Int. J. Econ. Environ. Geol. Vol. 12 (2) 87-91, 2021

Journal home page: www.econ-environ-geol.org

Open Access

ISSN: 2223-957X

Neural Network and Regression Methods for Estimation of the Average Daily Temperature of Hyderabad for the Years 2018-2020

Adeel Tahir,[*1] Mamnoon Akhter,[2] Zaheer Uddin,[1] Muhammad Sarim[3]

[1]Department of Physics, [2]Department of Applied Physics, University of Karachi, Pakistan
[3]Department of Computer Science, Federal Urdu University of Arts S&T, Karachi, Pakistan

*Email: adeel.tahir@fuuast.edu.pk

Received: *25 Febrary, 2021* ***Accepted:*** *30 March,2021*

Abstract: A qualitative study on temperature distribution has been executed in Hyderabad by several researchers. This study, however, is the first attempt to study temperature distribution quantitatively. Two different methods, i.e., Artificial Neural Network (ANN) and Regression Analysis (RA), have been used to determine the average daily temperature distribution for Hyderabad, a city in Pakistan. Both the methods are used to predict the average daily temperature of the years; 2018, 2019, and 2020. In Neural Network (NN) analysis, the network was trained and validated for three years with temperature recorded from 2015-2017. With the help of training and validation parameters of the hidden layer, the average d aily temperature was predicted for 2018-2020. Based on input parameters (dew point, relative humidity, and wind speed), a multiple regression model was developed, and average daily temperature for the years 2018-2020 was predicted again. For validation of the model statistical errors, Root Mean Square Error (RMSE), Mean Absolute Error (MABE), Mean Absolute Percent Error (MAPE), and coefficient of determination are calculated. The statistical errors show that multiple regression models and neural network models provide a good prediction of temperature distribution. However, the results of the neural network are better than the regression model.

Keywords: Artificial neural network (ANN), regression analysis, average daily temperature, Hyderabad city.

Introduction

The climatic change can be predicted by gathering information, observing, and forecasting weather events over many years (Ukhurebor et al., 2017a). It depends on sea, longitude, and latitude (Pondya et al., 2011) and impacts food security. The term "weather" describes the condition of air on the earth's surface at a particular time and place. Forecasting is the act of making an educated guess in computing future forecasts by relying on present and past data (Avazbeigi et al., 2010). The key service rendered by the metrological community is forecasting weather conditions. Though, it is still challenging for many government and private agencies (Sharma and Datta, 2007). Temperature is a significant metrological parameter in various fields, including tourism, agriculture, and economics. In these fields, the importance of weather forecasts is crucial, since forecasting future conditions is a frequently required phenomenon (Al-Matarneh et al., 2014).

Artificial neural networks are powerful techniques for non-linear models as they do not require a mathematical expression for any complex phenomena (Kumar, M., et al., 2002). Using the ANN technique, the maximum temperature of the winter season in Iran and Tehran was predicted. 70% of data was assigned for training and the remaining 30% for testing and validation. ANN model based on three input layers, nine hidden layers, and one output layer Hyperbolic tangent function used in the hidden layer. Several statistical parameters are calculated, which are in the acceptable range (Nezhad, et al., 2019).

The temperature prediction technique has been the subject of numerous studies. Radhika and Shashi (2009) reported that the support vector machine could forecast the weather by considering time-series data. Hayati and Mohebi (2007) used meteorological data from 1996-2006 to train and test an Artificial Neural Network based on human comfort. Thunderstorms probability, rain, fog, and frost forecasting meteorologists adopted the dew point temperature irrespective of relative humidity (Górnicki et al., 2017; Yousif and Tahir, 2013). A lower atmosphere greater than 60 °C of dew point shows intense thunderstorm probability (Ukhurebor et al., 2017b).

Multiple Regression Analysis is a statistical analysis to deduce a relation between a dependent and various independent parameters. They are widely used for two reasons which are conceptually and significantly different. It is generally used for forecasting, and its application overlaps considerably with that of machine learning. Also, it can be used to conclude causal associations between independent and dependent variables in a particular case. A regression analysis method may provide useful information to researchers (Cook et al 1982, Gupta et al, 2017)

In the last century, in the climate variation research field, the main focus was a variation of surface air temperature. In Pakistan, the air temperature was studied from 1882 to 2003. Normal distribution fit an analytical analysis and suggested temperature increases from 0.3 °C to 1.0 °C in the twentieth century. From 1882 to 1960 and 1961 to 2000, the Bayesian analysis computational technique showed that the average annual mean temperature for the first period is less than the second period. The general pattern of changing climate is like the current global warming configuration throughout Pakistan by Bayesian inference (Iqbal and Quamar, 2011).

Therefore, this study aims to forecast the temperature of Hyderabad by using artificial neural network technique and Multiple Regression Analysis, keeping in view the basic weather parameters such as humidity, dew point, and wind speed.

Study Area

Hyderabad is the capital of the Hyderabad Division, which is located in Pakistan's Sindh province (Fig. 1). It is located at 25.367 °N latitude and 68.367 °E longitudes and has a tropical desert climate with mild temperatures all year (Jatt, 2016, WS, 2021). As a result, typical Hyderabadi homes have "wind-catching" towers that funnel breezes down into living quarters to reduce heat (Talpur, 2007).

The winter season starts in December and continues till February. The Western Disturbance influences the winter season. Winters are mild, with highs about 25 °C (77 °F) and lows below 10 °C (50 °F). The Spring season mostly begins in March, which lasts until April. Hyderabad's spring season is almost unobserved due to the city's dry climate. Autumn starts from October to November. Autumn is characterized by hazy and dry weather. The perceived humidity in Hyderabad varies enormously throughout the year (https://weathers park.com).

Fig. 1 Geographical location of Hyderabad in Sindh, Pakistan.

Materials and Methods

Artificial Neural Network (ANN)

An Artificial Neural Network consists of interconnecting artificial neurons. These neurons are a model of a neuron in the biological brain. In Artificial Neural Network, the computer learns the behaviour of the input data and performs several tasks to train the machine. The process of training is the adjustment of weights of edges connecting neurons. Once the machine is trained efficiently, the output data is predicted (Agatonovic, 2000, Malone, 1955, Maqsood, 2004)

Artificial Neural Network (Agatonovic and Beresford, 2000; Malone, 1955; Maqsood et al., 2004) predicts the average daily temperature of Hyderabad city by taking three variables, i.e., dew point, relative humidity, and wind speed as input. The architecture of ANN used in this study is shown in Fig. 2.

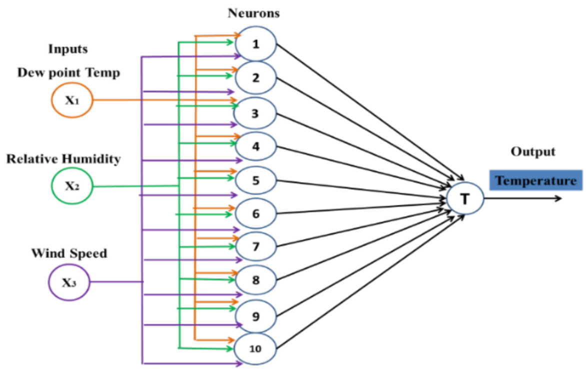


Fig. 2 Feedforward neural network with ten neurons in hidden layers

Three input parameters, i.e; dew point, relative humidity, and wind speed are used to estimate the temperature. This prediction model is developed on MATLAB, which is a three-layered feed-forward network. The input layer contains three neurons without any hidden layer. Ten neurons are included in the hidden layers to ensure good fitting of the input to output variables. Hidden layer neurons have sigmoidal transfer function while neuron in output layer has linear transfer function. The output of the only neuron in the output layer represents estimated temperature (T), and it is given by

$$T = \sum_{i=1}^{10} w_i H_i + B \tag{1}$$

Where bias B=1.366025 and weights w_i are given in table 1. Here, H_i can be calculated by the following equation, and E_i can be calculated using the following formula,

$$H_i = \frac{1}{1 + e^{-E_i}} \tag{2}$$

$$E_i = W_{1i}X_1 + W_{2i}X_2 + W_{3i}X_3 + b_i \tag{3}$$

Where X_1, X_2, and X_3 represent inputs dew point, relative humidity, and wind speed, respectively. b_i is the bias associated with ith neuron of the hidden layer, and W_{ji} is the weight of branch connecting ith neuron and jth input.

Multiple Regression Models

A multiple regression model (see equation 4) is suggested in which a predictor of a dependent variable (average daily temperature) is given as a function of three independent variables. The independent variables are dew point, relative humidity, and wind speed. The suggested model is linear. In equation (4) a_o, a_1, a_2 and a_3 are regression coefficients. T is the average daily temperature, T_d is the average dew point, Rh is the relative humidity, and V is wind speed.

$$T = a_o + a_1 T_d + a_2 Rh + a_3 V \quad (4)$$

Statistical Errors for Validation of Developed Models

To check the validity of both models, the Mean Square Error (MSE), Mean Absolute Error (MABE), Mean Absolute Percent Error (MAPE), and Coefficient of determination R^2 are calculated using (5) to (8) (Uddin et al., 2019).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(T_{c.i} - T_{m.i})^2 \quad (5)$$

$$MABE = \frac{1}{n}\sum_{i=1}^{n}|T_{ci} - T_{mi}| \quad (6)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(T_{ci} - T_{mi})}{T_{m.i}}\right| x100 \quad (7)$$

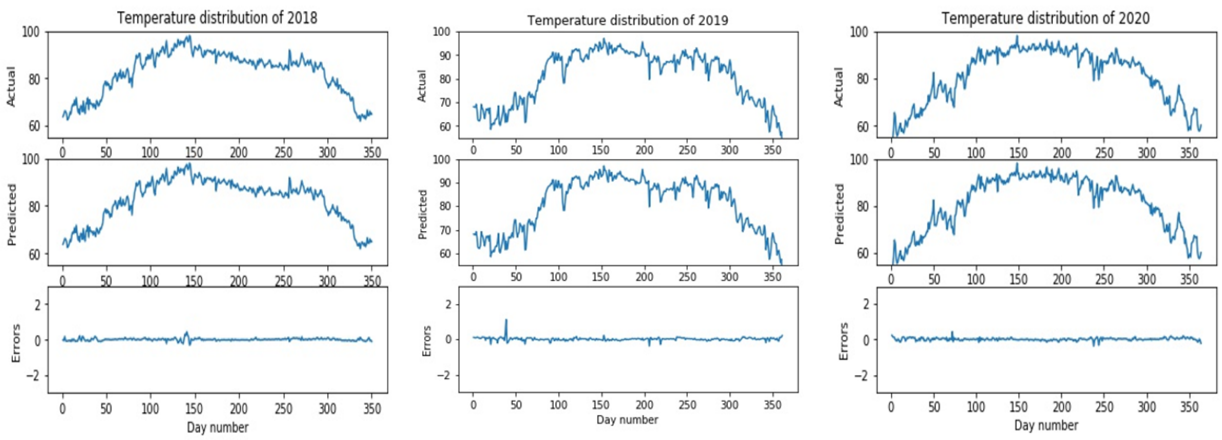
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(T_{ci} - T_{mi})^2}{\sum_{i=1}^{n}(T_{ci} - \overline{T_m})^2} \quad (8)$$

Where T_{ci} and T_{mi} are predicted and recorded temperatures for ith data point respectively.

Results and Discussion

The prediction of average daily temperature is one of the essential tasks in forecasting. Temperature plays a significant role for different sectors such as the industry of renewable energy, agriculture, and the daily life of the common man Water contents in the soil are also dependent on temperature, which directly affects processes like seed germination, plant growth, etc. In designing houses, hospitals etc., the variation in daily temperature is also considered. Scientists and researchers have employed various techniques to forecast temperature. The use of Artificial Neural networks has emerged as a modern technique to perform this task effectively. This study employs a feed-forward ANN to estimate average daily temperature. It consists of three input variables (dew point, relative humidity, and wind speed), a hidden layer with 10 neurons, and the output variable (Average Daily Temperature). The network is trained using 1092 data (2015-2017) points (average daily temperature, dew point, relative humidity, and wind speed) using Levenberg–Marquardt algorithm. 764 out of these 1092 samples were used to train this network, while 164 were used for validation and testing each.

Later, this trained network was used to estimate 1086 values of temperatures ranging from 2018 to 2020. Predicted and actual values along with errors are plotted in fig. 3(a) and 3(c). The weight vectors W_1 and W_2 of the neuron which connects different nodes are determined and presented in table 1. The table also gives the values of the bias vector in equation (3). The bias 'B' in equation (1) is found to be 1.366025. The weights w_i at nodes connecting neurons in the hidden layer and outside are also presented in table 1.



(a) (b) (c)

Fig. 3 Comparison of predicted and actual average daily temperature of Hyderabad for the years (a) 2018, (b) 2019, (c) 2020.

Table 1 Weights of the input variable, and neurons and bias associated with the neurons.

I	w_i	W_{1i}	W_{2i}	W_{3i}	b_i
1	0.195268	2.278503	-1.20393	0.097574	-1.72063
2	0.026207	0.842497	1.658171	-1.03971	-2.41446
3	-1.85264	-0.67074	0.352121	-0.00652	-0.20721
4	-0.00349	2.544566	-0.36235	2.646378	-1.5808
5	0.006889	1.111571	3.274038	2.140504	-0.16471
6	-0.00525	-1.91097	-2.21042	2.116281	0.020484
7	-0.01835	-3.03678	-1.78265	3.528702	-1.73505
8	-1.67435	-0.234	1.185217	0.087971	1.680052
9	0.048303	0.559745	1.563312	0.770618	1.766489
10	-0.36315	0.374902	2.190196	-0.43646	2.457451

A multiple regression model is used for the prediction of the average daily temperature for Hyderabad. To develop the model for predicting average daily temperature, six-year climate data from 2017-2020 is used. The multiple regression model consisting of three independent variables, which are: dew point, relative humidity, and wind speed are used for predicting average daily temperature. The estimated model is given below;

$$\boldsymbol{T = 47.162849 + 1.0832388T_d - 0.6146813Rh - 0.00433395\,V}$$

Using equation (9), the average daily temperature for three years i.e. (2018-2020) is predicted. To validate the regression model, the predicted values are compared with the actual adequacy of the model that is assessed by (MSE), (MABE), (MAPE), and R- square.

Figure 4(a) shows the graph of recorded temperature distributions, as predicted by the regression model for 2015-2017. The regression coefficients determined by data of 2015-2017 coefficients are used to predict temperature distribution for the period from 2018-2020 (see fig. 4(b)-4(d).

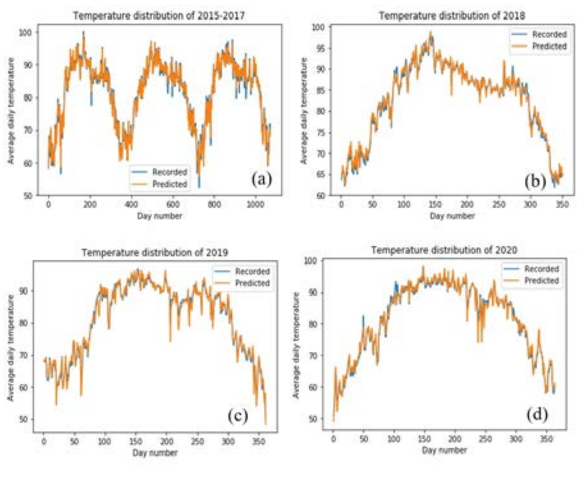


Fig. 4 The recorded and predicted temperature distribution for the years (a) 2015-2017, 2018, (c)2019. (d) 2020.

The statistical errors for both ANN and regression models are calculated and presented in table 2. The values of statistical errors are lower in the case of ANN; the coefficient of determination between predicted and actual average daily temperatures is higher for ANN. The maximum absolute difference between measured and calculated temperature by the model was found to be most significant (7.14 ºF) in 2020, and the lowest is 1.25 ºF for the year 2018. Table 2 shows the values of statistical errors for each model and each year from 2018-2020.

Table 2 Statistical Errors in the predicted temperature distribution of 2018-2020 for regression and ANN models.

	Year	RMSE	MABE	MAPE	R^2
Regression Model	2018	0.924770	0.717016	0.894937	0.991340
	2019	1.270538	0.923233	1.180832	0.986885
	2020	1.653165	1.319912	1.705941	0.981612
ANN Model	2018	0.072626	0.050767	0.063599	0.999971
	2019	0.096517	0.059109	0.079051	0.999965
	2020	0.073636	0.052224	0.069465	0.999981

Conclusion

In this study, two models are developed to estimate the average daily temperature for Hyderabad city, Pakistan. The first model is a linear regression model based on three predictors, i.e., dew point, relative humidity, and wind speed. Regression coefficients are estimated using data values of 2015-2017, and this model is used to estimate average daily temperatures of 2018-2020. The second model is ANN; the modelling process comprises two steps. First is the training phase in which the model was trained on temperature variations for 2015-2017 by taking dew point, relative humidity, and wind speed as inputs. The trained model predicts the average daily temperature for three years (i.e. from 2018 to 2020). The predicted results obtained by ANN were excellent. The prediction errors are in the range of ±0.5 for the years 2018 and 2020, whereas, found in the range of ±1 for the year 2020. The RMSE, MABE, and MAPE are lower in ANN (between 0 and 0.1), whereas the same errors have higher values for the multilinear regression model (between 0 and 2). The coefficient of determination is also calculated to check the correlation between predicted and actual daily temperatures, which is found as 0.981 to 0.999 for the multilinear regression model. At the same time, 0.9999 in the case of ANN confirms the suitability of the second method.

In conclusion, the quantitative study of temperature distribution for Hyderabad is done through ANN and multilinear regression. This type of study can also be carried out for any other city. Though both methods give good quantitative results, the ANN approach provides better results than multilinear regression.

References

Agatonovic-Kustrin, S., Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J. Pharm. biomed. anal*., **22** (5), 717-727.

Al-Matarneh, L., Sheta, A., Bani-Ahmad, S., Alshaer, J., Al-Oqily, I. (2014). Development of temperature-based weather forecasting models using neural networks and fuzzy logic. *Int. j. Multimedia Ubiquitous Eng.,* **9** (12), 343-366.

Avazbeigi, M., Doulabi, S. H. H., Karimi, B. (2010). Choosing the appropriate order in fuzzy time series: a new n-factor fuzzy time series for prediction of the auto industry production. *Expert syst. appl*., **37** (8), 5630-5639.

Bair, F. E., Ruffner, J. (1977). The weather almanac. 6th edition, Avon Books, USA.

Cook, R. D., Weisberg, S. (1982). Criticism and influence analysis in regression. *Soc. met.*, *13*, 313-361.

Górnicki, K., Winiczenko, R., Kaleta, A., Choiska, A. (2017). Evaluation of models for the dew point temperature determination. *Tech. sci.*, **20** (3), 241-257.

Gupta, A., Sharma, A., Goal, D. A. (2017). Review of Regression Analysis Models. *International Journal of Engineering Research & Technology (IJERT)*, **6** (8), 58-61.

Hayati, M., Mohebi, Z. (2007). Application of artificial neural networks for temperature forecasting. *World acad. sci. eng. technol.*, **28** (2), 275-279.

Iqbal, M. J., Quamar, J. (2011). Measuring temperature variability of five major cities of Pakistan. *Arab. J. Geosci.*, **4** (3-4), 595-606.

Jatt, Z. R. (2016). Aesthetics and organization of spaces: a case study of colonial era buildings in hyderabad, Sindh. *J. Archit. Plann. Res.,* **20**, 30-40.

Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., Pruitt, W. O. (2002). Estimating evapotranspiration using artificial neural network. *J. Irr. Dra. Eng***.**, **128** (4), 224-233.

Malone, T. F. (1955). Application of statistical methods in weather prediction. In *Proceedings of the National Academy of Sciences of the United States of America*, **41** (11), 806.

Maqsood, I., Khan, M. R., Abraham, A. (2004). An ensemble of neural networks for weather forecasting. *Neural. comput. appl.*, **13** (2), 112-122.

Nezhad, E. F., Ghalhari, G. F., Bayatani, F. (2019). Forecasting maximum seasonal temperature using artificial neural networks "Tehran case study". *Asia-Pacific J. Atm. Sci.*, **55** (2), 145-153.

Pondyal, K. N., Bhattarai, B. K., Sapkota, B., Kjeldstad, B. (2011). Solar radiation potential at four sites of Nepal. *J. Inst. Eng.*, **8** (3), 189-197.

Radhika, Y., Shashi, M. (2009). Atmospheric temperature prediction using support vector machines. *Int. J. Comput. Theory Eng*., **1** (1), 55.

Schoen, C. (2005). A new empirical model of the temperature–humidity index. *J. Appl. Meteorol.*, **44** (9), 1413-1420.

Sharma, A., Datta, U. (2007). A weather forecasting system using concept of soft computing: a new approach. In *AIP Conference Proceedings*, American Institute of Physics, **923** (1), 275-287.

Steadman, R. G. (1979). The assessment of sultriness. part i: a temperature-humidity index based on human physiology and clothing science. *J. appl. meteorol. climatol.,* **18** (7), 861-873.

Steadman, R. G. (1984). A universal scale of apparent temperature. *J. Appl. Meteorol. Climatol.*, **23** (12), 1674-1687.

Talpur, M. A. M. (2007). The vanishing glory of Hyderabad (Sindh, Pakistan). *Web J*., **1**, 47-65.

Uddin, Z., Khan, M. B., Zaheer, M. H., Ahmad, W., Qureshi, M. A. (2019). An alternate method of evaluating lagrange multipliers of MEP. *SN Appl. Sci.*, **1** (3), 1-10.

Ukhurebor, K. E., Abiodun, I. C., Bakare, F. (2017b). Relationship between relative humidity and the dew point temperature in benin city, Nigeria. *J. Appl. Sci. Environ. Manage.*, **21** (5), 953-956.

Ukhurebor, K. E., Batubo, T. B., Abiodun, I. C., Enoyoze, E. (2017a). The influence of air temperature on the dew point temperature in Benin city, Nigeria. *J. Appl. Sci. Environ. Manage.*, **21** (4), 657-660.

WS (Weather Spark). (2021). https://weatherspark. com/y/106562/Average-Weather-in-Hyderabad-Pakistan-Year-Round (Accessed on 1st January 2021).

Yousif, T. A., Tahir, H. M. (2013). The relationship between relative humidity and the dew point temperature in Khartoum state, Sudan. *J. Appl. Ind. Sci.*, **1** (5), 20-23.