Original Paper

Indoor and Built Environment

Indoor Built Environ 2012;21;6:772–781

Accepted: March 24, 2011

Models for Prediction of Daily Mean Indoor Temperature and Relative Humidity: Education Building in Izmir, Turkey

Türkan Göksal Özbalta[a] Alper Sezer[a] Yusuf Yıldız[b]

[a]Department of Civil Engineering, Ege University, Izmir, Turkey
[b]Department of Architecture, Izmir Institute of Technology, Izmir, Turkey

Key Words
Artificial neural network · Multiple regression · Modelling · Indoor temperature and relative humidity

Abstract
In this research, several models were developed to forecast the daily mean indoor temperature (IT) and relative humidity values in an education building in Izmir, Turkey. The city is located at a hot–humid climatic region. In order to forecast the IT and internal relative humidity (IRH) parameters in the building, a number of artificial neural networks (ANN) models were trained and tested with a dataset including outdoor climatic conditions, day of year and indoor thermal comfort parameters. The indoor thermal comfort parameters, namely, IT and IRH values between 6 June and 21 September 2009 were collected *via* HOBO data logger. Fraction of variance (R^2) and root-mean squared error values calculated by the use of the outputs of different ANN architectures were compared. Moreover, several multiple regression models were developed to question their performance in comparison with those of ANNs. The results showed that an ANN model trained with inconsiderable amount of data was successful in the prediction of IT and IRH parameters in education buildings. It should be emphasized that this model can be benefited in the prediction of indoor thermal comfort conditions, energy requirements, and heating, ventilating and air conditioning system size.

Introduction

After energy crises of 1973, global energy demand has increased rapidly. Therefore, countries started to constitute national strategies to use energy efficiently. These kinds of strategies are very important for Turkey, since there are significant differences between the energy demands and consumption rate. Thus, the deficit in energy requirements has been compensated by importing energy from neighbouring countries [1]. In 2005, Turkey's energy consumption was 92.4 MTOE (million tonnes of oil equivalent), while production was 26.81 MTOE from primary domestic sources. By the year 2020, energy

DOI: 10.1177/1420326X11422163
Accessible online at http://ibe.sagepub.com
Figure 5 appears in colour online

Yusuf Yıldız,
Department of Architecture, Izmir Institute of Technology, Izmir, Turkey. Tel. +90 232 750 7156, Fax +90 232 750 7012,
E-Mail yusifyildiz@gmail.com

consumption is considered to increase to a value of 222 MTOE, while the primary energy production is expected to rise to 70 MTOE [2].

Generally speaking, energy consumption and requirements in Turkey arise from four main sectors, which can be enlisted as industry, buildings, transportation and agriculture. According to the report prepared by The Scientific and Technical Research Council of Turkey (TÜBİTAK) [3], the building sector is the second biggest energy consumer sector succeeding industry and it is estimated that the energy consumption of the building sector will continue to rise up in the future [4].

Indoor thermal comfort conditions affect the amount of the energy consumption in buildings [5,6] and they are usually dependent on indoor temperature and relative humidity values. Their importance stems from the need to define correct size of the active heating and cooling systems; determination of these values are also extremely significant in calculation of heating and cooling energy consumptions in buildings [7]. Otherwise, high amount of energy can be consumed due to oversized heating, ventilating and air conditioning (HVAC) systems in buildings. Therefore, the size of the HVAC systems can be defined by the use of reasonable predictions of indoor temperature (IT) and IRH parameters.

Prediction of these values to a high level of accuracy is not a simple task. Use of building energy simulation tools is an effective way in the estimation of indoor thermal comfort parameters. However, they necessitate several detailed input data such as building thermal characteristics, true level of occupancy, lighting gains and equipment loadings [8]. This is also a time-consuming process. Several techniques existing in literature reveal that nonlinear models are more powerful than linear ones in terms of certain predictions [9]. Besides, usage of ANN is common in various building applications. A simple literature survey leads to a vast number of investigations including various models aiming to forecast indoor temperature, relative humidity and size of HVAC equipment in buildings. Thomas and Soleimani-Mohseni [10] compared different models such as nonlinear artificial neural networks (ANNs), and linear ARX-ARMAX (ARX, Auto-Regressive eXogenous input; ARMAX, Autoregressive Moving Average with Exogenous inputs) and Box–Jenkin models. In these models, the outdoor and indoor temperature, heating power, wall temperatures, ventilation flow rate, time of day and sun's radiation were used as input signals for two buildings. It was shown that ANN models give more accurate temperature prediction than linear ones for both buildings. Lu and Viljanen [11] used nonlinear autoregressive model with external input and generic algorithm to estimate indoor temperature and relative humidity values in a test house. As a result, correlation coefficients 0.998 and 0.997 were obtained in the testing stage. Parishwad et al. [12] estimated the outdoor ambient temperature, relative humidity and air velocity in India by the use of monthly–mean hourly values of these parameters with developed correlations. The method can be used to predict the weather parameters at different locations of India. Mohseni et al. [13] carried out a study on prediction of operative temperature in rooms and buildings *via* several ANN models. In addition, nonlinear ANN models could provide better estimations than linear ARX models but the most accurate prediction was obtained using feed-forward ANN models with one hidden layer and Levenberg–Marquardt training algorithm. Ruano et al. [14] evaluated the usage of neural network models for estimation of indoor air temperature. The results presented illustrated that multi-objective genetic algorithms for the off-line design of radial basis function neural networks and the neural models, can give better results than the state-of-the-art physical models. Yigit and Ertunc [15] predicted the air temperature and humidity at the outlet of a wire-on-tube type heat exchanger by applying ANN. Mean relative errors were found to be less than 1% and 2% for air temperature and humidity values, respectively. These values showed that the method could be useful for the manufacturer to estimate the performance of cooling coils in HVAC systems. Reginato et al. [16] focused on the definition of a multiple-input/single output orthonormal basis functions model to estimate indoor air temperature profile and energy consumption. It was found that there is a good agreement between temperature values produced by the model and simulation tool. Mustafaraj et al. [17] evaluated the potential of a neural Network-based Nonlinear Auto-Regressive model with eXternal inputs (NNARX), a Nonlinear Autoregressive Moving Average Model with eXternal inputs (NNARMAX) and a Nonlinear Output Error (NNOE) model to estimate the thermal behaviour (temperature and relative humidity) of an open-plan office. According to the results, all methods can provide acceptable predictions but the NNARX and NNARMAX methods could give more accurate outcomes, in comparison with NNOE.

Previous studies intended to predict the different parameters such as indoor or outdoor temperature, operative temperature, energy performance and size of HVAC equipment using various methods. In this research, an estimation of indoor relative humidity and air temperature in an existing education building in Izmir was evaluated by the use of ANN and regression models utilizing outdoor climatic parameters as inputs. It was initially intended to establish the input–output relationship by the use of linear and nonlinear regressions. Afterwards, ANN models were utilized to obtain better relationships in the prediction of ?internal relative humidity (IRH) and IT parameters. Small modifications in network architecture have led to a considerable number of models. The generated ANN models were trained and tested with daily mean indoor temperature and relative humidity values measured between 6 June and 21 September 2009 with HOBO data logger.

Description of Education Building

In this research, an existing four-storey university building was selected at the Ege University Campus in Izmir, Turkey. The reason in this selection is simple: indoor thermal comfort conditions and energy performance are amply significant for the lecture/seminar rooms, which are the basic requirements for educational buildings. Students spend most of their time in seminar rooms to concentrate on their studies [18]. It is also known that thermal characteristics of the building envelope have a deterministic role on indoor building conditions. External walls of the studied building are composed of external plaster, reinforced concrete and internal plaster with an U-value (heat transfer coefficient) of $2.856\,W{\cdot}m^{-2}K^{-1}$. A small part of the external walls consists of Bims block, and external and internal plasters ($U = 1.054\,W{\cdot}m^{-2}K^{-1}$). The ceiling of the building ($U = 0.45\,W{\cdot}m^{-2}K^{-1}$) is constructed using reinforced concrete with 4 cm of extruded polystyrene. The roof is the only structural part and it was built to meet the defined requirements of TS 825, Thermal Insulation Standard in Buildings in Turkey [19]. The ground floor has a U-value of $2.9\,W{\cdot}m^{-2}K^{-1}$; 29% of the all external walls are covered with double clear glazing windows ($U = 2.8\,W{\cdot}m^{-2}K^{-1}$). The typical floor plan of the university building is shown in Figure 1.

The investigated indoor thermal variables were indoor air temperature and relative humidity. These parameters were monitored using HOBO data logger. The instrument was placed at a height of 1.4 m from the ground surface on the indoor wall in an office room, which ensured few disturbances due to routine activities. Office room was monitored during a critical summer period from 6 June and 21 September 2009. A fan coil unit in the office was used to control the temperature during the winter and summer months. However, this was not effective for cooling due to its capacity.

Computational Methods

Artificial Neural Networks

ANNs were developed to simulate the substantial features of human nervous system to solve building problems in a supervised manner. ANN mimics the human brain

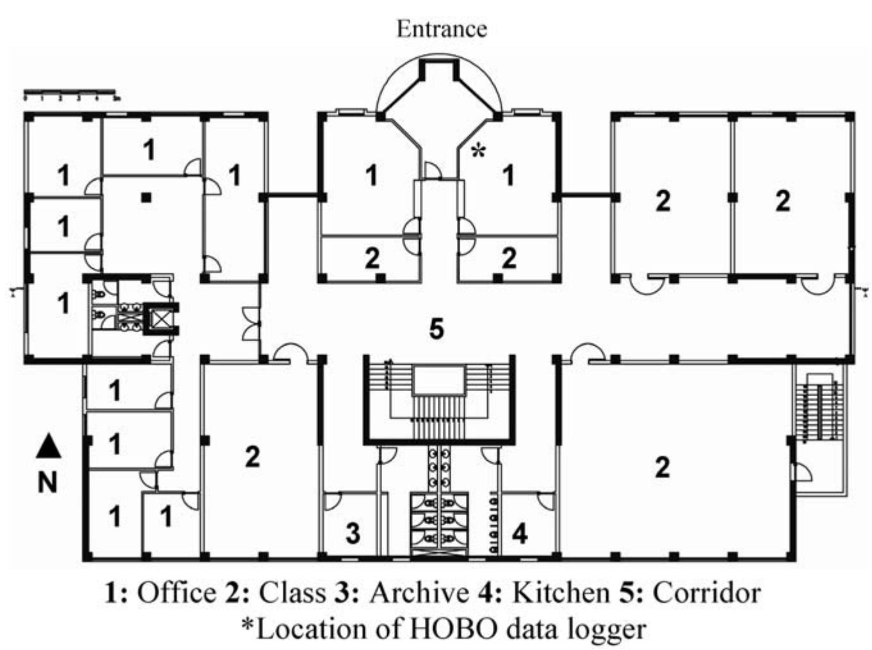

Fig. 1. Typical floor plan of university building.

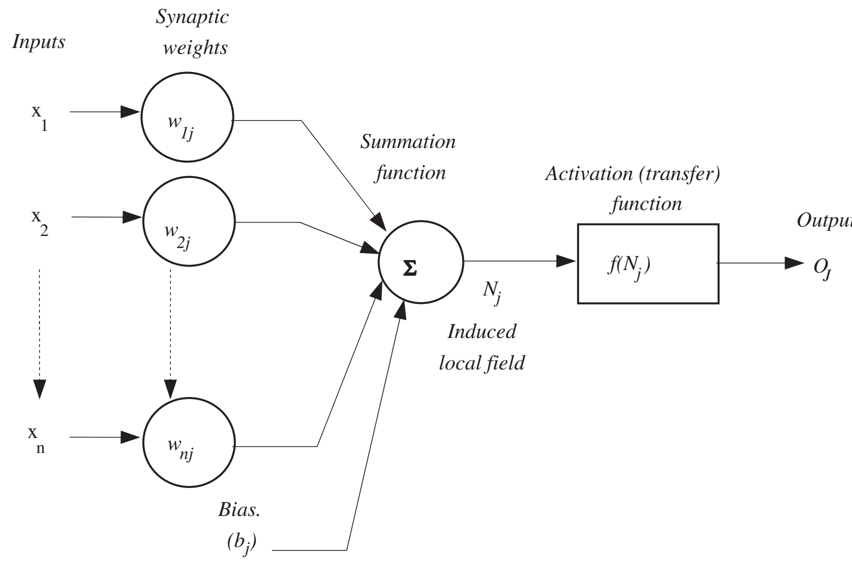

Fig. 2. Definition of processing element in ANN.

by the use of simple computational elements called artificial neurons that are connected by various weights. Multilayer perception structure is generally preferred for its precise computational ability and remarkable capability for classification, prediction and recognition of problems. An artificial neuron consists of five elements: inputs, weights, summation function, activation function and outputs (Figure 2).

A typical neural network model has three independent layers: input, hidden and output layers. Each layer is fitted with several operating neurons [20]. Input values are instructed to input neurons, while the outputs are assigned by the output neurons. The hidden layer performs an interface to fully interconnect input and output layers. A neuron is connected to all neighbouring neurons by a connection weight, which indicates the strength of the relationship between two connected neurons. The summation function is used to calculate the sum of inputs of the neurons in progress and the sum is converted to an output value based on a predefined activation function. The weighted sums of inputs are calculated by the use of Equation (1):

$$N_j = \sum_{i=1}^{n} \omega_{ij} x_i + b \tag{1}$$

where N_j is the weighted sum of the jth neuron, ω_{ij} the synaptic weight and b the bias value [20]. The sigmoid function is a kind of the activation functions, which are used to process the net input obtained from the summation operation, in order to calculate the outputs, as given by Equation (2) [21]:

$$O_j = f(N_j) = \left[1 + e^{-\alpha N_j}\right]^{-1} \tag{2}$$

where α is a constant. Although the range of the log-sigmoid function outputs is between −1 and 1, this interval can be modified. Moreover, tangent-sigmoid function is a powerful tool, especially when speeding up the computational process is essential:

$$O_j = f(N_j) = 2 \times \left[1 + e^{-2N_j}\right]^{-1} \tag{3}$$

At the end of the calculations, these parameters are inverse-normalized, as shown in Equation (3), to calculate the outputs. In order to make accurate predictions with ANN approach, the network needs a supervised learning algorithm, which enables the network to learn the nonlinear input–output relationship. The back-propagation algorithm with a gradient search technique (the steepest gradient descent method) would minimize a function equal to the mean square difference between the desired and the actual network outputs, which is widely used for the supervision of neural networks. Since ANN methodology is widely described by hundreds of studies in the past, the reader is advised to check out the literature for further information [20,21].

Table 1. Basic statistical analysis on both testing and training data

Property	DOY	OT	ORH (%)	WS	IT	IRH
TRAINING DATA						
Elements in dataset	94	94	94	94	94	94
Average	252.46	26.22	49.03	3.12	26.65	38.38
Median	250.00	26.60	47.20	3.10	27.33	37.32
Minimum	192.00	17.00	28.00	1.70	21.76	24.71
Maximum	322.00	31.90	86.50	5.00	30.48	55.47
Standard deviation	37.70	3.29	11.43	0.75	2.28	6.43
Variance	1421.50	10.82	130.75	0.56	5.21	41.36
Standard error	3.89	0.34	1.18	0.08	0.24	0.66
Skewness	0.21	−0.60	0.70	0.24	−0.55	0.41
Kurtosis	−1.04	−0.21	0.68	−0.38	−0.83	−0.01
TESTING DATA						
Elements in dataset	25	25	25	25	25	25
Average	253.68	25.98	49.01	3.11	26.58	37.87
Median	254.00	26.60	48.90	3.00	26.97	37.78
Minimum	195.00	18.30	23.50	1.50	21.97	24.37
Maximum	319.00	30.10	73.40	4.60	30.00	50.20
Range	124.00	11.80	49.90	3.10	8.03	25.83
Standard deviation	36.07	3.21	12.67	0.85	2.21	6.42
Variance	1300.73	10.28	160.60	0.73	4.89	41.20
Standard error	7.21	0.64	2.53	0.17	0.44	1.28
Skewness	0.17	−0.71	0.08	0.23	−0.63	0.06
Kurtosis	−0.81	−0.17	−0.44	−0.88	−0.57	−0.25

Regression Analyses for the Prediction of Thermal Comfort Parameters

In simplest words, regression analysis can be defined as the establishment of relationships between a dependent variable (output) and independent variables (inputs). Either linear or nonlinear multiple regression analyses can be employed to construct plausible relationships describing the variation of a dataset including predetermined inputs and output. Multiple regression analysis is a simple and powerful tool for prediction and forecasting. Meaningful relationships between inputs and outputs can be established as far as it is assumed that the sample under consideration is representative of the population, and the variance of errors is homogenously distributed within the dataset. Equation (4) is the generalized form of regression equations:

$$y = f(x,\alpha) \tag{4}$$

In Equation (4), y is the dependent variable, x the independent inputs and α the unknowns. Both linear and nonlinear models aim to minimize the sum of squared residuals, whereby smaller residuals would indicate better goodness-of-fit. Linear regression methods aim to fit a linear combination of functions to a predetermined dataset. Similarly, nonlinear regression methods have the same goal; they aim to fit a nonlinear model on given data which is ensured by minimization of an objective function. Linear regression equations are established by the use of least squares estimation, instrumental variables regression, maximum likelihood estimation, adaptive estimation, etc. On the other hand, nonlinear regression investigations necessitate numerical optimization algorithms to find out the unknowns. The nonlinear model uses an iterative approach to minimize the objective function, which is dependent on the initial estimates. It should be stated that the limited number of linear models can be established without transforming the data, which is a disadvantage of linear regression. As the iteration proceeds, objective function is minimized by the use of transformed data, which increases the performance of nonlinear regression in comparison with that of linear regression. The local minima problem should be overcome to reach the global minimum, which denotes the best nonlinear regression equation [22,23].

Results and Discussion

Model Parameters and Statistical Analysis

The data obtained from HOBO data logger have been subjected to linear and nonlinear regression analyses.

Table 2. Assessment of model performances in terms of output parameters (EN: epoch number, tan-sig: tangent-sigmoid and log-sig: logarithmic-sigmoid)

Model	Architecture/formula	EN	Transfer function type (in layer)		Training			Testing		
			Hidden	Output	R^2	VAF (%)	RMSE	R^2	VAF (%)	RMSE
IT (°C)										
MR1	Equation (5)	–	–	–	0.87	87.47	0.80	0.81	80.58	0.99
MR2	Equation (6)	–	–	–	0.87	87.46	0.80	0.80	79.69	1.00
MR3	Equation (7)	–	–	–	0.79	77.21	1.08	0.77	73.19	1.16
ANN1	$4 \times 10 \times 2$	500	Tan-sig	Tan-sig	0.97	96.98	0.39	0.93	93.06	0.57
ANN2	$4 \times 10 \times 2$	500	Log-sig	Log-sig	0.91	90.50	0.93	0.80	80.11	1.02
ANN3	$4 \times 10 \times 2$	500	Tan-sig	Linear	0.98	97.98	0.32	0.94	93.28	0.56
ANN4	$4 \times 10 \times 2$	500	Linear	Linear	0.87	87.47	0.80	0.82	81.92	0.94
ANN5	$4 \times 10 \times 2$	100	Tan-Sig	Linear	0.98	97.52	0.36	0.91	91.15	0.65
ANN6	$4 \times 10 \times 2$	100	Linear	Linear	0.87	87.47	0.80	0.82	81.92	0.94
IRH (%)										
MR4	Equation (8)	–	–	–	0.87	87.03	2.30	0.85	83.66	2.57
MR5	Equation (9)	–	–	–	0.87	86.98	2.31	0.84	80.42	2.79
MR6	Equation (10)	–	–	–	0.85	85.35	2.45	0.82	81.04	2.79
ANN1	$4 \times 10 \times 2$	500	Tan-sig	Tan-sig	0.97	96.23	1.15	0.91	90.29	0.99
ANN2	$4 \times 10 \times 2$	500	Log-sig	Log-sig	0.88	81.77	2.39	0.89	82.72	2.66
ANN3	$4 \times 10 \times 2$	500	Tan-sig	Linear	0.97	96.28	1.14	0.96	92.89	0.90
ANN4	$4 \times 10 \times 2$	500	Linear	Linear	0.87	84.85	2.30	0.86	85.82	2.17
ANN5	$4 \times 10 \times 2$	100	Tan-sig	Linear	0.97	96.21	1.15	0.91	91.98	1.03
ANN6	$4 \times 10 \times 2$	100	Linear	Linear	0.87	84.85	2.30	0.86	85.82	2.47

The dataset consists of 119 recordings, and they were separated into two parts: 94 for training sessions and 25 for testing sessions. The software used in the establishment of regression and ANN model calculations was MATLAB 7.0. The outcomes of several ANN models were compared with those of regression analysis outcomes. Evaluations were given in the following sections. The inputs for the models were determined as day of year (DOY), outdoor temperature (OT), outdoor relative humidity (ORH) and wind speed (WS). Indoor conditions are mostly dependent on the outdoor weather conditions as well as the features of the building envelope, level of occupancy, air tightness and internal heat gains [24]. For example, the indoor temperature in a space which does not have any heating and cooling system can usually vary based on the outside temperature and the sun which affects the space. The relative humidity can change based on the outside humidity and heating or cooling supplied by heating and cooling sources. In addition, outdoor weather conditions are measurable parameters and they do not change from one study to another. In other words, selected inputs are independent from the selected building. The input parameters were provided from Turkish State Meteorological Service in Izmir. Additional climatic parameters were measured, but parameters such as daily solar radiation could not be measured in the time period between 6 June and 21 September 2009. Therefore, the number of input parameters was inevitably constrained. On the other hand, the output parameters were IT and IRH, as expected. Relative humidity was especially preferred instead of moisture content because the use of relative humidity values is very common in the assessment of thermal comfort conditions. It should be reminded that input data were separated into two parts: training and testing datasets. Basic statistical analyses employed on two datasets and their results were tabulated in Table 1.

Investigations of the analysis results have shown that the average values of the same parameters in training and testing datasets are similar. The standard deviation values in the two datasets are 3.29/3.21 and 6.43/6.42 for OT and IRH parameters, respectively. The given results of maximum, minimum, standard error and variance values indicated that the testing and training databases are similar to each other. Skewness parameters in two databases had the same trends. However, it should be stated that the skewness values of IRH parameter in training and testing databases were 0.41 and 0.06. This means that the data belonging to IRH parameter in testing database were cumulating in the vicinity of the mean value in comparison with the training database, which could have been the cause of probable errors in modelling phase.

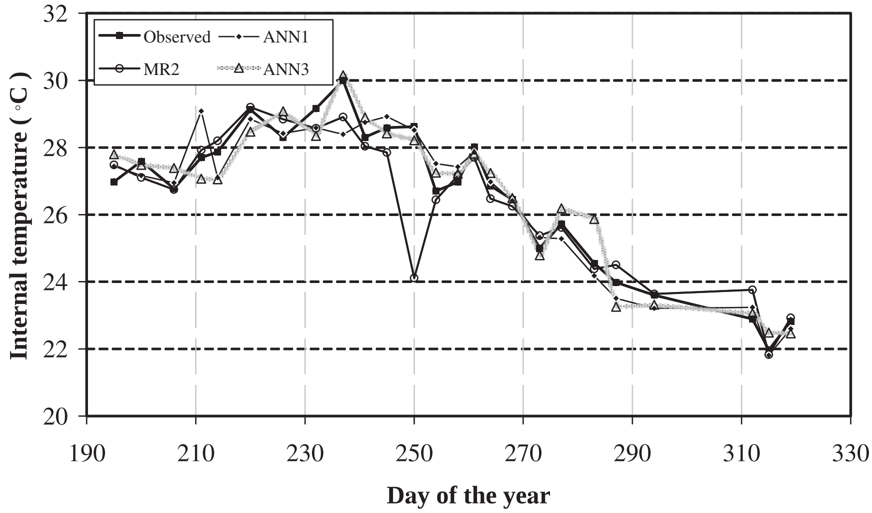

Fig. 3. Predicted and measured values of IT parameters in selected models.

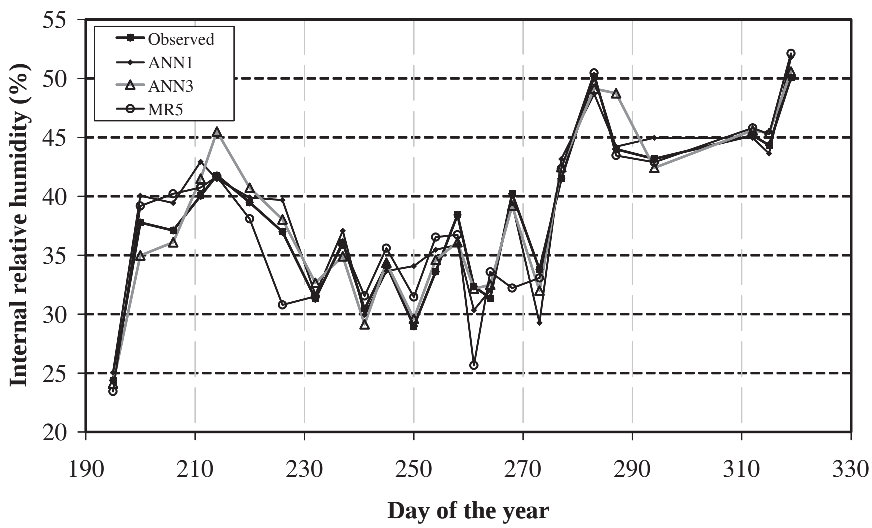

Fig. 4. Predicted and measured values of IRH parameters in selected models.

Evaluation of Multiple Regression and ANN Models

Several multiple regression equations were established for predicting IT and IRH parameters. The regression equations were constituted by the use of the training dataset, and their modelling ability is tested *via* training database. The regression equations (Equations (5)–(10)) for IT parameter prediction are as follows:

$$\begin{aligned}\mathrm{MR1}: \mathrm{IT} = &-0.022 \times \mathrm{DOY} + 0.492 \times \mathrm{OT} + 0.021\\ &\times \mathrm{ORH} - 0.023 \times \mathrm{WS} + 18.430\end{aligned} \tag{5}$$

$$\mathrm{MR2}: \mathrm{IT} = e^{-0.0008\times\mathrm{DOY}+0.0195\times\mathrm{OT}+0.0009\times\mathrm{ORH}+0.0003\times\mathrm{WS}+2.933} \tag{6}$$

$$\begin{aligned}\mathrm{MR3}: \mathrm{IT} = &-0.021 \times \mathrm{DOY} + 0.843 \times \mathrm{OT} + 0.078\\ &\times \mathrm{ORH} + 0.385 \times \mathrm{WS}\end{aligned} \tag{7}$$

Moreover, established models for IRH prediction are given in the following equations:

$$\begin{aligned}\mathrm{MR4}: \mathrm{IRH} = &\ 0.045 \times \mathrm{DOY} + 0.294 \times \mathrm{OT} + 0.487\\ &\times \mathrm{ORH} - 0.367 \times \mathrm{WS} - 3.472\end{aligned} \tag{8}$$

$$\begin{aligned}\mathrm{MR5}: \mathrm{IRH} = &\ 0.041 \times \mathrm{DOY} + 0.228 \times \mathrm{OT} + 0.476\\ &\times \mathrm{ORH} - 0.444 \times \mathrm{WS}\end{aligned} \tag{9}$$

$$\mathrm{MR6}: \mathrm{IRH} = e^{0.0013\times\mathrm{DOY}+0.0075\times\mathrm{OT}+0.0114\times\mathrm{ORH}-0.0087\times\mathrm{WS}+2.584} \tag{10}$$

The evaluation of the models was given in the subsequent lines. Three models were capable of predicting the IT and IRH parameters to a reasonable degree. The models were even more successful than a number of ANN models.

A number of ANN models have been designed for a comparative approach to thermal comfort parameter predictions. As explained before, four parameters were selected as inputs while two were selected as outputs. The number of hidden layers was altered by assigning values 10, 30, 50 and 70. Investigating the results, it can be stated that the performance of the models demonstrated insignificant changes. Therefore, the numbers of epochs and transfer function in hidden and output layers were changed to apply a variety of ANN architectures. As seen in Table 2, the numbers of iterations in the ANN calculation were selected as 100 and 500. The number of iterations was decided by a trial-and-error approach. During the trial-and-error analyses, overestimating and underestimating problems were considered for small iteration numbers, such as 50. As overtraining and inefficiency were observed at larger number of iterations, the results of trials on iteration numbers surpassing 500 were not included in the text. Nonetheless, the results were consistent with the findings, and selected iteration numbers 100 and 500 were found to be consistent in terms of optimization and calculation complexity. Furthermore, tan-sig, log-sig and linear transfer functions and their combinations were used in various network architectures. Six ANN architectures were selected; the networks and their architectures are given in Table 2.

The performances of the models in both training and testing databases were evaluated by comparing several parameters. The results were analysed in terms of coefficient of determination (R^2), variance factor (VAF) and root mean square error (RMSE) values, as given by Equation (11). The coefficient of determination would be a good measure of the proximity of real-life and predicted values [25]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - p_i)^2}{\sum_{i=1}^{n} \left[y_i - \frac{1}{n} \sum_{i=1}^{n} y_i \right]^2} \tag{11}$$

where y_i is the actual output, p_i the modelled value and n the length of the dataset. As the R^2 values verge on 1, a more successful model is obtained. On the other hand, the variance account for VAF parameter, given by Equation (12), is another descriptor of model performance, when this value converges to 100% [21]:

$$\mathrm{VAF}(\%) = \left[\frac{\mathrm{var}(y) - \mathrm{var}(y - p)}{\mathrm{var}(y)} \right] \times 100\% \tag{12}$$

where var(y) is the variance of real-life data and p the predicted value. Another measure of a model's performance is the RMSE, defined by Equation (13), as described by Gokceoglu [26]:

$$\mathrm{RMSE} = \left(\frac{\sum_{i=1}^{n} (y_i - p_i)^2}{n} \right)^{0.5} \tag{13}$$

It is clear that, when the RMSE value equals to 0, the performance of the model shows excellence.

As shown in Figures 3 and 4, the performance of the selected models can be observed to a reasonable degree. In Figure 3, IT parameter observations and predictions of MR2, ANN1 and ANN3 were graphically demonstrated.

ANN1 and ANN4 models were not the best of the best models; however, taking the two parameters into account, the two models were taken into account for a comparative approach. Figure 3 should be looked over using Table 2. In Table 2, the model and equation labels, epoch number, used transfer functions and the R^2, and VAF and RMSE values obtained in training and testing phases were given. As mentioned before, additional combinations of network architectures, transfer functions and epoch numbers were employed; nevertheless, because of space requirements, selected models are instructed here. Regarding the results of ANN3 and ANN5, the number of passes in learning phase does not have a remarkable effect on the performance of the models. ANN5 model best learned the behaviour of the input data, and the R^2, VAF and RMSE values are calculated as 0.98, 97.52 and 0.36, respectively. The multiple regression analyses showed that performance of MR1 and MR2 was better than MR3. Selected MR1 model gave R^2, VAF and RMSE values 0.87, 87.47 and 0.80, respectively. Training results implied that, ANN1, ANN3 and ANN5 models seemed to converge to the behaviour of input data. Analysing the results of the testing phase, as in the training phase, ANN1, ANN3 and ANN5 models seemed to model the IT parameters, appropriately. For instance, the R^2, VAF and RMSE values calculated in the testing phase of ANN1 model were 0.93, 93.28 and 0.56, respectively. Multiple regression models were worse-fit to observed values with R^2 values ranging between 0.77 and 0.81, respectively. The use of log-sig transfer function could adversely affect the

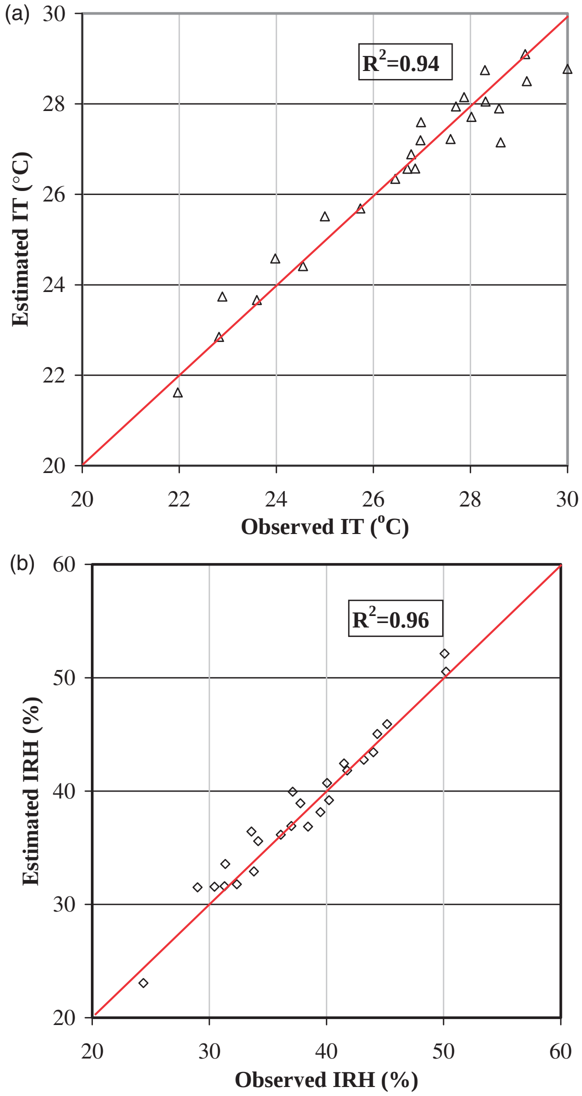

Fig. 5. Scatter plots of ANN3 model outputs and measured values of: (a) IT parameter and (b) IRH parameter.

model performance, as derived from the training and testing phase results of ANN2 model.

The second part of Table 2 includes the results related with the IRH prediction capabilities of the models under consideration. The comparative graphical representation of predictions of several models and observed values in Figure 4 imply that ANN models would be better in the estimation of IRH parameter in comparison with the multiple regression models, as expected.

A closer examination of the parameters characterizing the success of the models in training phase would give information similar to the results obtained in IT parameter prediction. Performance of ANN models seemed to be fairly better than multiple regression models. The R^2, VAF and RMSE values of multiple regression models were in the ranges 0.85–0.87, 85.35–87.03 and 2.30–2.45, respectively. Analysing the results of ANN models, the parameters lying in the range 0.87–0.97, 81.77–96.28 and 0.88–0.96, respectively could be determined. ANN1, ANN3 and ANN5 models would distinguish from the remaining models with their high R^2 values and minimized errors. In addition to these, as shown in Table 2, the best models in prediction of IRH parameter are ANN1, ANN3 and ANN5, respectively. The R^2, VAF and RMSE values of ANN1 model were 0.91, 90.29 and 0.99, respectively. The regression models concluded slightly worse results in IRH prediction, in comparison with the training phase. The use of tan-sig function as transfer function is recommended, as can be derived from the results. The most successful estimator of IT and IRH parameters was selected as ANN, and the scatter plots of measured and predicted values were given in Figure 5.

The results explicitly illustrated that ANN3 model is a useful tool for modelling IT and relative humidity in education buildings.

Conclusions

Several ANN models and multiple regressions were compared to estimate the daily indoor temperature and relative humidity values of an existing educational building to an acceptable accuracy level. It was concluded that performances of ANN models were better than those of multiple regression models. In the testing phase of the best ANN model, the R^2 values between predicted and actual values of indoor temperature and indoor relative humidity were computed as 0.94 and 0.96, respectively. The prediction of daily indoor temperature and relative humidity values generated using ANN model would be beneficial in:

- evaluation of indoor thermal comfort conditions and energy consumption level in selected buildings;
- determination of appropriate size of HVAC systems for the buildings and to make necessary modifications in existing HVAC systems; and
- prediction of daily indoor temperature and relative humidity using outdoor climatic variables.

References

1 Kaygusuz K: Energy situation, future developments, energy savings, and energy efficiency in Turkey: Energy Source Part A 1999;21(5):405–416.
2 Yüksel I: Energy production and sustainable energy policies in Turkey: Renew Energy 2010;35:1469–1476.
3 TÜBİTAK: Energy and natural resources panel-additional Reports: EK-A: Turkey, The World and OECD data [Enerji Ve Dogal Kaynaklar Paneli-Rapor Ekleri EK-A: Türkiye, Dünya ve OECD Verileri], 2003.
4 Hepbasli A, Utlu Z: Comparison of Turkey's sectoral energy utilization efficiencies between 1990 and 2000 Part 2: Residential- commercial and transportation sectors: Energy Source Part A 2004;26(14):1345–1355.
5 Holz R, Hourigan A, Sloop R, Monkman P, Krarti M: Effects of standard energy conserving measures on thermal comfort: Build Environ 1997;32(1):31–43.
6 Tham KW, Ullah MB: Building energy performance and thermal comfort in Singapore: ASHRAE Trans 1993;99(1):308–321.
7 Dombaycı ÖA, Gölcü M: Daily means ambient temperature prediction using artificial neural network method: a case study of Turkey: Renew Energy 2009;34:1158–1161.
8 Catalina T, Virgone J, Blanco E: Development and validation of regression models to predict monthly heating demand for residential buildings: Energy Build 2008;40:1825–1832.
9 Dodier RH, Henze GP: Statistical analysis of neural networks as applied to building energy prediction: Trans ASME 126 2004;126(1):592–600.
10 Thomas B, Soleimani-Mohseni M: Artificial neural network models for indoor temperature prediction: investigations in two buildings: Neural Comput Appl 2007;16:81–89.
11 Lu T, Viljanen M: Prediction of indoor temperature and relative humidity using neural network models: model comparison. Neural Comput Appl 2009;18:345–357.
12 Parishwad GV, Bhardwaj RK, Nema VK: Prediction of monthly-mean hourly relative humidity, ambient temperature, and wind velocity for India: Renew Energy 1998;13(3):363–380.
13 Mohseni MS, Thomas B, Fahlen P: Estimation of operative temperature in buildings using artificial neural networks: Energy Build 2006;38:635–640.
14 Ruano AE, Crispim EM, Conceicao EZE, Lucio MMJR: Predictions of building's temperature using neural networks models: Energy Build 2006;38:682–694.
15 Yigit KS, Ertunc HM: Prediction of the air temperature and humidity at the outlet of a cooling coil using neural networks: Int Commun Heat Mass Transfer 2006;33:898–907.
16 Reginato BC, Freire RZ, Oliveira GHC, Mendes N, Abadie MO: Predicting the temperature profile of indoor buildings by using orthonormal basic functions: Eleventh International IBPSA Conference Glasgow, Scotland, July 27–30, 2009.
17 Mustafaraj G, Chen J, Lowry G: Thermal behaviour prediction utilizing artificial neural networks for an open office: Appl Math Model 2010;34:3216–3230.
18 Perez YV, Capeluto IG: Climatic considerations in school building design in the hot–humid climate for reducing energy consumption: Appl Energy 2009;86:340–348.
19 TS 825: Thermal Insulation in Buildings. Ankara, Turkey, Turkish Standard Institution, 1998.
20 Haykin S: Neural Networks: A Comprehensive Foundation, 2nd edn. Englewood Cliffs, NJ, Prentice Hall, 1998, pp. 88–90.
21 Göktepe AB, Altun S, Altintas G, Tan O: Shear strength estimation of plastic clays with statistical and neural approaches: Build Environ 2008;43:849–860.
22 Fox J: Applied Regression Analysis, Linear Models, and Related Methods. Thousand Oaks, CA, Sage Publications, 1997, pp. 369–436.
23 Draper NR, Smith H: Applied Regression Analysis. New York, John Wiley & Sons, 1998, pp. 505–553.
24 Givoni B: Man, Climate and Architecture, 2nd edn. London, UK, Applied Science Publishers Ltd., 1976, pp. 50–100.
25 Everett BS: The Cambridge Dictionary of Statistics, 2nd edn. Cambridge, Cambridge University Press, 2002, pp. 102–105.
26 Gokceoglu C: A fuzzy triangular chart to predict the uniaxial compressive strength of the Ankara agglomerates from their petrographic composition: Eng Geol 2002;66:39–51.