Proceedings of the Pakistan Academy of Sciences: A Physical and Computational Sciences 59(3):45-53 (2022)

ISSN (Print): 2518-4245; ISSN (Online): 2518-4253
https://doi.org/10.53560/PPASA(59-3)654

Pakistan Academy of Sciences

Research Article

Determination of Temperature Distribution of Rohri, Sindh using Artificial Neural Network and Regression Analysis

Adeel Tahir[1*], Muhammad Ashraf[1], Zaheer Uddin[2], Muhammad Sarim[3], and Syed Masood Raza[1]

[1]Department of Physics, Federal Urdu University of Arts Science and Technology, Karachi, Pakistan
[2]Department of Physics, University of Karachi, Karachi, Pakistan
[3]Department of Computer Science, Federal Urdu University of Arts Science and Technology, Karachi. Pakistan

Abstract: As time passes, the world is facing the problem of global warming, which results in a rise in average daily temperature. Proper knowledge of temperature distribution and future prediction may help to cope with the situation in the near future. Climate forecasting has gone through various faces; in the early days' people used to predict the behavior qualitatively. Now environmental scientists have developed a quantitative method for forest climate behavior with certain uncertainties. Empirical models have been developed based on regression analysis to estimate temperature distribution. Two models, linear and non-linear, use dew point temperature and relative humidity as independent variables. In addition to regression analysis, Artificial Neural Network (ANN) has been utilized to predict the average daily temperatures of Rohri Sindh, a city in Pakistan in the Sindh province. Both empirical models and ANN estimates are in good agreement with the known values of average daily temperatures.

Keywords: Artificial Neural Network, Estimated Model, Temperature Distribution, Forecasting, Rohri

1. INTRODUCTION

People have been forecasting climate changes manually since ancient times. Today, daily climate information can be recorded automatically due to technological developments. Once the data is recorded, it is projected to forecast the climatic information for the future. But the chaotic state of climate still causes forecast errors which makes the forecast prediction unreliable for a long time period. Therefore, further investigation is required to address the challenges in predicting the weather forecast. Reliable weather forecast is beneficial for humans in different aspects and they can prepare themselves to face the challenges in upcoming disasters. Although the researchers have developed various tools and made a significant contribution to metrology, it still requires further studies to overcome the existing challenges in weather forecasts [1].

Changing climate is the World's first environmental issue that presents a risk to humans. Industries are highly influential in causing this issue and the recent global warming that the planet has experienced. Extreme changes in air temperature can harm plants and animals, so knowing the variability of ambient temperature is critical in agriculture [2,3]. Temperature plays a vital role in routine life. The prediction of temperature for the respective region highly depends on the accurate data observed. Forecasting air temperature helps determine the probability of tornadoes and floods in a given region [4]. Soil temperature, energy load consumption, and solar radiation also depend on accurate forecasting of temperature [2]. Artificial neural networks are powerful techniques, which do not require any mathematical expressions to solve complex non-linear problems [5]. ANN model was designed to predict the highest winter temperature in Tehran (Iran) which consists of three input

Received: July 2022; Accepted: September 2022
*Corresponding author: Adeel Tahir <adeel.tahir@fuuast.edu.pk>

layers, nine hidden layers with hyperbolic tangent function, and one output layer. The estimated statistical parameters were found in the acceptable range [6]. A temperature prediction model is proposed based on the long short-term memory (LSTM) neural network. The model also refines the missing data and may forecast the temperature time series up to 14 days [7]. Another model based on (LSTM) NN may forecast the short-term temperature of Bandung (capital of West Java) with high accuracy by using the long period data [8]. During the period 1998 to 2000, a neural network estimated Saudi Arabia's solar radiation by using temperature and relative humidity [9]. Sea surface temperature (SST) and numerical estimations by ANN were combined with a special wavelet neural network to study the forecast of six different locations in the Indian Ocean over three-time scales (daily, weekly, and monthly). The performance was assessed by statistical error analyses, which showed satisfactory results [10]. A statistical post-processing method was proposed based on the Land-Atmosphere Modeling Package Weather Research and Forecasting model, which uses a Generalized Linear model and parameter correction to increase the accuracy of numerical weather prediction (NWP). The spatial distribution of temperature and wind speed (July and January, East Asia) was analyzed to show their consistent relation [11]. The method is based on the Geographically Weighted Kigring Regression model, which was applied to forecast Poland's spatial distribution of air temperature. It extends Hengl's decision tree to select the suitable prediction model for varying temperature and environmental predictors. The Local Geographically Weighted regression model was chosen to model the deterministic part of spatial variation. Sixty-nine air temperature cases (with time aggregation from daily to mean annual temperature) were analyzed. The author claimed local prediction models for air temperature were better fitted in spatial distribution irrespective of data aggregation [12]. The daily max and min temperature series (1887 to 2019) were constructed and homogenized by Tianjin Meteorological Archive to estimate the warning trends (0.154 ± 0.013 °C per decade) for the last 130 years [13]. A physical model was developed to predict the runway surface temperature at Oslo Airport, Norway. A Now-casting model was designed to estimate the surface temperature for the next three hours. The predictions were more satisfactory at the beginning of winter [14]. A polynomial fitting technique for computing the mean daytime and a double-cosine method was introduced to estimate daily ambient temperature profiles for any place in Europe using a spatially continuous database [15]. Based on daily data set from more than 5000 meteorological stations all over the World concluded that during the last century, the mean temperature, the maximum and minimum temperature increased by 0.5 °C, 0.050 °C, and 0.018 °C, respectively. Also observed is that rate of increase in minimum temperature is greater than maximum temperature [16]. From 1977 to 1994, the air temperature of high-altitude regions such as the Middle Mountains and the High Himalayas of Nepal was studied. The mean temperature per year increases by 0.06 °C. In the Terai and Siwalik areas, a significant decreasing trend was found. In addition, relative to other seasons, winter was found to be warm [17]. In Nepal, during 30 years (1975-2005), an increasing trend of mean temperature was observed by a factor of around 0.04 °C [18]. From 1950 to 2004 period, analyzed maximum temperature of 4280 stations, minimum temperature of 4284 stations and Diurnal temperature change (DTR) of 4157 stations of the entire World and found an increase of 0.296 °C, 0.287 °C, and 0.296 °C per decade for mean, maximum and minimum temperatures respectively [19].

This paper presents the work to forecast the air temperature using an Artificial neural network and multi-regression models of Rohri, Sindh based on two independent parameters, namely relative humidity and dew point.

2. STUDY AREA

The study area namely "Rohri" is situated on the east bank of the river Indus. It is one of the famous cities of Sindh province (Pakistan). It is also the Taluka of the Sukkur district. Rohri is a metropolitan city with a population of around 70,000 people. The area of this ancient city is 1,318 km^2.

Its geographical location (see Figure 1) lies at 27° 27′ 18″ N latitude, 68° 55′ 53″ E longitude and an altitude of 66 m. Rohri has a tropical desert like weather with scorching summers and chill winters. It remains hot and dry all over the summer. The temperature usually falls between 46 °F to 111 °F in a year, with temperatures rising above 117 °F and rarely falling below 40 °F [20,21].

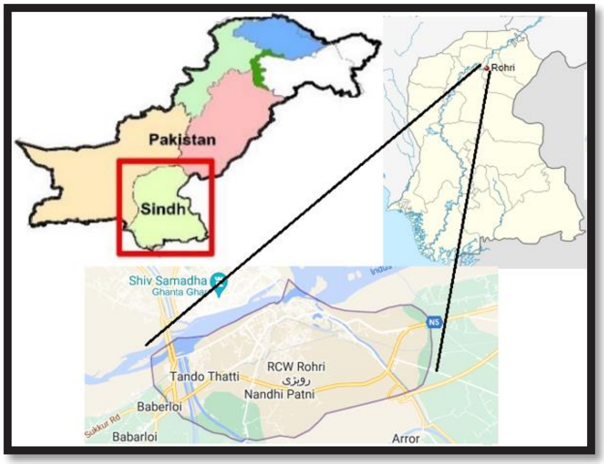


Fig. 1. Geographical location of Rohri, Sindh

3. MATERIALS AND METHODS

First weather parameters data was collected from the Pakistan Meteorological Department, Government of Pakistan, then two different models were designed to forecast the air temperature of Rohri (Sindh).

3.1 Artificial Neural Network

Artificial neural networks (ANN) are powerful mathematical tools developed from the inception of the human brain. They are widely used to solve complex problems involving non-linear behavior. ANN structure is based on three layers: the input layer, the hidden layer, and the out layer. The input layer receives data, which is then transferred to hidden layers, where the desired task is solved based on mathematical models internally designed on independent input variables. The result of hidden layers proceeds to the output layer [22]. The complete procedure involves two steps: First, an ANN is trained until it emulates according to the data provided; then, the network is used to predict the inputs not included in training data [23]. In this study, the designed ANN model is shown in Figure 2, which was utilized to estimate the mean daily temperature of Rohri city.

In this work, a feed-forward network model was constructed in MATLAB. The model predicts the temperature by taking in the dew point and relative humidity. It consists of three layers named as the input layer, hidden layer and output layer, where the input layer consists of two neurons, the hidden layer of ten neurons and the output layer of one neuron respectively. The non-linear complex mathematical model in the hidden layers was designed with a sigmoidal transfer function.

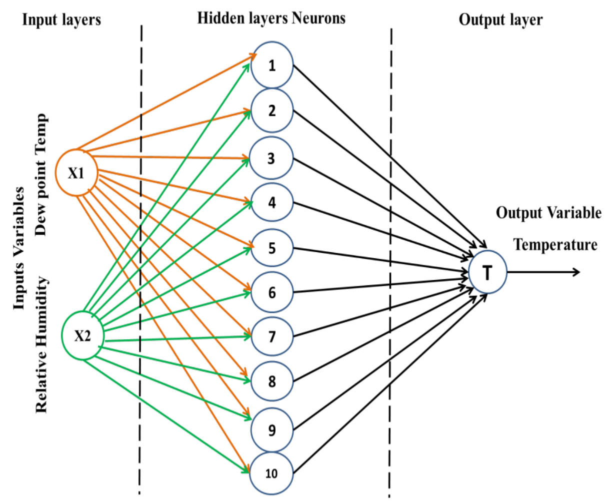


Fig. 2. 3 layered Feed-forward neural network

The output layer with one neuron contains the linear transfer function. The forecasted temperature T is given by

$$T = \sum_{i=1}^{10} \omega_i N_i + \beta \qquad (1)$$

Where bias β= 1.9578 and weights ωi are given in Table I.

Here, Ni can be calculated by the following equation

$$N_i = \frac{1}{1 + e^{-\Upsilon i}} \qquad (2)$$

ϒi can be calculated using the following formula,

$$\Upsilon i = \omega_{1i}\,\chi_1 + \omega_{2i}\,\chi_2 + b_i \qquad (3)$$

Where ωji indicates the jth synaptic link weight for ith neuron, bi the bias for that hidden neuron, χ_1 the dew point, and χ_2 the humidity.

3.2 Linear and Non-Linear Regression Models

The proposed linear and non-linear regression models are presented in the equations 4 and 5, where the bivariate functions give the daily mean temperature. Dew point and relative humidity are taken as independent input parameters.

The regression models in equation 4 are linear.

$$T = a + bT_d + c\,R_h \qquad (4)$$

Where regression coefficients are indicated by a, b and c respectively, R_h the relative humidity, while T and Td the daily mean temperature, and dew point respectively.

The regression models in equation 5 are non-linear.

$$T = a_o + a_1 T_d + a_2 R_h + a_3 \left(\frac{T_d{}^2}{R_h{}^2}\right) \qquad (5)$$

where regression coefficients are indicated by a_o,a_1,a_2 and a_3 respectively, R_h the relative humidity, while T and Td the daily mean temperature, and dew point respectively.

3.3 Statistical Measurements for Authentication of Designed Models

By using statistical techniques including coefficient of determination R^2, Mean Absolute Percent Error (MAPE), Mean Absolute Error (MABE) and Mean Square Error (MSE), we have calculated the accuracy of our designed models.

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(T_{c.k} - T_{m.k})^2 \qquad (6)$$

$$MABE = \frac{1}{n}\sum_{k=1}^{n}|T_{c.k} - T_{m.k}| \qquad (7)$$

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{(T_{c.k} - T_{m.k})}{T_{m.k}}\right| x100 \qquad (8)$$

$$R^2 = \left[1 - \frac{\sum_{k=1}^{n}(T_{c.k} - T_{m.k})^2}{\sum_{k=1}^{n}(T_{c.k} - \overline{T_m})^2}\right] \qquad (9)$$

4. RESULTS AND DISCUSSION

In this article, the temperature distribution for one of the cities (Rohri) province Sindh, Pakistan, has been investigated through two different directions, i.e., ANN and multiple regression analysis.

In ANN, the Levenberg–Marquardt algorithm with 10 hidden layers was used to train the network from 2015 - 2017 based on variables i.e dew point, relative humidity, and mean daily temperature. 70% data of this period were used to train, whereas the remaining 30% equally divide for validation and testing, respectively. As the network was trained, based on independent variable dew point and relative humidity, the mean daily temperature was forecasted for 2018 - 2020. Graphical analysis of actual and forecasted mean daily temperature with errors was found, shown in Figure 3. The ω1i and ω2i were the weights of neurons of input variables, whereas ωi represented the weights of hidden layers of neurons. The biases bi were also determined and represent all values in Table 1.

Furthermore, two empirical models, multilinear regression (model 2) and multi non-linear regression (model 3) were designed to forecast the mean daily temperature distribution. To determine regression coefficients, the related weather data from 2015 - 2017 was used to design the empirical models. The equations (see equations 10 and 11) for both models are formed by substituting equations 4 and 5, respectively. By these equations, the temperature-distribution for 2018-2020 (see Figure 4) is forecasted using dew point and relative humidity as independent input variables. To confirm the validation and performance of ANN and empirical models, statistical errors such as Mean Square Error (MSE), Mean Absolute Error (MABE), Mean Absolute Percent Error (MAPE), and R^2 were estimated. The best model was identified by comparing these calculated statistical errors.

$$T = 49.102658 + 1.061672\,T_d - 0.628384\,R_h \qquad (10)$$

$$T = 40.764341 + 0.939534\,T_d - 0.420574\,R_h + 2.988233\left(\frac{T_d{}^2}{R_h{}^2}\right) \qquad (11)$$

It has observed the statistical errors calculated for ANN model result having lowest values than both estimated (linear and non-linear regression) models as shown in Table 2. The maximum root mean square error was 4.35 ^{o}F 2.85 ^{o}F, and 1.47 ^{o}F in 2018, 2019 and 2020 by a linear model. Since the weather is unpredictable due to complexity and non-linearity, our estimated non-linear model 3 showed a better result than estimated linear model 2. The minimum errors less than 0.5 ^{o}F from 2018-2020 have been observed in non-linear model 3. The ANN model shows the more reliable results for Rohri city.

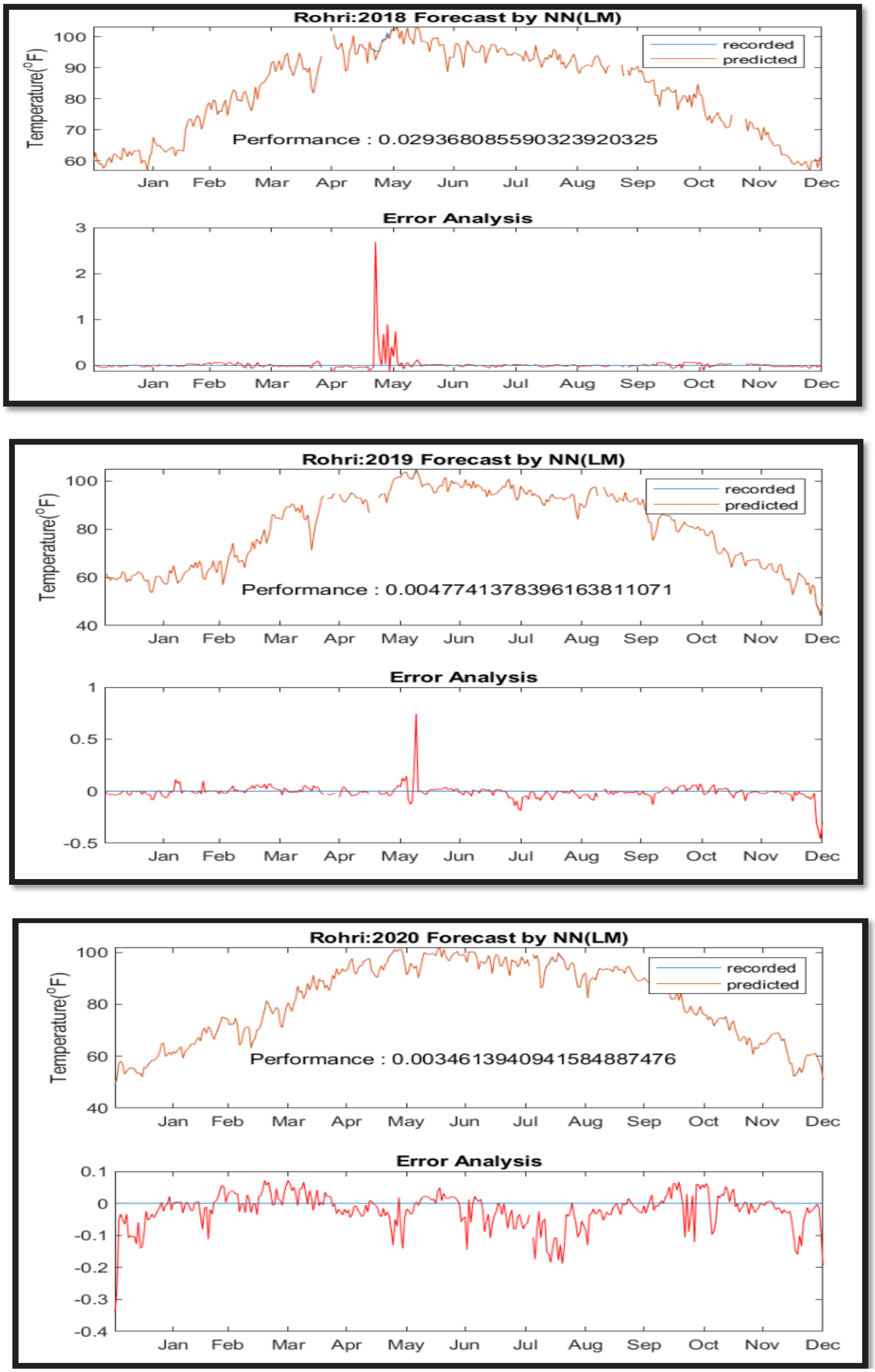
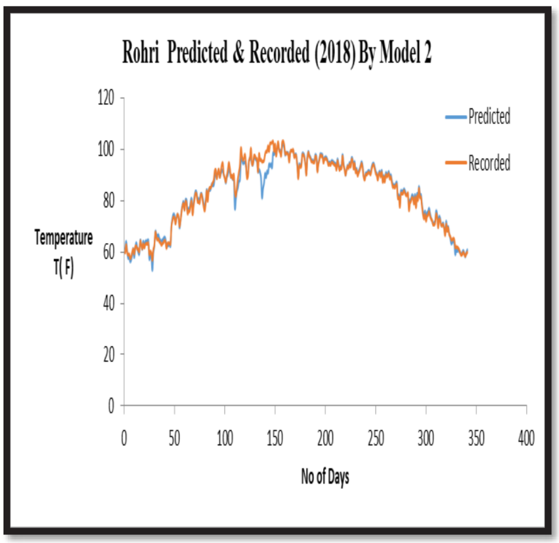
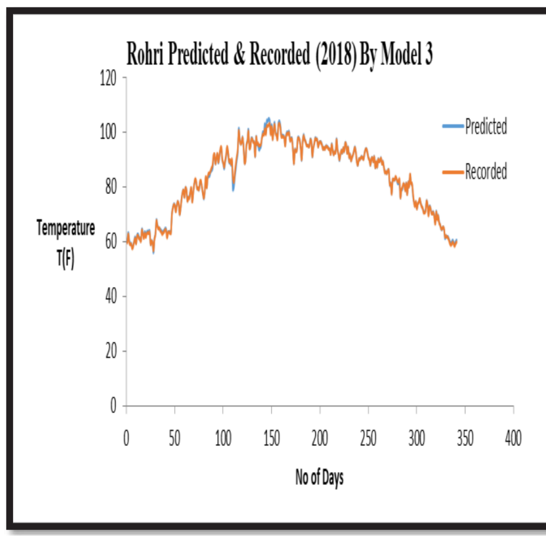


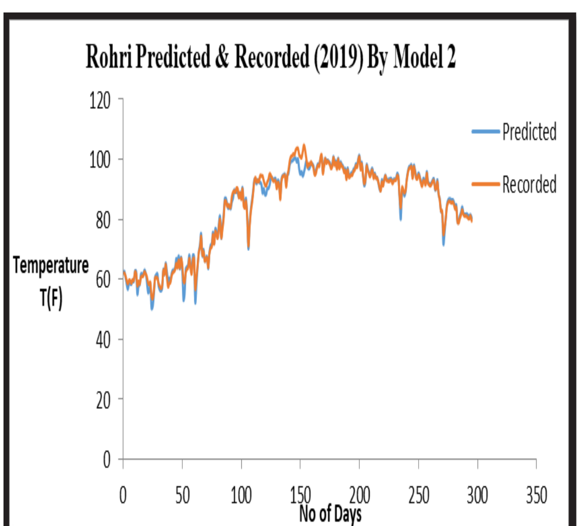
Fig. 3. Graphical analysis of recorded and predicted mean daily temperature of Rohri (2018-2020)

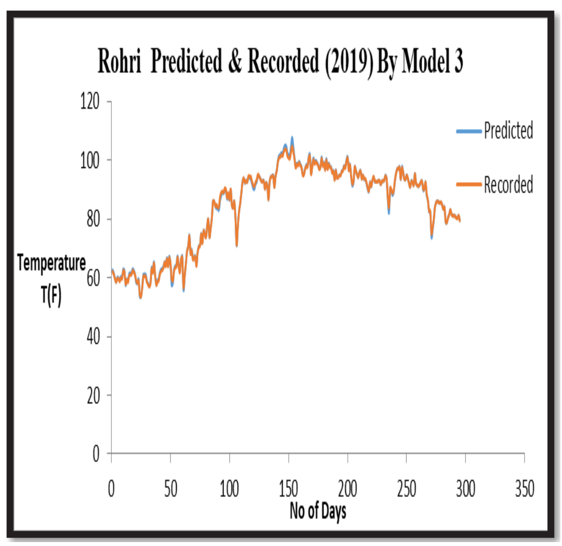
Table 1. Hidden layer neurons syanptic weights with biases

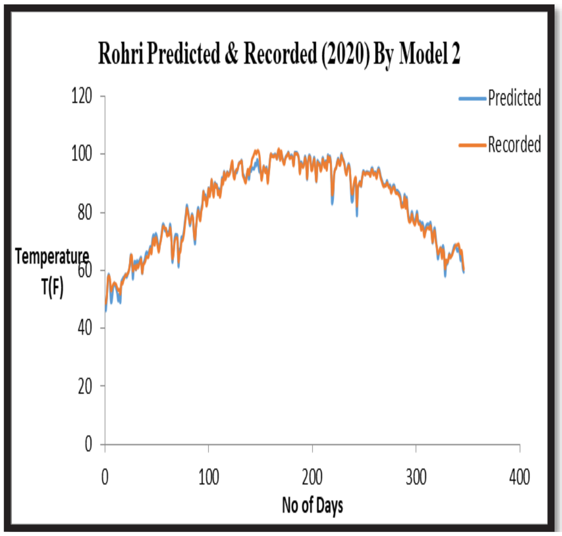
I	ω_{1i}	ω_{2i}	ω_i	b_i
1	-6.7467	-0.4713	3.5673	-7.7792
2	7.9793	0.0062	0.9250	-6.5946
3	0.7358	0.0290	-3.6938	2.7148
4	0.3354	-0.7788	-1.0415	-2.0080
5	-1.3409	0.5959	0.0524	-1.1601
6	0.1757	-3.3364	-0.3609	0.5331
7	-1.1484	0.0538	2.3892	1.1018
8	3.3007	0.0072	-2.6706	-1.9695
9	-4.1774	-0.0075	-4.6420	-3.4720
10	5.7911	-0.1191	-0.7732	4.9713











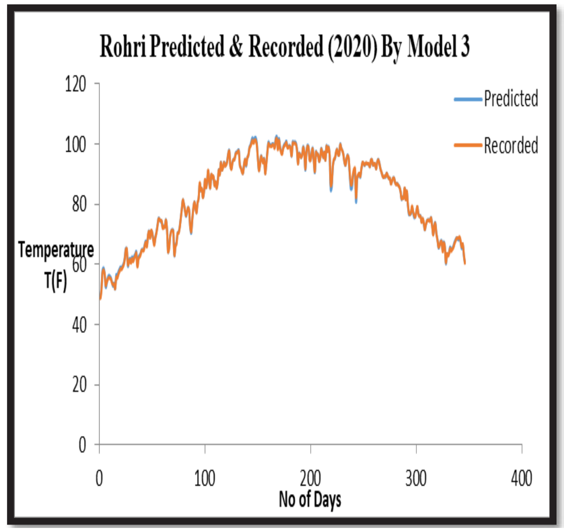


Fig. 4. The comparative analysis of temperature distribution (2018 – 2020) (a) model 2 (linear) and (b) model 3 (non-linear)

Table 2. Statistical Errors calculated between actual and forecasted temperature distribution (2018-2020) by ANN, Linear Regression, and Non-Linear Regression

Model	Year	MSE	MABE	MAPE	R^2
Model 1	2018	0.029368	0.04386	0.05016	0.99984
ANN	2019	0.00477	0.03492	0.04688	0.99997
	2020	0.00346	0.04029	0.05326	0.99998
Model 2	2018	4.3541594	1.14461	1.35716	0.97667
Linear Regression	2019	2.8494416	1.08434	1.46396	0.98886
	2020	1.4693335	0.87636	1.1713	0.99369
Model 3	2018	0.2549779	0.31374	0.38892	0.99866
Non Linear Regression	2019	0.2408168	0.32869	0.43837	0.99902
	2020	0.128415731	0.265981	0.35649	0.999484

5. CONCLUSION

This study was conducted to forecast the mean daily temperature distribution for Rohri, Sindh, Pakistan, based on three designed models including ANN model, a multi linear regression model and a non-linear multi regression. The ANN model takes in relative humidity and dew point as the input independent variables, and returns mean daily temperature as the output dependent variable. It consists of three layers including input, hidden and the output layers. We trained the machine to determine the temperature based on two independent variables, i.e., dew point and relative humidity, for the duration 2015-2017. As the machine trained, weights and bias were obtained, which were used to forecast the mean daily temperature for the next three years, 2018 - 2020. An excellent forecast was received and observed by ANN model 1. The statistical errors were less than ±0.1 for the years 2018- 2020.

Further, two estimation models were designed, multilinear regression (model 2) and non-linear multi regression (model 3). Three years 2015-2017, weather data were used to estimate the regression coefficients. During estimation, MABE for the linear and non-linear models were found ±1.07 and ±0.44, respectively. These coefficients are then employed for models 2 and 3 for forecasting temperatures for the years 2018-2020 on dew point and relative humidity.

To investigate the correlation between actual and forecasted mean daily temperature, the coefficient of determination R^2 was calculated. In all three models, R^2 determined close to 1, which shows a high correlation between them.

To see the authenticity of ANN and estimated multi regression models, statistical errors MSE, MABE, and MAPE were calculated. In ANN, the MSE, MABE, and MAPE were in the lowest range (between 0 and 0.1) as compared to non-linear model 3 (between 0.1 and 0.3) and linear model 2 (between 1.4 and 4.4).

All the model's actual and forecasted mean daily temperature distribution graphs were highly overlapped (see figure 3,4). So, we can conclude that for the Rohri city, non-linear regression model 3 showed better results to forecast the temperature than the linear regression model. Thus, it was proved that the ANN model forecast the temperature seemed more accurate than estimated empirical models.

6. ACKNOWLEDGEMENT

The metrological weather parameters data are gathered from the Pakistan Meteorological Department, Government of Pakistan. Data can be verified from website. https://en.tutiempo.net/climate/ws-417250.html.

7. CONFLICT OF INTEREST

The authors declare no conflict of interest.

8. REFERENCES

1. M. Hayati, and Z. Mohebi. Application of artificial neural networks for temperature forecasting. World Academy of Science, Engineering and Technology 28: 275-279 (2007).
2. Ö.A. Dombaycı, and M. Gölcü. Daily means ambient temperature prediction using artificial neural network method: A case study of Turkey. Renewable Energy 34: 1158-1161 (2009).
3. B.A. Smith, G. Hoogenboom, and R.W. McClendon. Artificial neural networks for automated year-round temperature prediction. Computers and Electronics in Agriculture 68: 52-61 (2009).
4. I. Tasadduq, S. Rehman, and K. Bubshait. Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia. Renewable Energy 25: 545-554 (2002)
5. M. Kumar , N. S. Raghuwanshi, R. Singh, W. W. Wallender, and W. O. Pruitt. Estimating evapotranspiration using artificial neural network. Journal of Irrigation and Drainage Engineering 128: 224-233 (2002).
6. E. F. Nezhad, G. F. Ghalhari, and F. B. Forecasting maximum seasonal temperature using artificial neural networks "Tehran case study". Asia-Pacific Journal of Atmospheric Sciences 55: 145-153 (2019).
7. I. Park, H.S. Kim, J. Lee, J.H. Kim, C.H. Song, and H.K. Kim. Temperature prediction using the missing data refinement model based on a long short-term memory neural network. Atmosphere 10: 718 (2019).
8. S. Zahroh, Y. Hidayat, R. S. Pontoh, A. Santoso, Sukono, and A. T. Bon. Modeling and forecasting daily temperature in bandung. Proceedings of the International Conference on Industrial Engineering and Operations Management Riyadh, Saudi Arabia 406-412 (2019).
9. S. Rehman, and M. Mohandes. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy 36: 571-576 (2008).
10. K. Patil, M.C. Deo, and M. Ravichandran. Prediction of sea surface temperature by combining numerical and neural techniques. Journal of Atmospheric and Oceanic Technology 33: 1715-1726 (2016).
11. J. Jeong, and S.J. Lee. A statistical parameter correction technique for WRF medium-range prediction of near-surface temperature and wind speed using generalized linear model. Atmosphere 9: 291 (2018).
12. M. Szymanowski, M. Kryza, and W. Spallek. Regression-based air temperature spatial prediction models: an example from Poland. Meteorologische

Zeitschrift 22: 577-585 (2013).

13. P. Si, Q. Li, and P. Jones. Construction of homogenized daily surface air temperature for the city of Tianjin during 1887–2019. Earth System Science Data 13: 2211-2226 (2021).
14. A.D.W. Nuijten. Runway temperature prediction, a case study for Oslo Airport, Norway. Cold Regions Science and Technology 125: 72-84 (2016).
15. T.A. Huld, M. Šúri, E.D. Dunlop, and F. Micale. Estimating average daytime and daily temperature profiles with in Europe. Environmental Modeling & Software 21: 1650-1661 (2006).
16. D.R. Easterling, B. Horton, P.D. Jones, T.C. Peterson, T.R. Karl, D.E. Parker, M.J. Salinger, V. Razuvayev, N. Plummer, P. Jamason, and C. Folland. Maximum and minimum temperature trends for the globe. Science 277: 364-367 (1997).
17. A.B. Shrestha, C.P. Wake, P.A. Mayewski, and J. E. Dibb. Maximum temperature trends in the Himalaya and its vicinity: an analysis based on temperature records from Nepal for the period 1971–94. Journal of climate 12: 2775-2786 (1999).
18. K.L. Shrestha. Global change impact assessment for Himalayan mountain regions for environmental management and sustainable development. Global Environmental Research-English Edition 9: 69 (2005).
19. R.S. Vose, D.R. Easterling, and B. Gleason. Maximum and minimum temperature trends for the globe: An update through 2004. Geophysical Research Letters 32: (2005).
20. M. Ali, and M. J. Iqbal. Extreme rainfall incidents over Sindh province, against different return periods. Proceedings of the Pakistan Academy of Sciences 50: 159-166 (2013).
21. K. Aftab, and S. A. Jilani. Sunspots Influence on Climatic Variability of Karachi and Rohri: Sunspots Influence on Climatic of Karachi. Pakistan Journal of Scientific & Industrial Research Series A: Physical Sciences 64: 52-58 (2021).
22. S. Mehdizadeh. Assessing the potential of data-driven models for estimation of long-term monthly temperatures. Computers and Electronics in Agriculture 144: 114-125 (2018).
23. J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses. Air temperature forecasting using machine learning techniques: a review. Energies 13: 4215 (2020).