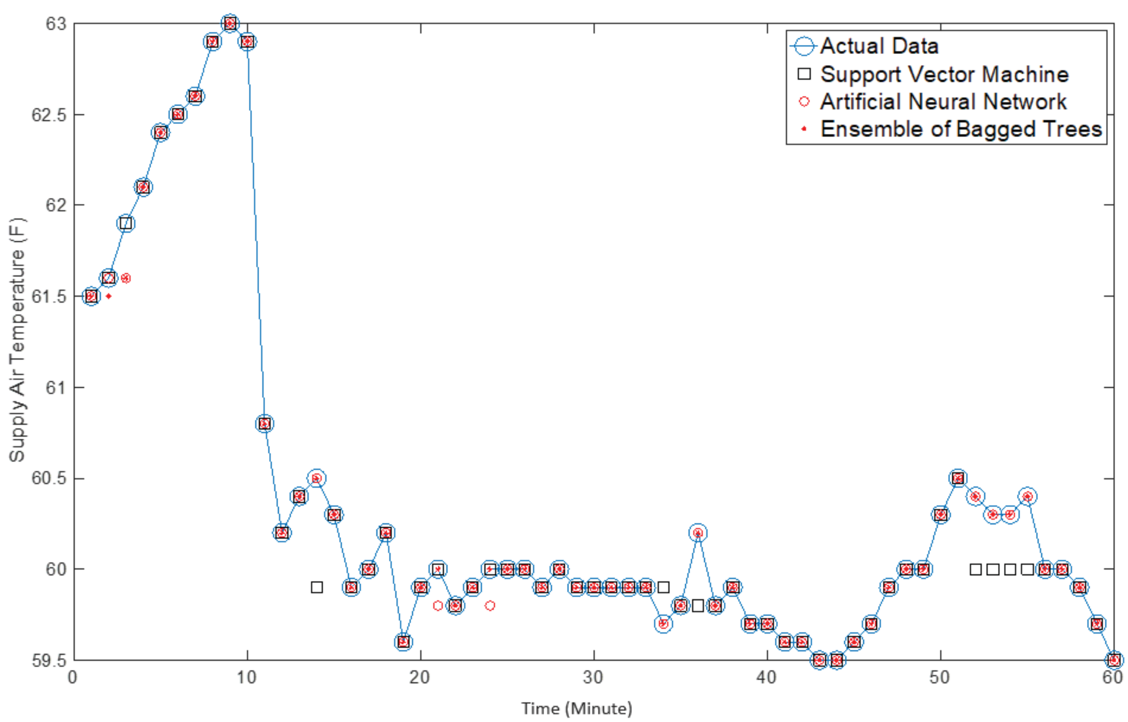


Figure 7 Comparison of actual supply air temperature over a period of one minute with outputs from all discussed models.

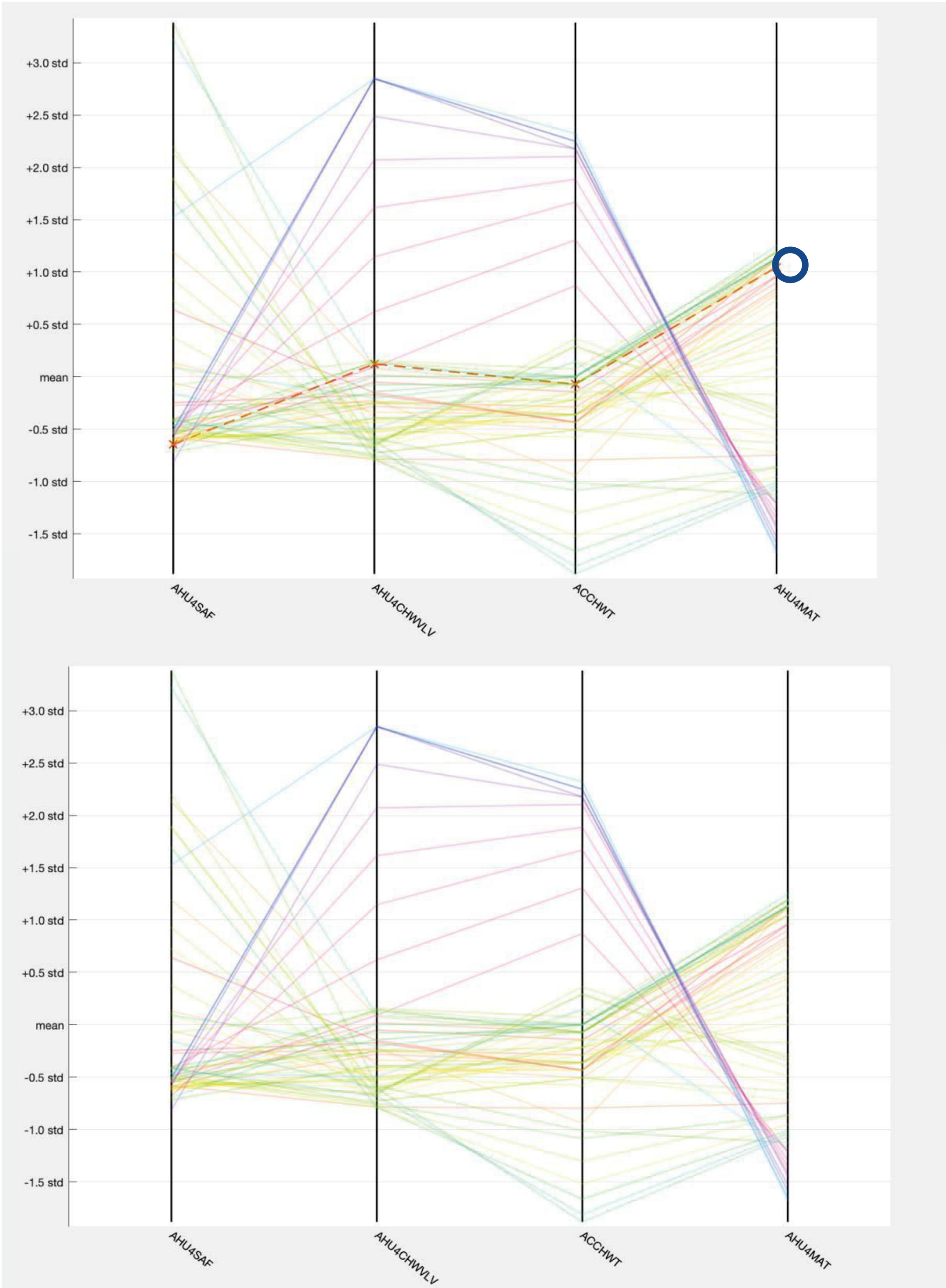


Figure 8 Parallel Coordinates plot for Model 1 and 3 respectively depicting training relations between data points among the inputs. Ideal training sample has only one error point for Model 1 and none for Model 3.

Figure 6 shows the training set with all the 26793 point of output values. While, figure 7 shows the prediction from all the models alongside the actual data over a period of a minute. This chosen prediction set had a range of prediction values therefore allows every model to be tested for a diverse set of inputs. Most of the data points overlap the actual data line i.e. the model was able to make a successful prediction for that set of inputs.

Conclusion

In order to validate the effectiveness of each model, an experimental data set was created with inputs which are supposed to produce a wide array of output values. While, the ANN model along with the Bagging Model had a 100% validation accuracy, the SVM ended up with one error point as shown in Figure 8. All models had a remarkably high accuracy in terms of their predictive capabilities ranging from 95-100% with Model 3 achieving the highest at 99.3%.

Author's Contributions

Authors contributed in the design concept, data-analysis and contributed to the writing of the manuscript.

Dr. Nassif has supervised the research and directed the authors through the research.

Rand Talib has layout the concept of the research and the main idea. She worked on collecting the data and transforming it, proposing the model input and outputs, the types of models used, and the models' structure.

Priyan Rai designed and programmed the models, conducted results, and constructed the tables for results evaluation.

Mostafa Tahmasebi helped throughout the research process.

References

[1] ASHRAE. "ASHRAE Handbook Applications". Chapter 41. Atlanta: American Society of Heating Refrigeration and Air Conditioning Engineers, Inc. Print. 2011.

[2] ASHRAE 62.1. "Ventilation for Acceptable Indoor Air Quality". American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc. 2016.

[3] EIA. "How Much Energy Is Consumed in Residential and Commercial Buildings in the United States?". U.S. Energy Information Administration – Independent Statistics and Analysis. Web. 2017.

[4] EPA. "Introduction to Indoor Air Quality". United States Environmental Protection Agency. Web. 2018.

[5] Leo Breiman. Bagging Predictors. Technical Report No.421. 1996.

[6] Nassif, N. "Modeling and Optimization of HVAC systems using Artificial Neural Network and Genetic Algorithm". International Journal of Building Simulation, 7 (3), 237.245. 2014.

[7] Pérez-Lombard, L., Ortiz, J., & Pout, C. "A review on buildings energy consumption information". Energy and buildings, 40 (3), 394-398. 2008.

[8] Rai. P, Nassif. N, Eaton. K & Rodrigues. A, ERJ. "Applications of Machine Learning in Building Energy Predictions and Savings". 2019.

[9] Talib, R., Nassif, N., Arida, M., & Abu-Lebdeh, T. "Chilled Water VAV System Optimization and Modeling Using Artificial Neural Networks". 2019.

[10] Werbos PJ. "Beyond regression: new tools for prediction and analysis in the behavioral sciences". [Ph.D. thesis]. Cambridge, (MA): Harvard University, 19 (1). 1974.