Journal of Construction Engineering, Management & Innovation
2022 5(2):119-139
DOI 10.31462/jcemi.2022.02119139

RESEARCH ARTICLE

Comparing the accuracy of ANN and ANFIS models for predicting the thermal data

Sami Shams Aldin, Hatice Sözer

Istanbul Technical University, Energy Institute, Istanbul, Türkiye

Article History

Received 27 May 2022
Accepted 24 June 2022

Keywords

Building energy performance prediction models
The accuracy of predictions
ANN
ANFIS

Abstract

The study aims to propose a suitable prediction model to deliver the full heating season's thermal performance dataset by using short-term measured data during the system operation period. Two machine learning-based models, BackPropagation Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System are compared by utilizing the measured data of indoor temperature and relative humidity. The independent variables of the prediction are obtained from the weather data, in addition to the building energy simulation model. Conversely, the data of the dependent variable are obtained from the real measurements from inside of the building for 31,5 days of the heating season, starting from February 22nd, which is called the first heating season. Moreover, the entire heating season of the building is evaluated between November 15th and March 21st, which is called the second heating season when the building's monthly consumption exceeds 14 kW/m^2. The first prediction approach is the feed-forward Artificial Neural Network (ANN) with Back Propagation Learning System (BPS). Four ANN models are structured by input-output and one hidden layer is performed. The second prediction approach is the Adaptive Neuro-Fuzzy Inference System (ANFIS). The Sugeno ANFIS method is utilized in this prediction work. Eight ANFIS models are structured by 6 layers are performed to achieve the prediction. Besides, the main motivation for approaching ANFIS is to avoid the stochasticity of the measured temperature and humidity data. The prediction results are compared with the measured data of the second heating season. The comparison showed that the ANFIS model is more efficient since it achieved an 85% accuracy rate for the indoor temperature and 81% for the humidity prediction. While the ANN prediction accuracy is 81%, 80% relatively for the temperature and humidity. Then the comparison is scaled by selecting the most ordinary period in the measured data to be the data sample that will be used in the comparison. The second comparison showed that the ANFIS model is once again better than the ANN model since the ANFIS prediction accuracy becomes 88% for temperature and 90% for humidity, while the ANN prediction accuracy becomes 83% for temperature and 87% for humidity. Nevertheless, the stochasticity of the measured affected the prediction results in accuracy rates. Hence, according to the achieved accuracy rates, both the ANFIS and ANN approaches are highly validated in this type of prediction.

Correspondence Hatice Sözer sozerh@itu.edu.tr

1. Introduction

Providing satisfactory thermal comfort conditions have a significant impact on buildings' energy consumption since the cooling and heating consumption correspond to almost haft of the overall energy consumption of the buildings [1]. Therefore, research that focuses on indoor thermal conditions would have multiple objectives as improving the energy performance of the buildings relatedly reducing the emissions.

Further, the thermal environment quality is highly related to human health and productivity since it has a direct effect on physical and psychological conditions [2]. Henceforward, measurement and verification of the building's indoor thermal conditions such as temperature and humidity have been investigated extensively in the last decades to sustain the quality of the indoor environment.

Measurement and verification of the indoor thermal comfort conditions to evaluate the building energy performance might take up to 2 years to cover the whole heating and cooling seasons. It requires detailed data on hourly, daily, monthly, and seasonal based on the indoor thermal environment. This long period cause inefficiency in the measurement and verification process. Granting, examining the thermal environment of the building during the operation of the heating and cooling systems provides the opportunity to report and fix any problem that the poor system performance may cause.

Predicting the indoor thermal conditions of a building by using short-term measured data could be an efficient way of investigating the building's thermal environment if the measurements were performed during the building system in operation. This predicted data could be a useful tool to support energy management in the building.

Artificial intelligence-based algorithms had been utilized widely in prediction works and studies. Machine learning could be effective since it doesn't need special infrastructures to be performed. The Artificial Neural Network algorithm (ANN) is one of the most used machine learning algorithms. ANN mimics the human brain process to convert the information and experiences into decisions, it had been utilized widely in both classification and prediction studies because of its highly accurate results and its ability to correlate the non-linearity between the variables.

On the other hand, the Adaptive Neuro-Fuzzy Inference System (ANFIS) is an algorithm that combines the Fuzzy approach and the ANN algorithm in order to perform better accurate results. Both ANFIS and ANN approaches have been used in both the classification and prediction studies in the literature.

The purpose of this work is to define the best effective model in terms of prediction accuracy rate. The focus persists on predicting the building's heating performance by using short-term measurements. A big-scale residential building was employed for measuring and collecting the short-term data along with developing the energy model for analyzing the heating season's consumption. The prediction is performed by the application of both the ANN and ANFIS approaches. The predicted data of ANN and ANFIS models are also validated by comparing them with the measured datasets.

1.1. Literature review

Research on predicting data by utilizing machine learning and deep learning approaches is applied to the diverse fields that have different objectives from weather data forecasting to currencies changes' estimations. There are altered approaches that provide precisely evaluated predictions that have been widely applied. ANN and ANFIS two well-known approaches were utilized to be tools to support the evaluation of diverse problems in literature.

For example, Piri et al. predicted the daily pan evaporation by developing two-hybrid models of the Cuckoo optimization algorithm (COA) with ANN and adaptive neuro-fuzzy inference system (ANFIS) with ANN. Further, their performances were compared with single ANN and ANFIS. The results didn't give a noticeable enhancement of combining the COA with ANN and ANFIS techniques. Consequently, hybridizing the COA with ANN and ANFIS cannot be a feasible option for estimating the daily evaporation. The ANN model provides better accuracy for the prediction of daily evaporation [3].

In another field, Golafshani et al. utilized both methods for predicting the compressive strength of normal and High-Performance Concretes. ANN and ANFIS techniques were hybridized by Grey Wolf Optimizer (GWO) for predicting the compressive strength (CS) of Normal Concrete (NC) and High-Performance Concrete (HPC). They concluded their study by indicating that the hybridization of the models with GWO improves the training and generalization capability of both ANN and ANFIS models [4].

Hajiand and Payvandy applied ANN and ANFIS to assess the ability of these methods for predicting the color strength of plasma-treated wool yarns dyed with a natural colorant. As a result, ANFIS had higher accuracy based on the obtained correlation coefficients [5].

Shahrak et al. functioned in ANN and ANFIS modeling for predicting the efficiency of water vapor adsorption capacity in porous metal-organic framework materials (MOF). In the model, the input parameters were selected as the surface area, pore-volume, and pore diameters and the outputs were calculated as the water vapor adsorption capacities of MOFs. The results show that real experimental data entailed the advantage of the ANFIS and ANN models to predict the water vapor adsorption capacity into MOFs with a mean squared error (MSE) of 0.005 and 0.002. This reveals a great potential for the application of both ANN and ANFIS methods to quickly monitor the MOF's suitability for water vapor adsorption [6].

Similarly, Sahoo et al. on the other hand utilized both ANN and ANFIS to predict the green sand mold permeability. Even though, the predicted permeability by both models was found to be very close to experimental values; the predictability of the ANFIS model was better because the error percent was less [7].

Respectively, there is some research just to compare the accuracy of multiple prediction models. Poul et al, have done a comparative study of multi-linear regression (MLR), K-nearest neighbors (KNN), ANN and ANFIS models. In the research, the prediction of the monthly flow in the St. Clair River was done by applying the MLR as a statistical method, ANN and ANFIS as non-linear ones, and KNN as a non-parametric regression. Results indicated that the implementations of three nonlinear models of ANN, ANFIS, and KNN are extremely convincing, the ANFIS model which benefits from the advantages of both fuzzy inference systems and neural networks was excellent. On the other hand, the MLR models that attempt to generate a linear relationship between the inputs and outputs were not successful to estimate the monthly flows accurately. KNN models were close to the ANN and ANFIS models, further, it consumes a simple structure and has an easy learning procedure [8].

Likewise, Taşan and Demir compared MLR, ANN, and ANFIS models for predicting the field capacity and permanent wilting point for Bafra plain soils. In this study, the performance of MLR, ANN, and ANFIS methods with different input parameters in the prediction of field capacity and the permanent wilting point from easily obtained soil characteristics were compared. Validation results revealed that the ANN model with the greatest (Coefficient of determination) R2 and the lowest MAE and RMSE value exhibited better performance for the prediction of FC and PWP than the MLR and ANFIS models [9].

Moosavi et al. comparatively used ANN-MLP, ANN-RBF, ANFIS, and GMDH methods to predict the thermal conductivity enhancement of nanofluids. In this work, four types of data mining methods, namely adaptive neuro-fuzzy inference system, artificial neural network—multilayer perceptron algorithm (ANN-MLP), artificial neural network—radial basis function algorithm (ANN-RBF), and group method of data handling (GMDH) have been used to predict the enhancement of the relative thermal conductivity of a wide range of nanofluids with different base fluids and nanoparticles. The total number of experimental data used in this work is 483 from 18 different nanofluids. The input parameters are thermal conductivity of base fluid and nanoparticles, volume fraction percent, the average size of nanoparticles, and temperature. Although the results showed that all four models are in relatively good agreement with experimental data, the ANFIS method is the best. Accordingly, the ANFIS method can able us to predict the relative thermal conductivity of new nanofluids in different conditions with good accuracy [10].

To extend the research further on building energy consumption, ANN has been experimented with to predict indoor comfort conditions.
Sozer and Sems took advantage of ANN by utilizing short-term monitored data to predict the whole heating season's data to be used in a real-time calibration process. The experiment was developed for a big-scale building and a detailed building energy model was developed. The results of the prediction were used as feedback data to improve simulation accuracy [11].

Comparing the prediction results of both ANN and ANFIS models for indoor thermal data in terms of accuracy would be a sufficient method to understand the ability of these machine learning models to provide accurate predictions in a such study. Different comparative studies for ANN and ANFIS prediction accuracy also have been introduced in the literature. Masoudi et. al. developed ANN and ANFIS models to predict the temperature in order to understand the effect of machining process parameters on temperature change. They found that the ANFIS model results, when compared to experimental data, were more accurate with a 3.17% error rate than the ANN model results which had a 5.83% error rate [12]. Riahi-Madvar and Seifi used 360 data points for 21 parameters to develop ANN and ANFIS methods aiming to predict the bedload transport in gravel-bed rivers. Their results showed that both ANFIS and ANN models are suitable for this kind of prediction and proved the superiority of the ANFIS model [13]. The same year, Soni published a comparison study between ANN and ANFIS predictive models developed to predict the turnover of the employees. The results of this study showed that the sensitivity of the ANN model is higher and it is more suitable for this prediction work since the RMSE of the ANN model was 0.1 for the training dataset and 0.4 for the unseen dataset while the RMSE for the ANFIS model were 0.19 for the training dataset and 0.5 for the unseen dataset [14]. In this study, ANN is compared with ANFIS to provide better accuracy in predicting the heating season indoor thermal comfort data. This prediction aims to provide a full heating season's thermal comfort dataset by using short-term measured data while the heating system is performing. Even though there are other comparative studies of ANN and ANFIS in the literature, their application of them to a building's thermal performance, in which the accuracy of prediction is crucial, is not realized yet. Furthermore, both modeling results are compared by scaling their accuracy, which would be considered a valuable outcome of this study.

2. Methodology

The prediction of the indoor thermal temperature and humidity data is performed by two supervised machine-learning algorithms ANN and ANFIS. Both of the algorithms need a sufficient amount of historical data to be trained to provide an accurate prediction. In addition, the period of the data to be predicted must be clearly defined. Fig. 1 shows the Methodology's steps followed.

2.1. Data gathering

Data gathering is a deliberate cycle to gather and measure the required information of the factors to be utilized in the forecast model. For the supervised machine learning predictive models, the quality and quantity of the available data directly affect the result of the prediction. The available data in this study was gathered from three different sources measurements, simulation, and weather stations.

2.2. Evaluating the heating season

The heating season of the building is the time period during which the prediction models are performed. To be able to customize the heating season of the building this study depended on the simulation result of heat consumption by setting a value that the heating period starts when the heating consumption exceeds it.

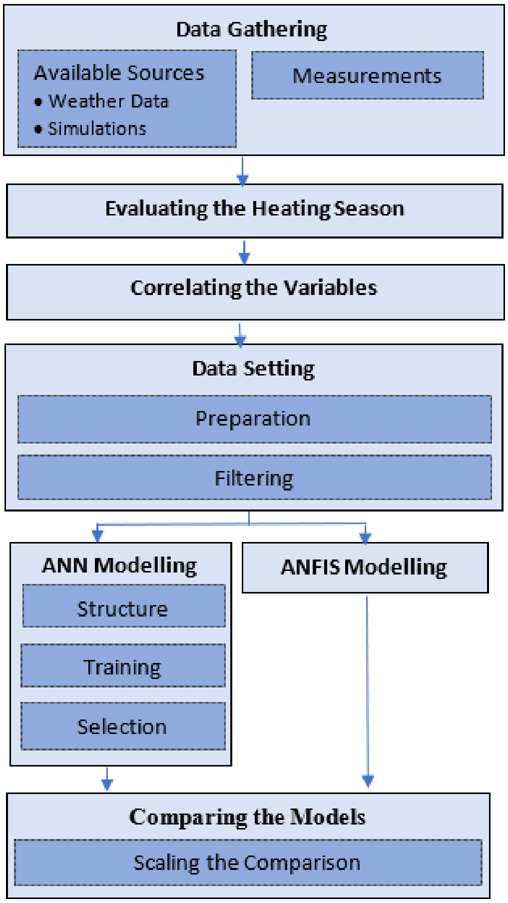

Fig.1. Methodology flowchart

2.3. Correlating the variables

It is the process to discover the linear relationship between the input and targeted variables. It is important that the alterations in each input variable correlated to the alterations of each targeted variable in the correlation analysis. This correlation generates a coefficient to definite the dependency between those two variables. The coefficient value is between 1 and -1. The relationship of the variables is direct when the coefficient value is 1, and it is reversed when the coefficient value is -1, and there is no relation when the coefficient value is 0. Since all of the gathered data are continuous the linear correlation, the coefficient formula in Eq. 1 is applied for the correlation process.

$$r(xy) = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} \quad (1)$$

2.4. Data setting

Data management is a crucial task to establish the base for the prediction model. The precision of the collected data leads to an accurate model. There are several constraints for preparing and unifying the gathered data such as initial time, frequency, unit, and structure.

Thenceforward, the considered data must be filtered to avoid or reduce the noises by un-using the instances which have unexpected values out of the range of the maximum and minimum values that are specified by the Eqs. 2-3:

$$Max = \mu + 1.5 * \sigma \quad (2)$$

$$Min = \mu - 1.5 * \sigma \quad (3)$$

where: μ: the mean – σ: the standard deviation.

2.5. ANN modeling

ANN is a method, often employed in various fields to correlate data and variables that do not have a clear algorithm to solve or link to each other to predict their outputs. The ANN model possesses different types, as the most utilized one is the linear multi perceptron. Besides, the most introduced model in the prediction exertion is the Feedforward Backpropagation model.

The composition of the structure of the ANN is very important that must be defined by the setting of the various parameters: the number of the hidden layers and the order of the nodes in each layer in addition to the method of the loss index error, learning rate, and distributing the datasets for training and selection phases.

The ANN model has been organized for training by setting the number of iterations and defining the activation function between the layers, which defines the numerical calculations between the neurons of the network. The network calculations will depend on the sigmoid (logistic) activation function to avoid the non-linearity between the variables. Sigmoid activation function could be expressed by Eqs. 4-5:

$$\text{sigmoid}(x) \text{ or } \text{sig}(x) = 1/(1 + e\text{^}(-x)) \quad (4)$$

$$Zj = \text{sig}(\sum(xi \times wij) - \theta) \quad (5)$$

where: Zj is the set that is received by the artificial neuron, xi is the input value, wij is the weight and θ is the use of a threshold.

Then, the assortment is the process of measuring the features and performance of the model. By a defined number of iterations, the ANN compares the selection data prediction results with its actual targeted data and calculates the losses of the model, then it improves the parameters of the model until it reaches the minimum losses. Then the final structure of the model would be defined and the last training performed would be iterated.

2.6. ANFIS modeling

Adaptive Neuro-Fuzzy Inference Systems interpolate between multi-linear equations in order to model an accurate nonlinear system. The ANFIS modeling has been done based on the Sugeno fuzzy inference [15] method since its computation process is more efficient and tighter than its counterpart Madami fuzzy inference [16] method. The Sugeno method has the ability to produce the best fuzzy modeling performance by customizing the membership functions through adaptive fuzzy model construction techniques.

To fuzzy, the data Takagi-Sugeno fuzzy if-then rules were used to perform a first-order Sugeno model. The ANFIS rules and its x and y inputs and f output are represented in Eqs. 6-7:

Rule (1): If x is A_1 and y is B_1,
then $f_1 = p_1x + q_1y + r_1$ (6)

Rule (2): If x is A_2 then y is B_2
then $f_2 = p_2x + q_2y + r_2$ (7)

where A_1 and A_2 are the input membership functions (MFs) for the input layer, B_1 and B_2 are the input membership functions (MFs) of y. The output function parameters are p_1, q_1, r_1, p_2, q_2, and r_2.

The ANFIS model is structured by the following five layers:

Layer 1: This layer produces the membership grade of the input variable in each node. Eq. 8 represents the membership functions values for each $i\ th$ node:

$$Qi^1 = \mu Ai\ (x) = \frac{1}{1+\left[\left(\frac{x-Ci}{ai}\right)^2\right]^{bi}} \tag{8}$$

where x is the input to node i and Ai if the linguistic label associated with this node function, ai,bi,ci is the parameter set that changes the shapes of the membership function.

Layer 2: In the second layer the value of each node multiplies by input weight, as shown by the Eq. 9:

$$Qi^2 = wi = \mu Ai\ (x)\ \mu Ai\ (x) \times \mu Bi\ (y),\ I = 1, 2 \ldots \tag{9}$$

Layer 3: This layer is responsible for the normalized firing strength for the membership values in node $i\ th$ by Eq. 10:

$$Qi^3 = wi = \frac{w1}{(w1+w2)}\ i = 1,2 \ldots \tag{10}$$

Layer 4: This layer represents the input-output values model which is represented by the Eq. 11:

$$Qi^4 = wi\ (pi\ x + qi\ y + ri) \tag{11}$$

where wi is the output from layer 3 and pi, qi, and ri are the design parameters that are determined during the training process. Parameters in this layer will be referred to as consequent parameters.

Layer 5: The fifth layer unifies the output in a single node by summation of the incoming values from nodes in the previous layer by the Eq. 12:

$$Qi^5 = \frac{\sum iwi\ fi}{\sum iwi} \tag{12}$$

The learning rule of ANFIS is the same as the back-propagation learning rule used in the common feed-forward neural networks. The optimization parameters are, bi ,ci which are the premise parameters, while pi ,qi ,ri are the consequent parameters. A hybrid-learning rule was employed in this research, which involves gathering the gradient descent and the least-squares method in order to find the appropriate set of preceding and consequent parameters [17]. The advantage of using a hybrid-learning rule was that it also seemed to be significantly faster than the classical back-propagation method [18]. Fig. 2 shows the structure of ANFIS.

The hybrid-learning procedure includes two passes, namely the forward pass and the backward pass. In the forward pass, the functional signals will go forward till layer 4 and the least-squares technique will identify the consequent parameters.

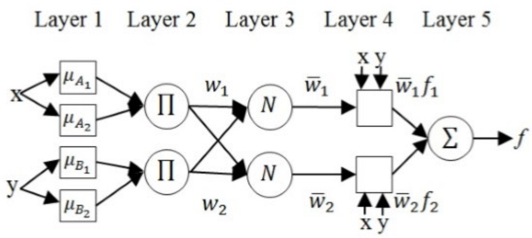

Fig. 2. Structure of ANFIS [19]

In the backward pass, the error rates transmit backward and the gradient descent will update the premise parameters. While the values of the premise parameters are fixed, it's possible to express the overall output as a linear combination of the consequent parameters [20].

2.7. Comparing the models

The prediction accuracy between measured and predicted data of both the ANFIS and ANN models are evaluated. the comparison is scaled by selecting the most ordinary period of the measured data as a sample to be compared with the predicted data and then providing the accuracy based on the selected sample.

3. Case study

The case building is utilized as an elderly home and has 8 stories with an 18,108 m^2 conditioned floor area, located in Kartal, Istanbul. Thermal comfort conditions were studied during the heating season. Fig. 3 demonstrates the image of the building.

3.1. Data gathering

The initial stage of the process is to define the accessible data sources. Because the prediction models are highly dependent on the quality and quantity of the available data.

3.1.1. Measurements

The measurements were the essential source of the data. In the building, four points were selected to measure the temperature and relative humidity for the period of a year to cover the dates between February 22nd of 2018 to February 29th of 2019.

Four sensors were allocated to the selected four points for measurements in the basement, on the third floor, in the lobby of the ground floor level, and on the first-floor levels. The sensors collected the data for every 15 minutes interval. Those four points are demonstrated in Fig.4.

3.1.2. Weather data

The targeted variables that are dependent, the temperature and humidity data were collected. Furthermore, prediction models need independent variables as input. The independent variables are available data, which are non-linearly correlated with dependent variables besides being available for the whole period to be predicted. The outside weather data is a prominent variable that has a great impact on the indoor thermal environment. Additionally, the outside dry-bulb temperature, outside dew-point temperature, wind speed, wind direction, atmospheric pressure, and solar azimuth could be obtained from the weather data.

3.1.3. Energy performance simulation

The heating system was operated during the measured period, therefore, heating consumption with related data sources was required to be included in the prediction model to provide a balance between the input datasets with the real case and to achieve better prediction in terms of accuracy. An energy performance simulation model was developed for the building by utilizing the DesignBuilder® software [21]. The simulation provides diverse categories of predicted results, one of those is the hourly heating consumption which can be used as an independent variable in the prediction model.

3.2. Evaluating the heating season

The scope of the prediction is to cover the indoor thermal data of the heating season. As known, the heating season is distinctive for each region, besides presenting diversity based on the building characteristics. For the case building, the heating period was defined by tracking to building's heating (gas) consumption as the period exceeds 14 kW/m^2 per month. As represented in Fig. 5, that is the point where the consumption raised dramatically. Thus, the heating period was defined roughly between the 15^{th} of November and the 31^{st} of March (137 days)

Fig. 3. Kartal elderly home

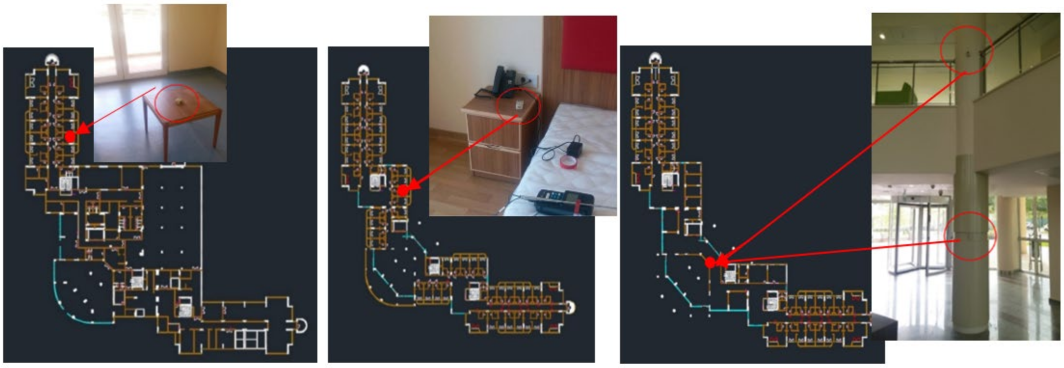

Fig. 4. Measurements points: Basement point, 3rd-floor point, and lobby points

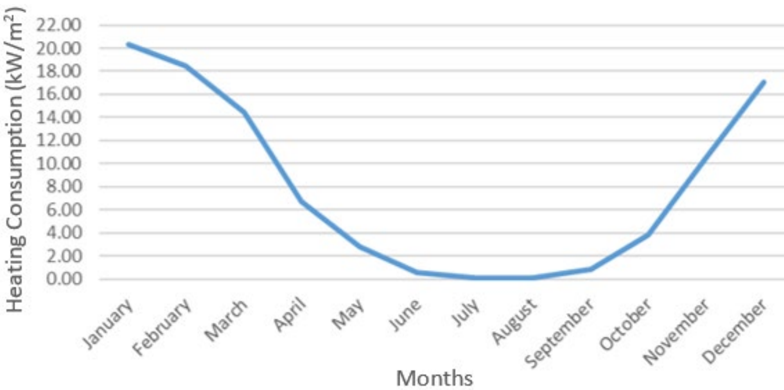

Fig. 5. Monthly heating consumption graph from the building energy model

3.3. Data setting

3.3.1. Data preparation

The measured and simulated data were collected within different structures and frequencies. The datasets had been unified in terms of frequency and prepared to be distributed in one datasheet in an appropriate form to be imported to the prediction models.

3.3.2. Data correlation

After collecting the available data and preparing it to be usable and comparable, the independent variables and dependent correlation were analyzed to discover the numerical impacts of the independent variables on the dependent outputs. Table1 represents the correlation factor of each of the independent variables with each of the dependent variables. The results of the correlation analysis showed that each of the selected variables was correlated to the dependent variables with an assorted effect that cannot be formulated linearly.

Table 1. Input-targeted variables correlation

Independent Variables (Inputs)	Basement room: Inside humidity rate	Basement room: Inside temperature	Upper room: Inside humidity rate	Upper room: Inside temperature	Upper Lobby: Inside humidity rate	Upper Lobby: Inside temperature	Ground lobby: Inside humidity rate	Ground lobby: Inside temperature
Outside Dry-Bulb Temperature	0.057	0.0925	0.151	0.205	0.385	0.428	0.29	0.392
Outside Dew-Point Temperature	0.239	0.251	0.437	0.456	0.461	0.542	0.416	0.57
Wind Speed	0.0969	0.272	0.084	0.134	0.368	0.135	0.272	0.0419
Wind Direction	0.0216	0.0469	0.0686	0.116	0.039	0.115	0.00865	0.0896
Atmospheric Pressure	0.464	0.367	0.392	0.375	0.0288	0.0324	0.203	0.246
Solar Azimuth	0.0265	0.0294	0.0252	0.039	0.00152	0.0377	0.0571	0.0916
Heating (Gas)	0.0224	0.000904	0.00826	0.0189	0.0195	0.0506	0.006	0.0515

Dependent variables (Outputs)

3.3.3. Data filtering

The outstrip points in the measured data that make noise for the model were filtered respectively, 68 data sets in the basement's point, 88 data sets in the upper room, 138 data sets in the upper lobby, and 178 data sets in the ground lobby. The considerable amount of data sets was reduced while accuracy was increased.

3.4. ANN modeling

3.4.1. ANN structure

Four Artificial Neural Networks were operated to predict thermal comfort data. Each network was built to characterize one of the four points (Basement, Upper room, Upper lobby, and Ground lobby). Each network was structured to have three layers, the first layer is the input layer, which contains 7 neurons that represent the independent variables. The second layer is the hidden layer, which contains 3 neurons. The third layer is the output layer, which are 2 neurons representing the targeted data (Temperature and Relative humidity) for each point that was represented in Fig. 6.

3.4.2. Training the artificial neural network

To start with, the first process was to match the frequency of the hourly weather data with the heating consumption data where the measured data was prearranged and sorted in an hourly data format. This process was, which was also minimized the amount of data to be trained. As a result, only 753 measured datasets for each point in the heating season were able to be trained. Afterward, the obtained data were scaled by means of the Minimum-Maximum method to be utilized in the activation function. Based on the short-term measured data, 85% of the datasets were used for the training phase, which used the Quasi-Newton algorithm and sigmoid activation function to go from the input to the hidden layer and linear function to move to the input layer. Python code was used to perform the ANN model.

3.4.3. Selection

In pursuit of training the ANN, 15% of the data was exploited into the selection method of 10 iterations to test the performance of the model and the parameters. The selection phase has changed the structure of the Network by increasing the hidden nodes to 8 nodes and minimizing the losses of the model. By using the selection phase's results, the model was trained for a second time to achieve the minimum losses as represented in Table 2. The final structure of the ANN model was represented in Fig. 6.

Table 2. Final losses

	Basement room	Upper room	Upper Lobby	Ground Lobby
Final Losses	0.295	0.369	0.38	0.56

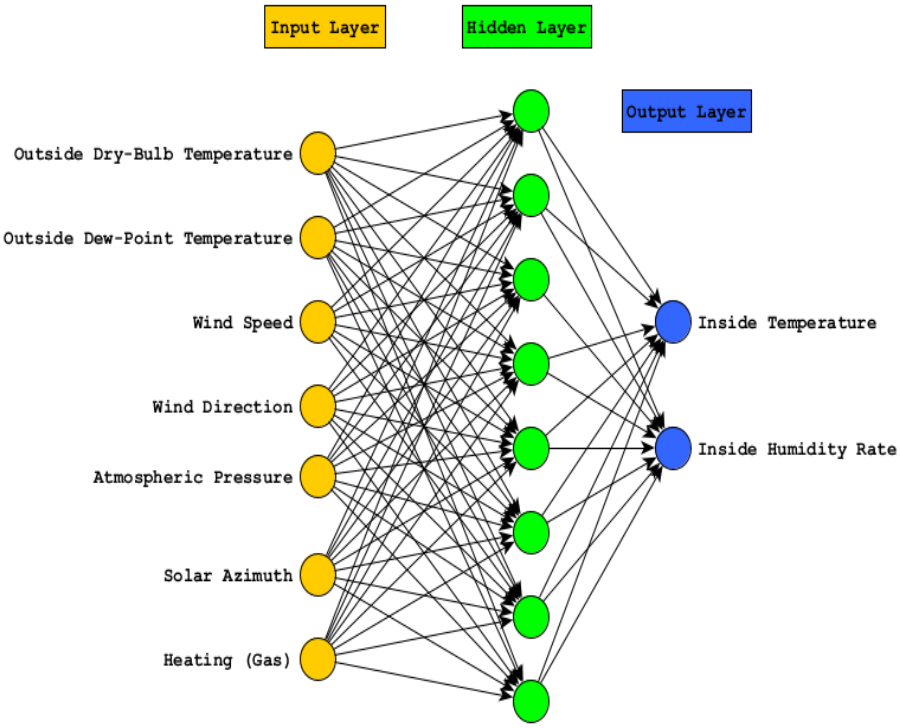

Fig. 6. ANN final structure

The performed ANN showed that filtering the data reduced the obscured results of the model and improved the accuracy. Furthermore, the number of hidden layers, which altered through the selection analysis, had a significant effect on the final losses.

3.5. ANFIS modeling

3.5.1. ANFIS structure

Eight adaptive neuro-fuzzy inference system models had been developed to predict the indoor temperature and the relative humidity of the building. Each model was used to predict one of the targeted parameters in one of the four points. Each of the utilized models is structured by six layers, the first layer is the input layer which includes 7 nodes each node representing one of the independent variables. The second layer is the input membership function layer, this layer contains 14 adaptive nodes, each pair of adaptive nodes receives the value of one of the independent variables to use as input in its function. The third layer includes fixed nodes that receive signals from the input layer, the output of this layer is the product of the received signals and it's called the firing strength of the rules. The fourth layer is the normalization layer, in the nodes of this layer the ratio of each rule's firing strength has been calculated. Layer 5 is the output membership function layer, and the last layer is the single output layer which is the temperature or relative humidity for each point. The ANFIS final structure is shown in Fig. 7.

3.5.2. Training the adaptive neuro-fuzzy inference system

The data set of the ANFIS is the same asthe ANN model, which means that the available datasets to be trained are 753 hourly datasets for the heating season.

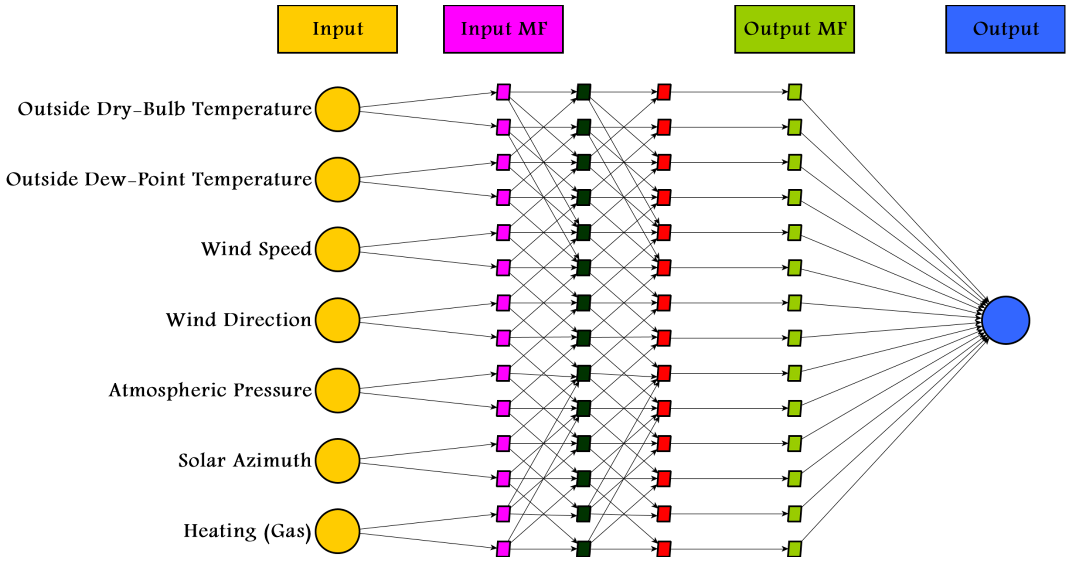

Fig. 7. ANFIS final structure

The ANFIS model utilized the Sugeno method which has output membership functions as explained in the methodology section in detail. The hybrid-learning algorithm's feedforward backpropagation procedures had been used as a learning algorithm for the 20 iterations ANFIS model. The number of iterations was defined by the testing phase since the datasets were distributed in 85% for the training phase and 15% for the testing phase. Matlab's ANFIS tool was used to perform the model.

3.6. Comparing the models

In this part, the ANFIS and ANN prediction results will be compared with the second-year heating season measurements. This comparison will show whether the approaches are applicable to predict the indoor thermal data, and which one is more appropriate in terms of accuracy.

Figs. 8-15 show that both the predicted data of ANN and ANFIS are relatively matching the measured data. For both temperature and humidity, The ANN predicted results' variation seems to be more realistic and closer to the measured data. While the ANFIS results seem to be more stable with fewer variations.

In Fig. 8, it is recognized that in the period between December 27th and January 1st the measured data seems to be far less than the average, in some points it went down less than 10 °C, which shows that the heating system wasn't properly performing. This deviation in some periods may cause mistakes in estimating the errors of the prediction work. It will decrease the accuracy of the prediction models since the errors were calculated by comparing the predicted data with the measured data. Therefore, the prediction model will carry responsibility for the heating system's bad performance.

Table 3, shows that the ANFIS model prediction of the temperature has less than the ANN model prediction in 3 points out of four. The ANFIS temperature prediction results' RMSEs (Root Mean Square Error) are 4, 2.9, 2.7, and 1.4 respectively in the basement room, upper room, upper lobby, and ground lobby, while in the same order the ANN prediction errors are 5.2, 3.9, 3.8, and 1.5.

It is significant the small error of both of the models in the ground lobby. This can be explained by observing the measured temperature dataset in Fig. 16. The measured data at this point is more stable with fewer variations compared to other points. Even though the average measured temperature is 20.74°C, which is less than the comfort range, the stability of the measured data in the point shows that the heating system is performing properly in the lobby.

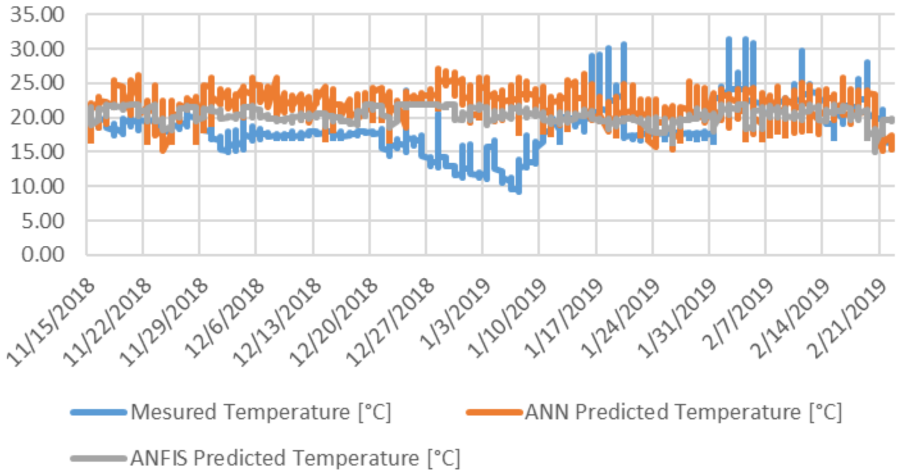

Fig. 8. Basement - Hourly predicted and measured temperature

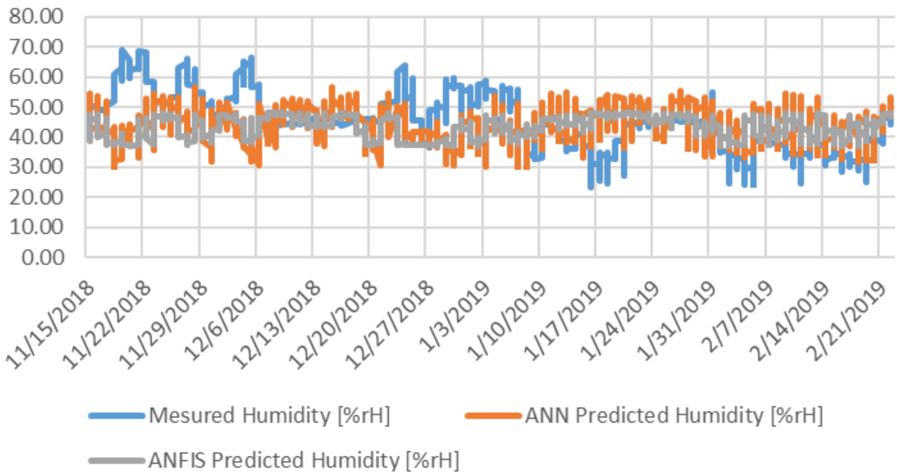

Fig. 9. Basement - Hourly predicted and measured humidity

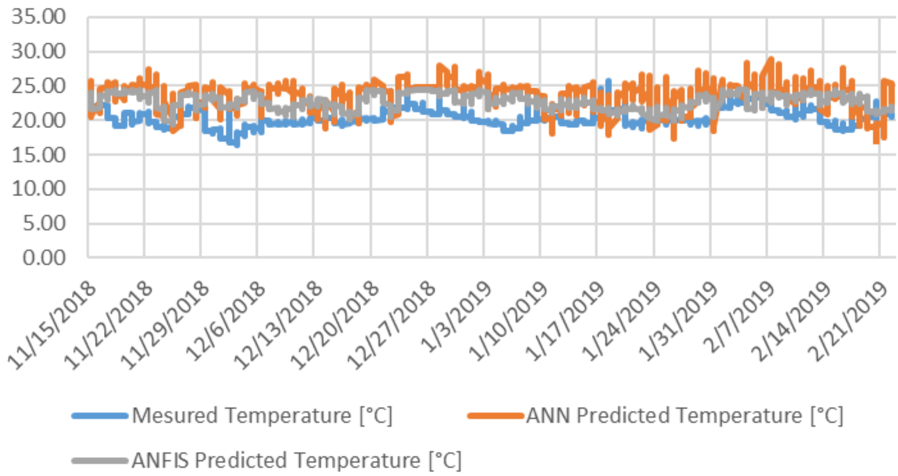

Fig. 10. Upper Room - Hourly predicted and measured temperature

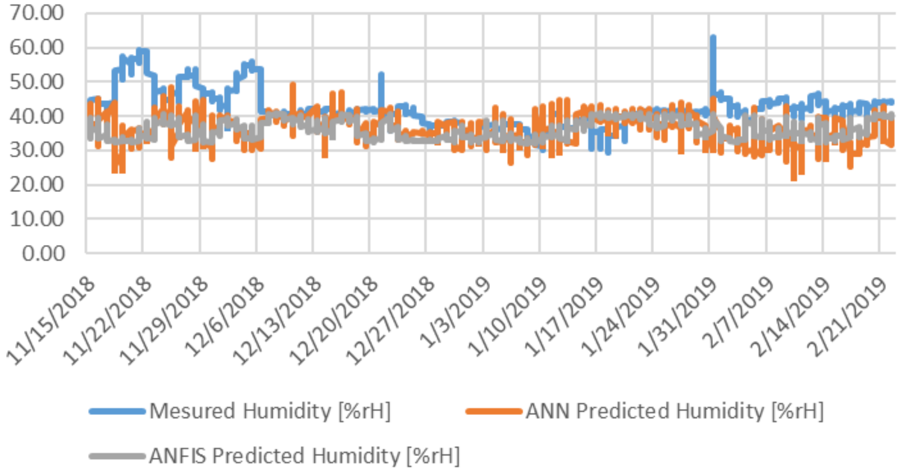

Fig. 11. Upper Room - Hourly predicted and measured humidity

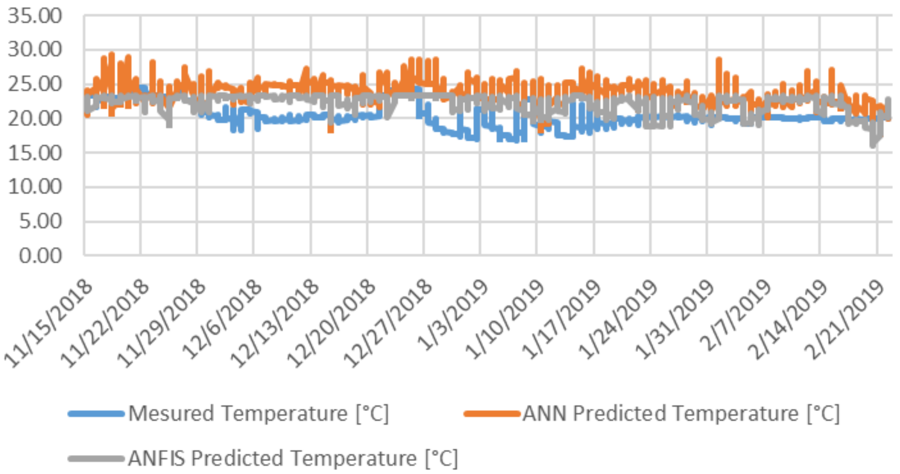

Fig. 12. Upper Lobby - Hourly predicted and measured temperature

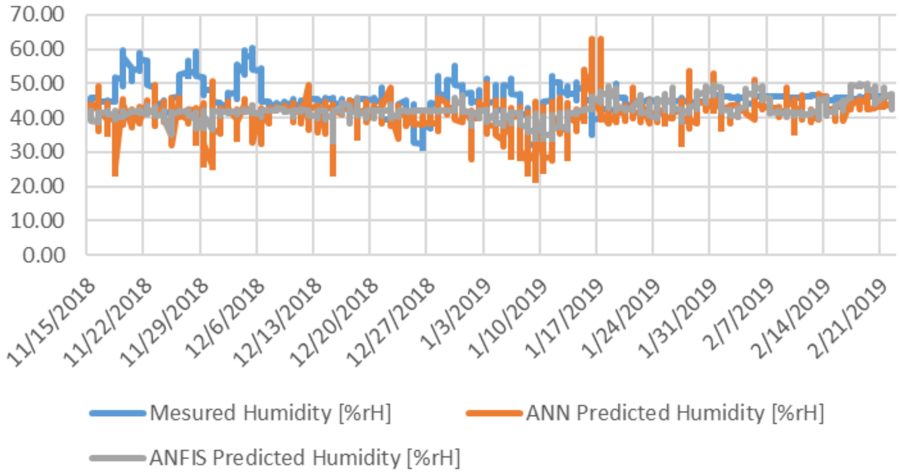

Fig. 13. Upper Lobby - Hourly predicted and measured humidity

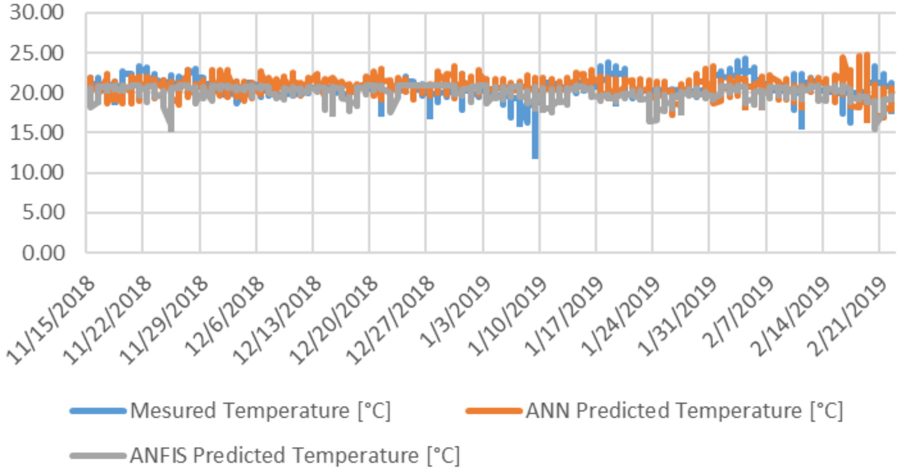

Fig. 14. Ground Lobby - Hourly predicted and measured temperature

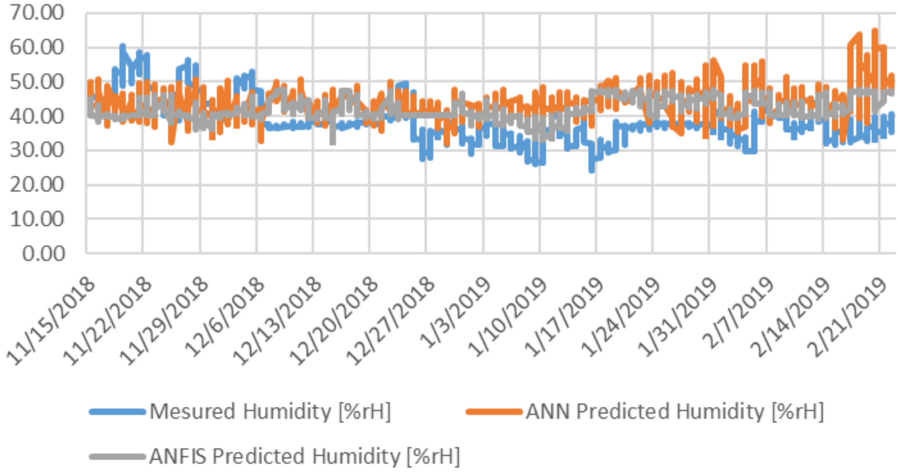

Fig. 15. Ground Lobby - Hourly predicted and measured humidity

Table 3. Predicted and measured temperature parameters

		Inside Temperature [°C]			
		Basement room	Upper room	Upper Lobby	Ground Lobby
Measured Average		18.5	20.4	20.6	20.7
ANN Predicted	Average	21.7	23.7	23.7	20.7
	RMSE	5.2	3.9	3.8	1.5
	Final RMSE	3.8			
ANFIS Predicted	Average	20.4	22.8	22.4	20.1
	RMSE	4.0	2.9	2.7	1.4
	Final RMSE	2.9			

However, these results show that the ANFIS prediction has better accuracy when it is compared with these stochastic measured data of the four points since the overall temperature prediction accuracy of the ANFIS model has been calculated to be 85%, on the other hand, the accuracy of the temperature ANN prediction was 81%.

According to Table 4, both the ANFIS and ANN predictions' errors are close to each other. The ANFIS humidity prediction results' RMSEs were 10.4, 8.3, 5.8, and 7.9 respectively in the basement room, upper room, upper lobby, and ground lobby, while in the same order the ANN prediction errors were 10.5, 8.5, 6.8, and 9. The ANFIS prediction was better in the four points. However, for overall humidity prediction again the ANFIS model which had an 81% accuracy rate was slightly more accurate than the ANN model which had an accuracy rate of 81%.

3.6.1. Scaling the comparison

The irregular performance of the heating system produced stochastic measured data in some periods. This stochasticity affected the comparison between measured and predicted data and the prediction accuracy of both the ANFIS and ANN models. To avoid this effect, the comparison is scaled by selecting the most ordinary period of the measured data as a sample to be compared with the predicted data and then providing the accuracy based on the selected sample.

By observing Figs. 16-23, it is recognized that the most stable period for the measured data was the period between 10/12/2018 and 20/12/2018. The data of this period will be the sample that is used for the comparison scaling.

Table 4. Predicted and measured humidity parameters

		Inside Humidity Rate [%rH]			
		Basement room	Upper room	Upper Lobby	Ground Lobby
Measured Average		45.4	42.0	45.6	38.6
ANN Predicted	Average	44.1	36.5	41.3	43.7
	RMSE	10.5	8.5	6.8	9.0
	Final RMSE	8.8			
ANFIS Predicted	Average	43.4	36.2	42.4	42.3
	RMSE	10.4	8.3	5.8	7.9
	Final RMSE	8.3			

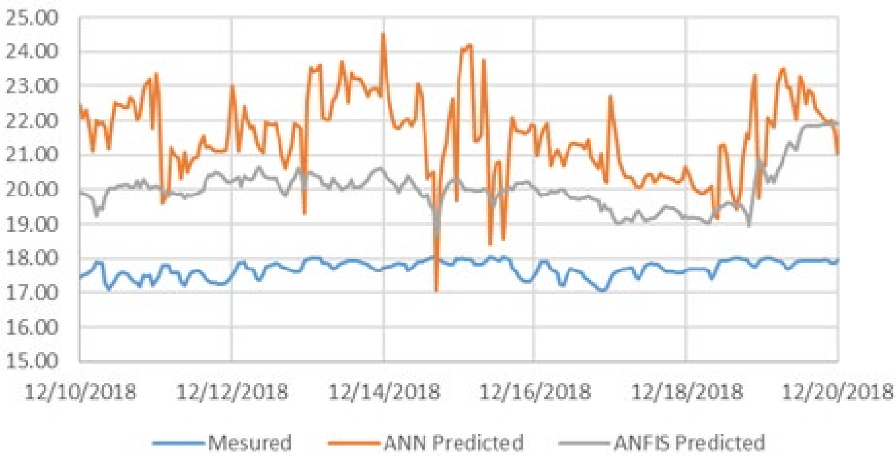

Fig. 2. Basement temperature data sample

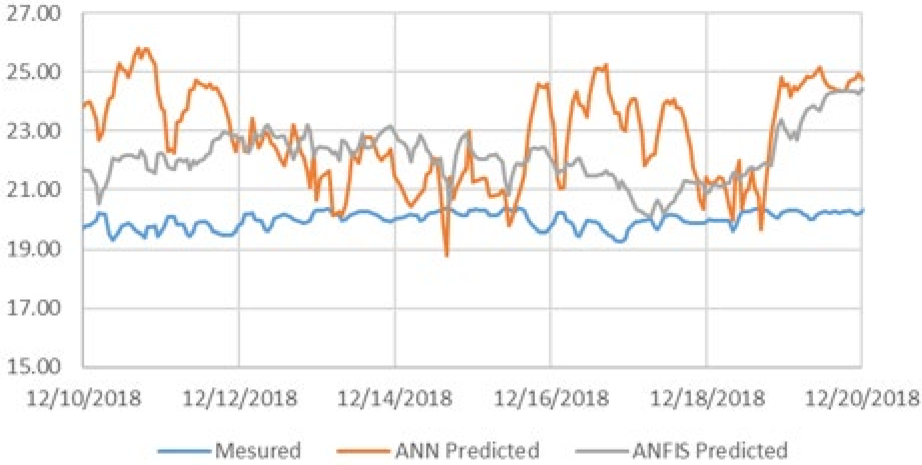

Fig. 17. Upper room temperature data sample

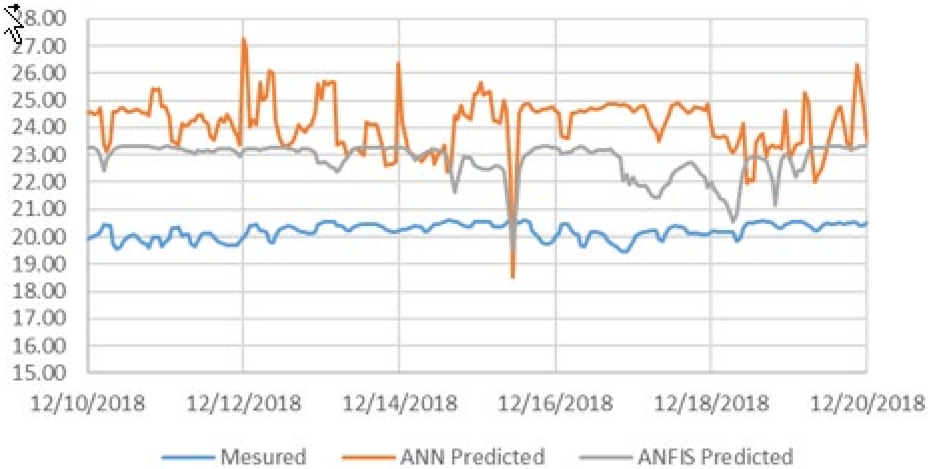

Fig. 18. Upper lobby temperature data sample

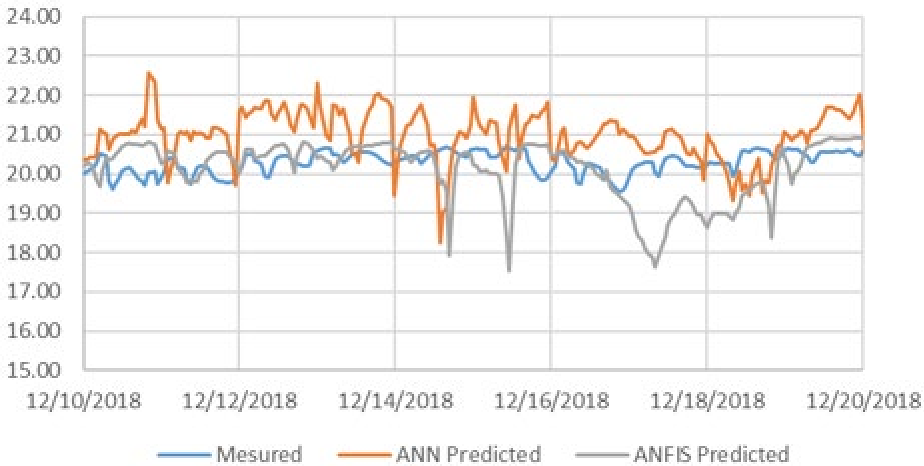

Fig. 19. Ground lobby temperature data sample

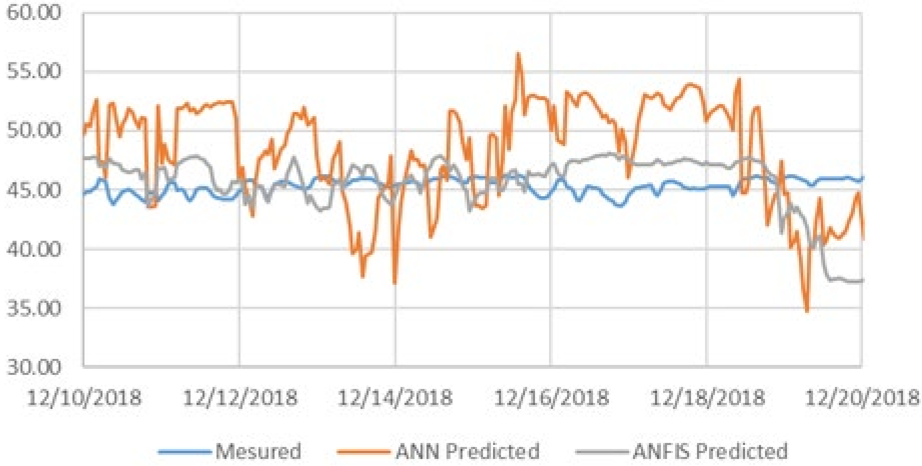

Fig. 20. Basement humidity data sample

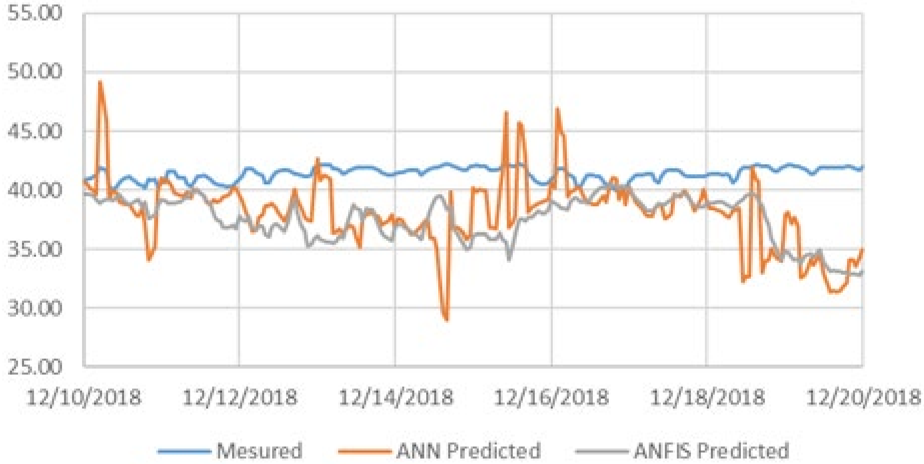

Fig. 21. Upper room humidity data sample

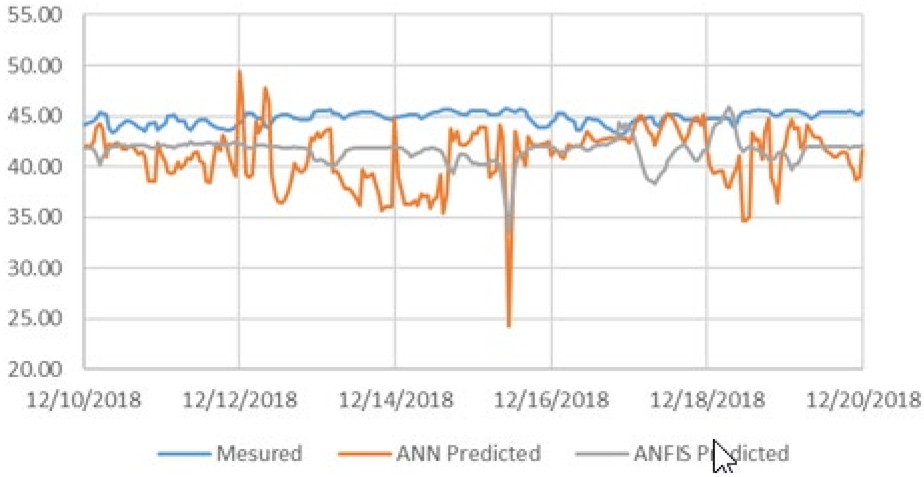

Fig. 22. Upper lobby humidity data sample

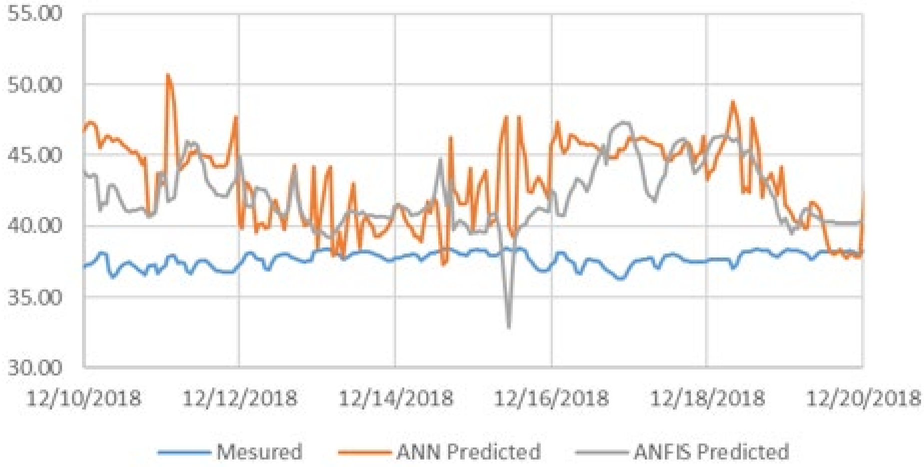

Fig. 23. Ground lobby humidity data sample

Figs. 16-19, and Table 5 show that the temperature prediction error increased in the most stable sample of the measured data for both ANN and ANFIS prediction models. For ANN the accuracy rate after scaling increased from 81% to 83%, but this increasing rate is still affected by the quality of the measured data, because even though the measured data is stable but it doesn't match the expected results, especially in the basement where the average of the measured data was less than 18°C, which is far away under the comfort zone. So the more the measured data is closed to the comfort zone, the less prediction error is achieved. The ANN maximum RMSE was in the basement at around 4.1, while it was minimum in the ground lobby point which is around 1. For the ANFIS model, the temperature prediction accuracy after scaling was increased by 3%, again the accuracy rate was affected by the heating system performance, and the average error decreased when the measured temperature was closer to the temperature comfort zone. The accuracy of ANFIS prediction became 88% after scaling while the ANN prediction accuracy increased to 85% which make the ANFIS model more eligible in term of accuracy to perform the kind of prediction work.

Figs. 20-23, and Table 6 showed that the prediction accuracy of the humidity was improved after scaling for both ANN and ANFIS prediction models. For ANN results the final prediction accuracy was increased by 6% after scaling. For ANFIS results also the final prediction error was decreased and the accuracy rate increased. The final humidity prediction accuracy of ANN is 87% while it is 90% for ANFIS, which shows that the ANFIS model's accuracy is better than the accuracy of the ANN model in the whole cases in this study.

The results also showed that the point with the lowest average measured temperature had the highest temperature prediction errors for both ANN and ANFIS, while the same point had the best average measured humidity and least humidity prediction errors. In opposite, the point with the best average measured temperature which is the closest to the comfort level, had the least temperature prediction errors, the lowest average measured humidity, and the highest humidity prediction errors.

4. Methodology

The thermal environment is the main index of the building energy performance and efficiency since it is the most important factor to improve the comfort level of the building, the main mission of most applied systems in any building is to provide a comfortable indoor environment. Hence, most of the building energy consumption is for heating or cooling. Therefore, it is necessary to manage certain comfort conditions such as temperature and humidity to manage and sustain the indoor environment's comfort level.

Table 5. Data sample temperature parameters

		Inside Temperature [°C]			
		Basement room	Upper room	Upper Lobby	Ground Lobby
Measured Average		17.7	20.0	20.2	20.3
ANN Predicted	Average	21.7	23.0	24.0	21.0
	RMSE	4.1	3.4	3.9	1.0
	Final RMSE	3.4			
ANFIS Predicted	Average	20.2	22.3	22.9	20.2
	RMSE	2.6	2.5	2.7	0.8
	Final RMSE	2.3			

Table 6. Data sample humidity parameters

		Inside Humidity Rate [%rH]			
		Basement room	Upper room	Upper Lobby	Ground Lobby
Measured Average		45.4	41.4	44.9	37.8
ANN Predicted	Average	47.5	37.9	41.1	42.8
	RMSE	5.7	4.8	4.8	6.0
	Final RMSE	5.3			
ANFIS Predicted	Average	45.0	37.2	41.6	42.0
	RMSE	3.5	5.0	3.6	4.9
	Final RMSE	4.3			

This study aims to compare two prediction models to provide better accuracy in predicting the heating season indoor thermal comfort data. This prediction aims to provide a full heating season's thermal comfort dataset by using short-term measured data while the heating system is performing. The heating season of the building was evaluated by defining a critical monthly heating consumption, which was 14 kW/m^2 per month, and selecting the period when the building's monthly consumption exceeds this value to be the heating season. Based on it the heating season was evaluated to be between November 15th and March 21st.

The ANFIS and ANN approaches had been used as predictive models. The two approaches were trained based on the measured indoor temperature and relative humidity data. The measurements inside the building were taken for one year which started on the 22nd of February 2018. Therefore, short-term data were collected during the first heating season, and these data were used in the model's training phase. While the data collected during the second heating season was used in validating the prediction results. In addition, the independent variables were obtained from the weather data and heating consumption simulated data.

The measurement and prediction works were done at four different points inside the building. The prediction results showed that the temperature averages should be in the comfort level for two points out of four, while the measured data showed that the four points are under the comfort condition. This was because of the poor performance of the heating system in some periods. This poor performance caused stochasticity in the measured data which affected the prediction results accuracy.

The ANN prediction errors for temperature varied between 1.5 and 5.2, and between 6.8 and 10.5 for humidity in the four points. The ANFIS prediction errors have recognized variations too since the temperature prediction errors were between 14 and 4, and for the humidity, prediction errors were between 5.8 and 10.4. These results showed that the ANN and ANFIS models achieved the best prediction with a minimum error rate at the point where the measured data was more stable with fewer variations.

However, the ANFIS prediction was more accurate in general since its prediction final accuracy rate was 85% for temperature and 81% for humidity, while the ANN prediction's final accuracy rates were 81% for temperature and 80% for humidity.

These results were significantly affected by the heating system's poor performance, in order to minimize this effect, the comparison was scaled by selecting the best measured period to be the data sample that will be used in the comparison. After scaling, the prediction accuracy was increased for both ANN and ANFIS models, to be 83% and 88%, respectively for temperature prediction. For humidity the accuracy rate of 87% for ANN and 90% for ANFIS. According to the results, the ANFIS model was the best fit for all of these prediction work cases. Considering the measured data stochasticity, both the ANFIS and ANN approaches are highly validated in this type of prediction work. Since the building is an elderly home, these results can be an indicator to improve the thermal environment inside the building Taking into account its impact on the health and well-being of older persons.

The results of this study offer the opportunity to extend in different directions as further work. The results could support the monitoring system which could be implemented inside the building to perform real-time calibration besides reporting the unexpected results which could assist to improve the indoor comfort level. The prediction results could also be used as an index to calibrate and develop the accuracy of the energy performance simulation of the building by improving the set points.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

The specifications of the buildings were applied from the European Union project that was called "Renovation of Residential urban spaces: Towards nearly zero energy CITIES" (R2CITIES) Grant Agreement No 314473 [22].

References

[1] Afroz Z, Shafiullah G, Urmee T, Higgins G (2017) Prediction of indoor temperature in an institutional building. Energy Procedia 142: 1860–1866. doi: 10.1016/j.egypro.2017.12.576.

[2] Lu T, Viljanen M (2009) Prediction of indoor temperature and relative humidity using neural network models: Model comparison. Neural Computing and Applications 18(4): 345–357. doi: 10.1007/s00521-008-0185-3.

[3] Piri J, Mohammadi K, Shamshirband S, Akib S (2016) Assessing the suitability of hybridizing the Cuckoo optimization algorithm with ANN and ANFIS techniques to predict daily evaporation. Environmental Earth Sciences 75(3): 1–13. doi: 10.1007/s12665-015-5058-3.

[4] Golafshani EM, Behnood A, Arashpour M (2020) Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer. Construction and Building Materials 232: 117266. doi: 10.1016/j.conbuildmat.2019.117266.

[5] Haji A, Payvandy P (2020) Application of ANN and ANFIS in prediction of color strength of plasma-treated wool yarns dyed with a natural colorant. Pigment & Resin Technology 49(3): 171–180. doi: 10.1108/PRT-10-2019-0089.

[6] Niknam Shahrak M, Esfandyari M, Karimi M (2019) Efficient prediction of water vapor adsorption capacity in porous metal–organic framework materials: ANN and ANFIS modeling. Journal of the Iranian Chemical Society 16(1): 11–20. doi: 10.1007/s13738-018-1476-y.

[7] Sahoo PK, Pattnaik S, Sutar MK. Parametric Optimization of Permeability of Green Sand Mould Using ANN and ANFIS Methods. Springer Singapore, 2020.

[8] Khazaee Poul A, Shourian M, Ebrahimi H (2019) A comparative study of MLR, KNN, ANN and ANFIS models with wavelet transform in monthly stream flow prediction. Water Resources Management 33(8): 2907–2923. doi: 10.1007/s11269-019-02273-0.

[9] Taşan S, Demir Y (2020) Comparative analysis of MLR, ANN, and ANFIS models for prediction of field capacity and permanent wilting point for bafra plain soils. Communications in Soil Science and Plant Analysis 51(5): 604–621. doi: 10.1080/00103624.2020.1729374.

[10] M Moosavi, K Firoozi Rad, A Daneshvar (2019) A comparative study in the prediction of thermal conductivity enhancement of nanofluids using ANN-MLP, ANN-RBF, ANFIS, and GMDH methods. Journal of the Iranian Chemical Society 16(12): 2629–2637. doi: 10.1007/s13738-019-01725-z.

[11] Sözer H, Aldin SS (2019) Predicting the indoor thermal data for heating season based on short-term measurements to calibrate the simulation set-points. Energy and Buildings 202: 109422. doi: 10.1016/j.enbuild.2019.109422.

[12] Masoudi S, Sima M, Tolouei-Rad M (2018) Comparative study of ann and anfis models for predicting temperature in machining. Journal of Engineering Science and Technology 13(1): 211–225.

[13] Riahi-Madvar H, Seifi A (2018) Uncertainty analysis in bed load transport prediction of gravel bed rivers by ANN and ANFIS. Arabian Journal of Geosciences 11: 688. doi: 10.1007/s12517-018-3968-6.

[14] Soni U, Singh N, Swami Y, Deshwal P. A comparison study between ANN and ANFIS for the prediction of employee turnover in an organization. International Conference on Computing, Power and Communication Technologies (GUCON), 28-29 September 2018, Greater Noida, India.

[15] Yang X, Song X, Huang L (2015) Some general inequalities for Choquet integral. Applied Mathematics 6(14): 2292–2299. doi: 10.4236/am.2015.614201.

[16] Ralescu D, Adams G (1980) The fuzzy integral. Journal of Mathematical Analysis and Applications 75(2): 562–570. doi: 10.1016/0022-247X(80)90101-8.

[17] Aqil M, Kita I, Yano A, Nishiyama S (2007) A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. Journal of Hydrology 337(1–2): 22–34. doi: 10.1016/j.jhydrol.2007.01.013.

[18] Roger Jang JS, Sun CT, Mizutani E. Neuro-Fuzzy and Soft Computing. Prentice Hall, USA, 1997.

[19] Talpur N, Salleh MNM, Hussain K (2017) An investigation of membership functions on performance of ANFIS for solving classification problems. IOP Conference Series: Materials Science and Engineering 226: 012103. doi: 10.1088/1757-899X/226/1/012103.

[20] Aydin AC, Tortum A, Yavuz M (2006) Prediction of concrete elastic modulus using adaptive neuro-fuzzy inference system. Civil Engineering and Environmental Systems 23(4): 295–309. doi: 10.1080/10286600600772348.

[21] DesignBuilder Software Ltd., "DesignBuilder." GL5 2AD, UK, 2018, [Online]. Available: https://www.designbuilder.co.uk/.

[22] "R2CITIES- Residential Renovation towards nearly zero energy CITIES." http://r2cities.eu/ (accessed Mar. 28, 2019).