International Journal of Statistics and Applications 2019, 9(5): 143-152
DOI: 10.5923/j.statistics.20190905.03

Stochastic Models and Neural Networks with Prediction Equations: A Comparative Study Using Weather Data of Quetta, Pakistan

Summiya Malik[1,2,*], Yasmin Zahra Jafri[1], Azhar Marri[1], Shabana Yasmeen[3], Zahra Khanum[1], Hassan Jafri[4]

[1]Department of Statistics, University of Balochistan, Quetta, Pakistan
[2]Agriculture Research Institute, Quetta, Pakistan
[3]Agriculture Research Institute, Mastung, Pakistan
[4]Solar Energy Solutions, Karachi, Pakistan

Abstract We construct stochastic time series models like Auto Regressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA) and Auto Regressive Moving average (ARMA) to analyze and forecast weather data. The weather parameters are maximum, minimum temperatures and wind speed of five years from January 2012 to December 2016 of Quetta, Pakistan. Daily variations has been taken to forecast data. ARIMA models are used to forecast and predict the equations for monthly data while the SARIMA models were used on seasonal data and it provides better results for short run forecasting. Weibull Distribution (WD) shows better results on wind data as compare to ARIMA. An Artificial Neural Network (ANN) models for prediction of weather parameters are studied and results are found better as compared to the classical statistical method. The experimental results show that ANN gives better predictive values then traditional stochastic modeling techniques due to their ability to deal with non-linear stochastic data.

Keywords Forecasting, Artificial Neural Networks, Time Series Box-Jenkins Model, Weibull Distribution

1. Introduction

Weather forecasting is mostly used to solve the system of complex equations. Weather forecasting is the application of technology and science to foresee the condition of atmosphere for a given position. Several of the live systems depend upon weather situations to make necessary modifications in their systems. Forecasting aid to take essential measures to prevent destruction to life and property to a great extent. Satellites, sensors and ground stations that are located surrounding of our planet, which have data of great quantity which send information on a daily basis and used the weather situation to foresee in the next day [1]. The weather forecasting is live forecasting where out-turn of the model may be required for everyday weather guide or weekly or monthly weather plans.

Time series analysis is widely used in weather data. Time series modeling is the procedure of forecasting using historical records. Time series analysis have been widely used in huge number of practical problems comprising modeling and forecasting economic time series and process. To estimate the time series data, we have used data mining techniques and also used patterns in parameters of weather data such as; wind direction, relative humidity and temperature. However, we only used temperature (low, high) and wind speed of weather forecasts. We need to keep them together for absolute forecast. All the data is collected from the meteorological department, Quetta city. Parameters of weather are as follows [2]:

1) **Temperature:** is an analysis of coldness and hotness in the air and is recorded in the Celsius (C^0).
2) **Wind speed:** is related to convert in pressure of air at 1200 UTC (knots)".

Sequential order is taken to measure the time series data within a certain time [3]. We have used Auto Regressive Integrated Moving Average (ARIMA) model in our study because the characteristics of streaming stationarity (Calculations do not depend on lag, mean, variance and covariance) [4]. The model is also called Box-Jenkins model, which is developed in 1976 [3].

Seasonal ARIMA (SARIMA) model is the extension of ARIMA model that is relevant of seasonal time series. Construction of SARIMA is considered as series of seasonal

* Corresponding author:
summiya.malik@outlook.com (Summiya Malik)
Published online at http://journal.sapub.org/statistics

model [5]. There is relationship between ARIMA and SARIMA, in which the daily variation can be considered [6]. Weibull Distribution (WD) has been considered for expressing the wind speed variation. It represents various distribution characteristics when its parameter shape and scale are appropriately tuned. Therefore the Weibull model can be applied for modelling wind speed changes and forecasting future wind speed. Artificial Neural Network (ANN) is used in our paper for forecasting and analyzing the tendency and prediction of the data. ANN has the non-linear relationship between prediction and predictor, also able to general uncertainty of the function. That is similar to forecast the actual weather data and can be predicted large number of data. In this research Multi-Layer Perceptron (MLP) architecture is used. Back Propagation Neural Network (BPNN) is proposed to forecast the values of the parameter [7].

We studied the data to analyse the best model with least Akaike Information Criteria (AIC) value of ARIMA and the least Root Mean Square Error (RMSE) of the ANN to evaluate temperature and wind in the last five years. This research has benefits to select better methods amongst all, and develop a method for prediction to overcome the quantitative modeling forecasting. Data processing is used which is done by some software such as Minitab, MS excel, R software and NeuroSolutions software [8]. In this research, we used R software NeuroSolutions, MS excel and that has also facility to solve the problem of ARIMA. The objective of this study to forecast the future values of the parameters [9]. In this paper, the execution of ANN and ARIMA models is considered to decipher an instance of parameter including temperature and wind [8]. This paper is organized such as: Section 2 shows literature review and section 3 describes methodology. Computational results are shows in section 4. Future work and conclusion of this work are described in section 5.

2. Literature Review

Purnomo et al. emphasize models using ANN month's rainfall data [7]. Murat. et al. used Box-Jenkins ARIMA, SARIMA and regression models on weather data to produce sensible forecast. Their experimental results show that monthly mean temperature surface of India vary [10]. Narvekar and Fargose used different techniques of ANN in which BPP Algorithm performs best prediction with minimal error [11]. Adebiyi et al. made comparison of ARIMA and ANN model on stock data. The empirical results obtained by the authors reveal the superiority of ANN models over ARIMA model [8]. Basheer et al. also compared ANN and classical Box-Jenkins models on Consumer price data of Yemen. The experimental results show ANN better predictive values than natural stochastic models [12].

Murat M., Krzyszczak J et al. found the forecasting of metrological materials which depicts the courses of future on the basis of previous time series model, and is beneficial for agro physical models [13,14]. El-Mallah E.S et al. interpreted the performance of quadratic ARIMA and linear ARIMA which make annual short-term forecasting of temperature [15]. Balyani Y et al. studied temperature of air-surface of monthly basis mean and used SARIMA model to forecast [16]. Hu proposed ADALINE system to forecast weather which is newly proposed applications of ANN in forecasting of weather [17].

Kaur, A et al used ANN to forecast temperature on an hourly basis, 24 hours ahead wind speed and relative humidity. The data is divided into four seasons. The experimental results show that Radial Basis Function Network (RBFN) is used to compare with MLP, Hopfield Network Model (HFN) and Elman Recurrent Neural Network (ERNN) [18]. Ch.Jyosthna Devi, developed an algorithm to forecast the temperature. The BPNN is used to fairly approximate the function of large class. Authors used a model which have real time data set with maximum and minimum normalization scale of the data between (0 to 1) then trained and tested the data by using BPNN. The experimental results are compared to validate and check model accuracy and least error [19]. Kamal, L., & Jafri, Y. Z used simulations of stochastic and forecasting models of hourly average wind speed. They found ARMA (p, q) which is appropriate for probability forecast and prediction intervals [20]. Mehrdad. et al used artificial neural network and stochastic models to forecast the monthly flow discharge of the Ghara-Aghaj River the results revealed that MLP and RNN had superior performance than ARIMA [21]. Kumar Abhishek et al. studies the applicability of ANN by constructing non-linear models for weather forecast, the researcher also compared the performance of established models using different transfer function, neurons and hidden layer to predict maximum temperature for 365 days [22]. Anosh Graham et al. used seasonal ARIMA model to forecast future rainfall and found that ARIMA model yield comparatively better forecast than the simple models [23]. Athraa Kadhem et al. used the wind speed data and apply WD model to predict the next day forecasting and the proposed model utilizes the ANN to predict the wind speed data. The results indicate that the proposed ANN model is capable of depicting the fluctuating wind speed during different seasons of the year at different locations [24].

3. Methodology

Methodology adopted for conducting this study includes stochastic models and ANN with R software and Neuro solution Version 7.

3.1. Statistical Models

Box-Jenkins is considered in this research followed by [25]. This model is dependent on different steps:

- Appropriate model is identified from the ARIMA model family.
- Model of estimation.
- To verify the model we check the suitability of under-study-series, when that is not suitable then we return back to the first step, if not then forward to the next step.
- Model that is selected is used for prediction.

3.1.1. ARIMA

These are models used for forecasting a stationary time series by one and more times differencing. An ARIMA model generally denoted as ARIMA (p, d and q) where p is the order of auto-regressive, d is the degree of differencing and q is the number of lagged forecast.

In term of X the general forecasting equation is

$$X_t = \mu + \theta_1 X_{t-1} + \cdots + \theta_p X_{t-p} + \varphi\, e_{t-1} + \cdots \varphi\, e_{t-q}$$

Here auto regressive parameter (θ) and moving average parameters (φ) are defined.

3.1.2. SARIMA

ARIMA is the extension form of these models and is used for seasonal time series. It is used for forecasting of time series with univariate data having seasonality and trends. The SARIMA (p, d, q) (P, D, Q) m process as SARIMA (p, d, q) (P, D, Q) m is given by

$$\theta(B^m)\theta(B)(1-B^m)^D(1-B)^d Y_t = c + \varphi(B^m)\,\varphi(B)e_t$$

In SARIMA we analyze the long term trend and seasonal effect. SARIMA is based on ARIMA models to change time series data.

3.2. Artificial Neural Network

Neural Network (NN) is important domain of Artificial Intelligence (AI) that is used in complex and modern applications of modern science, which is as follows: industry of robotic systems, systems of decision support, automated control systems, prediction and identification systems.

ANN is a tool of effective forecasting [26] and consist of algorithm, which mimic the feature of human being brain that explore and generate basic knowledge by research [27]. ANN contains different components, which need to be carefully calculated because it effects the performance of forecasting method. ANN define some different elements such as Machine Learning Algorithm and architecture structure. The architecture is defined by different number of layers, number of neurons in all the layers and the rules that determine the architecture. Feed-Forward Back Propagation Neural Network (FFBPNN) is one of the type of neural network which is mostly used for forecasting. FFBPNN algorithm is mostly used as learning algorithm which updates the weights dependence on the variation in to the output value of the NN and desired real value. Figure 1, Shows ANN model the given input layer shows the action which is fed in to next layer until the output layer.

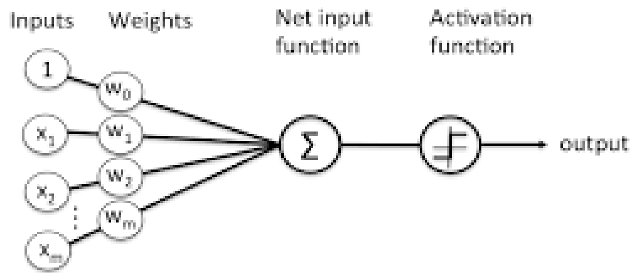


Figure 1. Single layer Neural Network

3.3. Weibull Distribution

We calculated the Weibull probability density functions. There are serval methods [28] for determining the Weibull parameters c and k using Justus relations for c and k, i.e. the scale parameter and the shape parameter we obtained the following.

$$k_{th} = \left(\frac{\sigma}{\bar{\mu}}\right)^{-1.086}$$

And

$$c = \frac{\bar{\mu}}{\left|(1+1/k)\right.}$$

4. Computational Results

The proposed method is provided in this section by results that are experimented. First of all we described data description that are used in the experimental results. The results are presented by using the Box-Jenkins method of statistics and then ANN results are compared and analyzed. The error measurements are forecasted by the use of estimating the method for forecasting. These error measurements that are used such as: RMSE for ANN and AIC value for ARIMA and SARIMA models.

4.1. Data Description

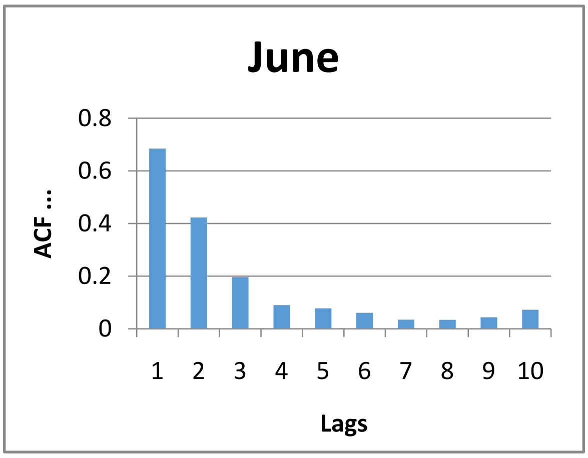
For conducting this study we obtained daily recorded data of five years period (2012-16) from Meteorological Department, Quetta. The obtained record includes daily maximum, minimum temperature and wind speed observations.

4.2. Prediction Using Statistical Model (Box-Jenkins)

The R software is used as statistical program, which identify the suitable model for data by using ARIMA model for AIC value, Auto-Correlation Function (ACF) and Partial Autocorrelation (PACF). Table 1, 2 and 3 show all results of Temperature Maximum, Minimum and wind speed respectively.

Table 1. ARIMA modelling for maximum temperature and identification of the model

Months	AIC	ACF	PACF	Suggested Model	θ	φ
January	782.78	At lag 10 gradually dice away.	At lag 2 different Significantly from Zero.	(1,1,2)	0.6304 -----	-0.7651 -0.1978
February	749.91	At lag 7 gradually dies away.	At lag 2 different Significantly from Zero.	(0,1,0)	------ ------	------ ------
March	819.91	At lag 4 gradually dies away.	At lag 2 different Significantly from Zero.	(1,0,1)	0.4922 -----	0.331 -----
April	714.82	At lag 10 gradually dies away.	At lag 2 different Significantly from Zero.	(0,1,0)	------ ------	------ ------
May	704.36	At lag 5 gradually dies away	At lag 2 different Significantly from Zero	(0,1,0)	------ ------	------ ------
June	570.2	At lag 7 gradually dies away	At lag 2 different Significantly from Zero	(2,1,2)	1.5615 -0.6754	-1.8488 0.8739
July	578.24	At lag 4 gradually dies away	At lag 2 different Significantly from Zero	(1,0,0)	0.6260 -----	----- -----
August	572.4	At lag 5 gradually dies away	At lag 2 different Significantly from Zero	(2,0,0)	0.6317 -0.1063	------ ------
September	572.69	At lag 5 gradually dies away	At lag 2 different Significantly from Zero	(1,1,2)	0.2527 ------	-0.3059 -0.5433
October	669.52	At lag 6 gradually dies away	At lag 3 different Significantly from Zero	(1,0,0)	0.8193 ------	----- -----
November	749.72	At lag 3 gradually dies away	At lag 2 different Significantly from Zero	(1,0,0)	0.5615 -----	----- -----
December	800.62	At lag 7 gradually dies away	At lag 2 different Significantly from Zero	(0,1,0)	------ ------	------ ------



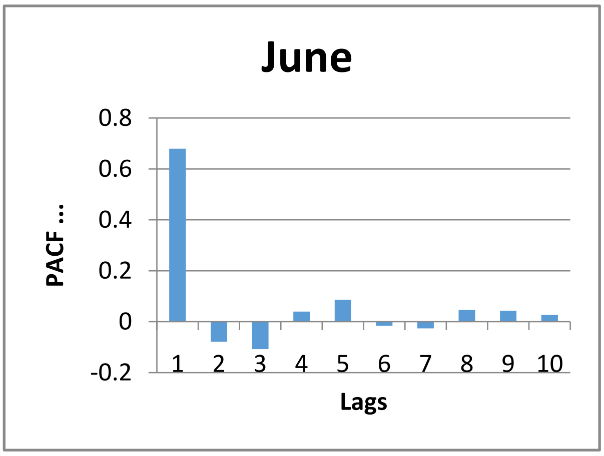


Figure 2. (a) Shows autocorrelation function verses lag by ARIMA model. (b) Shows partial-autocorrelation function verses lag by ARIMA model

4.2.1. Prediction Equations of Maximum Temperature

ARIMA (0, 1, 0) is the model that is used for the months of February, April, May and December. ARIMA (1, 0, 0) for July, October and November. ARIMA (1, 1, 2) for January and September. ARIMA (2, 1, 2) for June. ARIMA (2, 0, 0) for August. And ARIMA (1, 0, 1) for March.

• The Non-Seasonal ARIMA equations are as follows

ARIMA (0, 1, 0)

$$X_t = a + X_{t-1}$$

ARIMA (1, 0, 0)

$$X_t = a + \theta(X_{t-1} - X_{t-2})$$

ARIMA (1, 1, 2)

$$X_t = a + X_{t-1} + \theta(X_{t-1} - X_{t-2}) - \varphi_1 e_{t-1} - \varphi_2 e_{t-2}$$

ARIMA (2, 1, 2)

$$X_t = a + X_{t-1} + \theta_1(X_{t-1} - X_{t-2}) + \theta_2(X_{t-1} - X_{t-2}) - \varphi_1 e_{t-1} - \varphi_2 e_{t-2}$$

ARIMA (2, 0, 0)

$$X_t = a + \theta_1(X_{t-1} - X_{t-2}) + \theta_2(X_{t-1} - X_{t-2})$$

ARIMA (1, 0, 1)

$$X_t = a + \theta(X_{t-1} - X_{t-2}) - \varphi\, e_{t-1}$$

Respectively where "a" is constant, e is the error at period (t-1), = θ AR and = φ MA.

Table 2. ARIMA modelling for minimum temperature and identification of the model

Month	AIC	ACF	PACF	Suggested Model	θ	φ
January	828.96	At lag 6 gradually dies	At lag 4 different Significantly from Zero	(1,1,1)	0.5367 ----- -----	-0.9577 ----- -----
February	764.86	At lag 5 gradually dies	At lag 2 different Significantly from Zero	(1,1,1)	0.6444 ----- -----	-0.9757 ----- -----
March	832.36	At lag 4 gradually dies	At lag 3 different Significantly from Zero	(0,1,3)	----- ----- -----	-0.4411 -0.2074 -0.2326
April	707.04	At lag 4 gradually dies	At lag 3 different Significantly from Zero	(0,1,3)	------ ------ ------	-0.5249 -0.0653 -0.2856
May	707.87	At lag 6 gradually dies	At lag 5 different Significantly from Zero	(1,1,1)	0.5329 ------ ------	-0.9414 ----- -----
June	730.3	At lag 3 gradually dies	At lag 7 different Significantly from Zero	(0,1,1)	------ ------ ------	-0.6630 ------ ------
July	643.05	At lag 4 gradually dies	At lag 6 different Significantly from Zero	(1,1,2)	0.6839 ----- -----	-1.2457 0.2930 -----
August	715.1	At lag 5 gradually dies	At lag 5 different Significantly from Zero	(1,0,1)	0.6549 ----- -----	-0.2780 ------ ------
September	683.81	At lag 3 gradually dies	At lag 2 different Significantly from Zero	(2,0,0)	0.8078 -0.2123 ------	------ ------ ------
October	759.77	At lag 7 gradually dies	At lag 2 different Significantly from Zero	(0,1,0)	------ ------ ------	------ ------ ------
November	729.78	At lag 5 gradually dies	At lag 2 different Significantly from Zero	(1,0,0)	0.6963 ----- -----	------ ------
December	818.55	At lag 6 gradually dies	At lag 4 different Significantly from Zero	(1,1,2)	0.6910 ----- -----	-1.3263 0.3410 -----

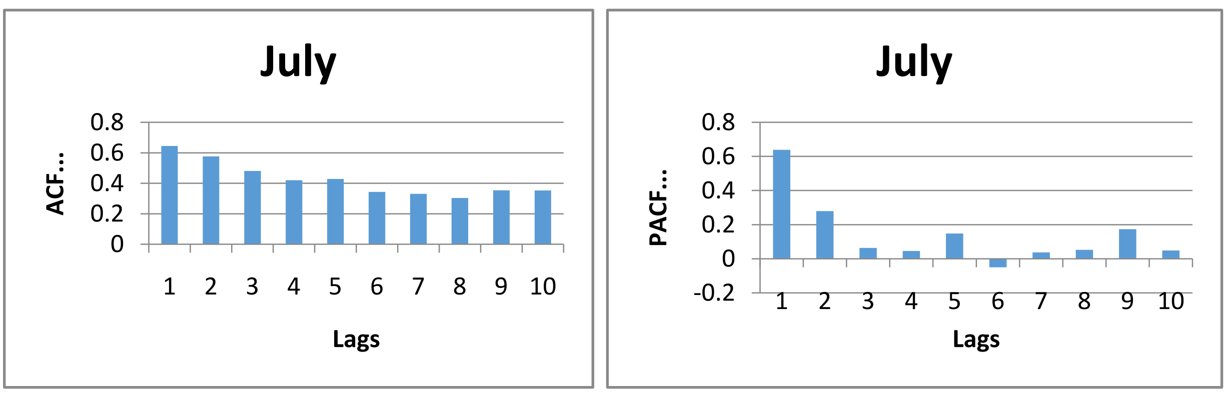


Figure 3. (a) Shows autocorrelation function verses lag by ARIMA model. (b) Shows partial-autocorrelation function verses lag by ARIMA model

4.2.2. Prediction Equations of Minimum Temperature

ARIMA (1, 1, 1) model is taken for the months of January, February and May. ARIMA (0, 1, 3) March, April. ARIMA (1, 1, 2) for July, December. ARIMA (0, 1, 1) for June. ARIMA (1, 0, 1) for August. ARIMA (2, 0, 0) for September. ARIMA (0, 1, 0) for October. And ARIMA (1, 0, 0) for November.

- The Non-Seasonal ARIMA equations are as follows

ARIMA (1, 1, 1)

$X_t = a + \theta(X_{t-1} - X_{t-2}) - \varphi\, e_{t-1}$

ARIMA (0, 1, 3)

$X_t = a + X_{t-1} - \varphi_1\, e_{t-1} - \varphi_2\, e_{t-2} - \varphi_3\, e_{t-3}$

ARIMA (1, 1, 2)

$X_t = a + \theta(X_{t-1} - X_{t-2}) - \varphi_1\, e_{t-1} - \varphi_2\, e_{t-2}$

ARIMA (0, 1, 1)

$X_t = a + X_{t-1} - \varphi_1\, e_{t-1}$

ARIMA (1, 0, 1)

$X_t = a + \theta(X_{t-1} - X_{t-2}) - \varphi\, e_{t-1}$

ARIMA (2, 0, 0)

$X_t = a + \theta_1(X_{t-1} - X_{t-2}) + \theta_2(X_{t-1} - X_{t-2})$

ARIMA (0, 1, 0)

$X_t = a + X_{t-1}$

ARIMA (1, 0, 0)

$X_t = a + \theta(X_{t-1} - X_{t-2})$

Respectively where "α" is constant, e is the error at period (t-1), θ =AR and φ =MA

Table 3. ARIMA modelling for wind speed and identification of the model

Month	AIC	ACF	PACF	Suggested Model	θ	φ
January	614.82	At lag 2 gradually dies	At lag 3 different Significantly from zero	(1,0,0)	0.2155 -----	----- -----
February	562.52	At lag 3 gradually dies	At lag 3 different significantly from Zero	(1,0,0)	0.2519 -----	------ ------
March	610.31	At lag 3 gradually dies	At lag 6 different significantly from Zero	(0,1,1)	------ ------	0.9125 -----
April	578.94	At lag 2 gradually dies	At lag 3 different Significantly from Zero	(2,0,2)	0.0006 0.4781	0.3297 0.7358
May	655.12	At lag 3 gradually dies	At lag 2 different Significantly from Zero	(0,0,1)	------ ------	0.2679 -----
June	520.26	up from lag 1	At lag 4 different Significantly from Zero	(0,0,0)	------ ------	------ ------
July	535.11	At lag 2 gradually dies	At lag 2 different Significantly from Zero	(0,0,1)	------ ------	0.1218 ------
August	624.49	At lag gradually dies	At lag 2 different Significantly from Zero	(0,0,0)	------ ------	------ ------
September	476.05	Up from lag 1	At lag 9 different Significantly from Zero	(0,1,1)	------ ------	-0.8558 ------
October	530.59	At lag 5 gradually dies	At lag 5 different Significantly from Zero	(1,0,1)	0.8989 ------	-0.7407 ------
November	453.74	Dies gradually up to lag 4 At lag 4 gradually dies	At lag 4 different Significantly from Zero	(2,1,2)	-0.9863 -0.1424	0.1994 -0.7551
December	890.11	At lag 2 gradually dies	At lag 2 different Significantly from Zero	(0,0,1)	----- -----	0.2664 -----

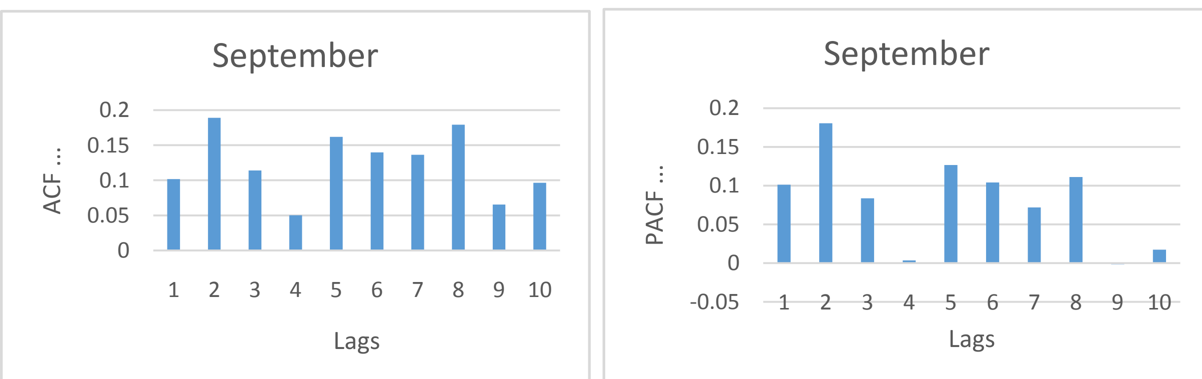


Figure 4. (a) Shows autocorrelation function verses lag by ARIMA model. (b) Shows partial-autocorrelation function verses lag by ARIMA model

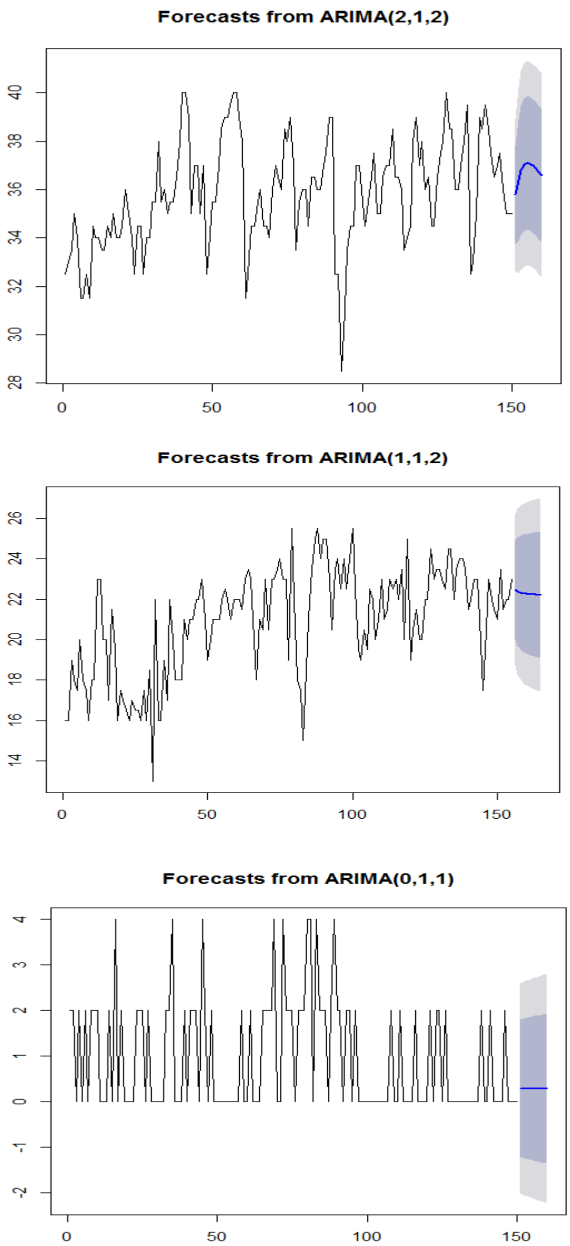


Figure 5. Prediction using Statistical Models (Box-Jenkins)

4.2.3. Prediction Equations of Wind Speed

ARIMA (1, 0, 0) for January and February. ARIMA (0, 1, 1) for March and September. ARIMA (0, 1, 1) May, July, December. ARIMA (2, 0, 2) for April. ARIMA (0, 0, 0) for June and August. ARIMA (1, 0, 1) for October. ARIMA (2, 1, 2) for November.

- The Non-Seasonal ARIMA equations are as follows

ARIMA (1, 0, 0)

$$X_t = a + \theta(X_{t-1} - X_{t-2})$$

ARIMA (0, 1, 1)

$$X_t = a + X_{t-1} - \varphi\, e_{t-1}$$

ARIMA (0, 0, 1)

$$X_t = a - \varphi\, e_{t-1}$$

ARIMA (2, 0, 2)

ARIMA (2, 0, 2)

$$X_t = a + \theta_1(X_{t-1} - X_{t-2}) + \theta_2(X_{t-1} - X_{t-2})$$
$$-\varphi_1\, e_{t-1} - \varphi_2\, e_{t-2}$$

ARIMA (1, 0, 1)

$$X_t = a + \theta_1(X_{t-1} - X_{t-2}) - \varphi\, e_{t-1}$$

ARIMA (2, 1, 2)

$$X_t = a + X_{t-1} + \theta_1(X_{t-1} - X_{t-2}) + \theta_2(X_{t-1} - X_{t-2})$$
$$-\varphi_1\, e_{t-1} - \varphi_2\, e_{t-2}$$

Respectively Where "a" is constant, e is the error at period (t-1), θ =AR and φ =MA.

The model that is identified for the data is ARIMA (2,1,2), (1,1,2) and (0,1,1). These results are achieved in approximated parameters test of significant and also achieved analysis of residential test (in other words, it is achieved for this model in test of diagnostic). Table 4. Represents the best results among the all parameters with least AIC value.

Table 4. Best Models of ARIMA that shows the suitability of time series data

Month	ARIMA	AIC	RMSE
Temperature Maximum (June)	(2,1,2)	570.2	1.92
Temperature Minimum (July)	(1,1,2)	643.05	1.89
Wind Speed (September)	(0,1,1)	476.05	1.45

Figure 5 the graphs of the forecasting on under study data by using Box-Jenkins model in which the recorded data is shows in black curve and forecasted data is shows in blue curve.

Table 5 to 7 show the predicted values of maximum temperature, minimum temperature and wind speed having 95% C.I of Lower and Upper limits. While Tables 8 to 10 defines AIC, RMSE values and seasonal SARIMA classification on the basis of all four seasons of Quetta from spring to winter having their corresponding months for all the three selected parameters.

Table 5. Shows the Forecasting Temperature Maximum values of ARIMA Model (June)

Point Forecast	Low 95%	High 95%
35.78233	32.64849	38.91616
36.38039	32.53219	40.22858
36.78586	32.67387	40.89785
37.01507	32.81960	41.21055
37.09913	32.88461	41.31364
37.07557	32.85871	41.29243
36.98201	32.76506	41.19897
36.85184	32.63480	41.06888
36.71176	32.49349	40.93003
36.58095	32.35700	40.80490

Table 6. Shows the Forecasting Temperature Minimum values of ARIMA Model (July)

Point Forecast	Low 95%	High 95%
22.46761	18.71429	26.22092
22.38839	18.29057	26.48621
22.33422	18.03443	26.63401
22.29718	17.86668	26.72767
22.27184	17.74922	26.79445
22.25451	17.66192	26.84710
22.24266	17.59339	26.89194
22.23456	17.53688	26.93223
22.22902	17.48828	26.96976
22.22523	17.44494	27.00551

Table 7. Shows the Forecasting Wind Speed values of ARIMA Model (September)

Point Forecast	Low 95%	High 95%
1.267275	-1.363356	3.897905
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411
1.148333	-1.501745	3.798411

Table 8. SARIMA modelling for maximum temperature and identification of the model

Seasons	AIC	RMSE	Model
Spring	2247.62	2.763	$(1,1,2)(1,0,0)_4$
Summer	1725.7	1.564	$(1,1,2)(1,0,0)_4$
Autumn	2103.31	2.411	$(2,1,2)(1,0,1)_4$
Winter	2302.29	3.082	$(1,0,1)(0,1,1)_4$

Table 9. SARIMA modelling for minimum temperature and identification of the model

Seasonal	AIC	RMSE	Model
Spring	2276.13	2.849	$(2,1,2)(1,0,0)_4$
Summer	2125.61	2.419	$(1,1,2)(0,0,1)_4$
Autumn	2169.80	2.605	$(1,1,2)(1,0,0)_4$
Winter	2424.11	3.504	$(1,1,2)(0,0,1)_4$

Table 10. SARIMA modelling for wind speed and identification of the model

Seasonal	AIC	RMSE	Model
Spring	1862.22	1.816	$(1,0,0)(1,0,0)_4$
Summer	1696.21	1.515	$(1,1,1)(1,0,0)_4$
Autumn	1457.92	1.187	$(1,0,1)(1,0,0)_4$
Winter	2228.93	2.823	$(1,0,0)(1,0,0)_4$

Table 11. ANN results of Maximum Temperature

Temp.Max		January	February	March	April	May	June	July	August	September	October	November	December
RMSE													
Neurons	20	4.7717	6.5091	4.485	5.189	2.60668	1.817	2.0344	1.945	2.456	2.2495	6.483	3.656
	50	3.7224	6.85	3.9938	4.622	3.1705	2.3846	2.1345	2.088	2.7205	2.4600	4.596	5.582
	80	3.275	7.835	5.636	5.934	2.7137	2.3600	2.0949	2.0855	2.4238	9.0924	3.128	5.2922

Table 12. ANN Results for Minimum Temperature

Temp.Min		January	February	March	April	May	June	July	August	September	October	November	December
RMSE													
Neurons	20	4.488	5.734	4.4715	4.5534	3.3733	6.128	1.867	3.1401	2.8162	5.545	2.7164	3.7628
	50	5.6711	6.3903	4.106	4.7906	3.7086	6.0561	2.888	3.7971	9.1248	3.9411	6.2713	4.2755
	80	5.239	7.206	6.247	5.3137	4.1524	3.870	2.3049	2.4986	2.5657	4.0458	4.7172	4.9687

Table 13. ANN Results for Wind Speed

Wind speed		January	February	March	April	May	June	July	August	September	October	November	December
RMSE													
Neurons	20	1.959	2.2125	1.909	1.588	2.2152	1.6657	1.5012	1.8707	1.207	1.9164	1.6932	1.5490
	50	1.9524	3.7837	1.9793	1.6635	2.4990	2.7285	1.892	3.4408	3.2376	2.4635	1.2444	1.6138
	80	2.2766	2.756	2.0406	2.2835	3.9856	1.6644	1.9048	3.1556	3.1974	3.1032	2.016	1.7268

Table 14. Results for the proposed ANN method

Month	RMSE	ANN
Temperature Maximum (June)	1.817	[20,1,1]
Temperature Minimum (July)	1.867	[20,1,1]
Wind Speed (September)	1.207	[20,1,1]

4.3. Forecasting Using ANN Model

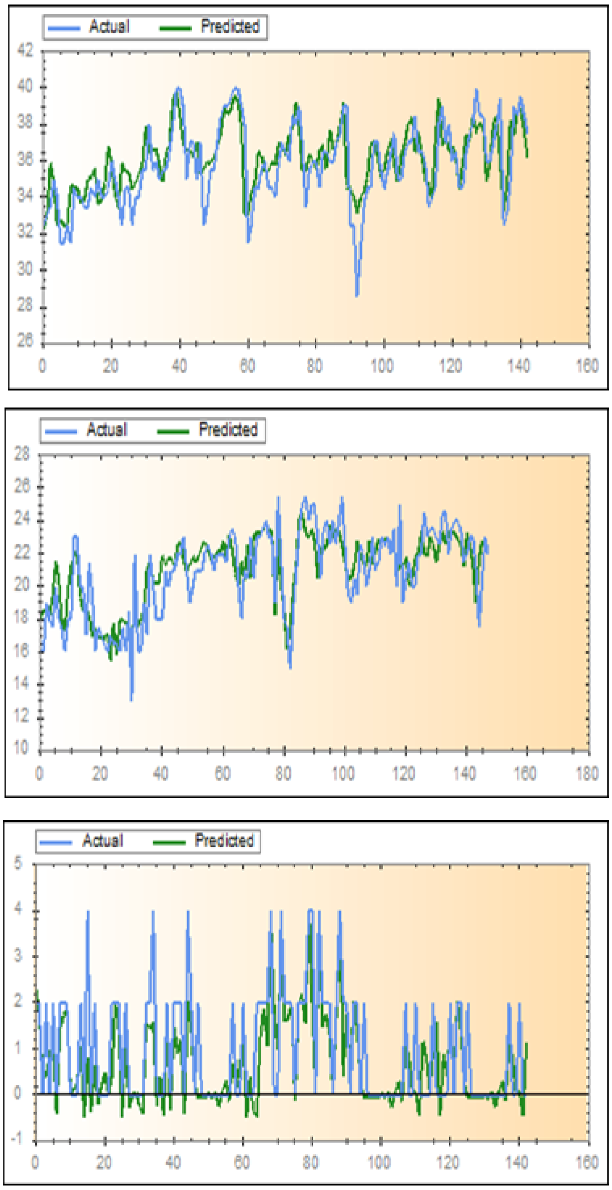


Figure 6. (a) Shows the neural network prediction of temperature maximum data and one hidden layer. (b) Shows the neural network prediction modeling of temperature minimum data and one hidden layer. (c) Shows the neural network prediction modeling of wind speed data and one hidden layer

Table 15. Results of WD for the month of September

Model	RMSE
ARIMA	1.450
ANN	1.207
WD	1.288

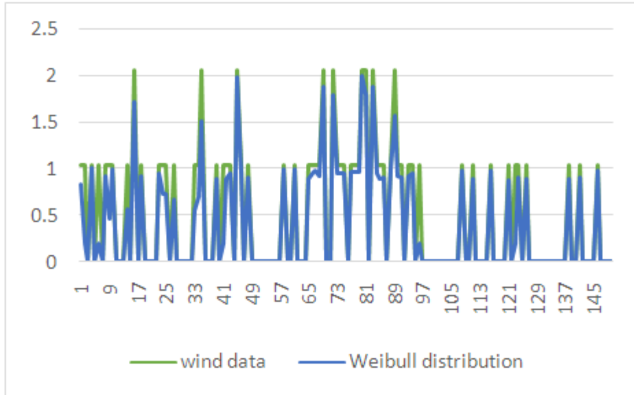


Figure 7. Predicted and actual values for the month of September

The accomplishment in execution of Neural Networks relies upon the comprehension and suitable decision variable of input. If there should arise an occurrence of accomplishing an anticipating with respect to the time series, the Neural Network will have one yield provided the determined esteem and the data sources might be spoken to through estimations of the factors investigated at various past minutes. Table 11 to 13 shows the values of ANN in the way that system presents the input layer as 20, 50 and 80 neurons having single layer of output with Training =45%, cross validation =15%, Testing =40%, Layer =1, Epoch =150 and bold values represent best results. The outcomes of RMSE utilizing this technique are in table 14 and actual and predicted values are shown in Figure 6.

4.4. Comparative Study

The Box-Jenkins model used for time series prediction on weather data assumes that there is a linear relationship between input and output. NN approximate the non-linear functions and have been successfully used for that future forecasting. Table 14 shows the proposed NN models, for June, July and September [20,1,1] perform better than classical Box-Jenkins models (2, 1, 2), (1, 1, 2) and (0, 1, 1) for the months of June, July and September. This can be shown in table 4. Using forecast error measurement like RMSE. We also used WD on wind data from January 2012 to December 2016. The table 15 shows the comparative study of ARIMA, ANN and WD. Figure 7 shows the predicted and actual values for the month of September. The comparison confirms the superiority of the proposed ANN to ARIMA and WD.

5. Conclusions

In this paper two methods for model identification and forecasting are used, one is based on stochastic models like ARIMA and SARIMA for short and long term variations on weather data, respectively and the other proposed model using ANN. Comparison between the statistical model and proposed ANN model showed that the proposed model gave lower error and higher accuracy of time series data. We also apply Weibull distribution on wind data. Table 15 shows the comparison of ARIMA, NN and WD. Comparison is made on the basis of RMSE. NN is best amongst three.

REFRENCES

[1] Tektaş, M. Weather forecasting using ANFIS and ARIMA models. Environmental Research, Engineering and Management, (2010). 51(1), 5-10.

[2] L. Sunitha, M. Balraju, J. Sasikiran, B. Anil Kumar. Finding relation between parameters of weather data using linear regression method. International Journal of Research in engineering and Technology, (2016). 2319-1163.

[3] Lütkepohl, H. New introduction to multiple time series analysis. Springer Science & Business Media. (2005).

[4] Gujarati, D. N. Basic econometrics. Tata McGraw-Hill Education. (2009).

[5] Afrifa-Yamoah, E. Application of ARIMA models in forecasting monthly average surface temperature of Brong Ahafo Region of Ghana. International Journal of Statistics and Applications, (2015). 5(5), 237-246.

[6] Sami, M., Waseem, A., Jafri, Y. Z., Shah, S. H., Khan, M. A., Akbar, S., ... & Murtaza, G. Prediction of the rate of dust fall in Quetta city, Pakistan using seasonal ARIMA (SARIMA) modeling. International Journal of Physical Sciences, (2012). 7(10), 1713-1725.

[7] Purnomo, H. D., Hartomo, K. D., & Prasetyo, S. Y. J. Artificial neural network for monthly rainfall rate prediction. In IOP Conference Series: Materials Science and Engineering, (2017). (180)1-012057.

[8] Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. Comparison of ARIMA and artificial neural networks models for stock price prediction. Journal of Applied Mathematics, (2014).

[9] Sidiq, M. Forecasting Rainfall with Time Series Model. IOP Conference Series: Materials Science and Engineering. (2018). 407-012154.

[10] Murat, M., Malinowska, I., Gos, M., & Krzyszczak, J. Forecasting daily meteorological time series using ARIMA and regression models. International agro physics, (2018). 32(2), 253-264.

[11] Narvekar, M., Fargose, P., “Daily weather forecasting using artificial neural network”, International Journal of Computer Applications, (2015). 121, 22, 9-13.

[12] Basheer M. Al-Maqaleh, Abduhakeem A. Al-Mansoub, Fuad N. Al-Badani. Forecasting using Artificial Neural Network and Statistics Models. I.J. Education and Management Engineering, (2016). 3, 20-32.

[13] Murat M., Malinowska I., Hoffmann H., and Baranowski P. Statistical modeling of agro meteorological time series by exponential smoothing. Int. Agrophys, (2016). 30(1), 57-66.

[14] Krzyszczak J., Baranowski P., Hoffmann H., Analysis of Climate Dynamics Across a European Transect Using a Multifractal Method, In: Advances in Time Series Analysis and Forecasting. Springer Int. Publishing, Cham. (2017). 978-3-319-55789-2_8.

[15] El-Mallah E.S. and Elsharkawy S.G. Time-series modeling and short term prediction of annual temperature trend on Coast Libya using the box-Jenkins ARIMA Model. Advances Res, (2016). 6(5), 1-11.

[16] Balyani Y., Niya G.F., and Bayaat A. A study and prediction of annual temperature in Shiraz using ARIMA model. J. Geographic Space, (2014). 12(38), 127-144.

[17] Hu, M.J.C., “Application of ADALINE system to weather forecasting”, Technical report, Stanford Electron, 1964.

[18] Kaur, A., Sharma, J.K., Agrawal, S., “Artificial neural network in forecasting maximum and minimum relative humidity”, International journal of Computer Science and Network Security, (2011). 11, 5, 197-199.

[19] Ch.Jyosthna Devi, B.Syam Prasad Reddy, K.Vagdhan Kumar, B.Musala Reddy, N.RajaNayak, “ANN Approach for Weather Prediction using Back Propagation,” International Journal of Engineering Trends and Technology- Volume 3 Issue 1- (2012).

[20] Kamal, L., & Jafri, Y. Z. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. Solar Energy, (1997). 61(1), 23-32.

[21] Fereydooni, M., Rahnemaei, M., Babazadeh, H., Sedghi, H., & Elhami, M. R. Comparison of artificial neural networks and stochastic models in river discharge forecasting, (Case study: Ghara-Aghaj River, Fars Province, Iran). African Journal of Agricultural Research, (2012). 7(40), 5446-5458.

[22] Abhishek, K., Singh, M. P., Ghosh, S., & Anand, A. Weather forecasting model using artificial neural network. Procedia Technology, (2012). 4, 311-318.

[23] Graham, A., Sahu, J. K., Sahu, Y. K., & Yadu, A. Forecast future rainfall & temperature for the study area using seasonal auto-regressive integrated moving averages (SARIMA) model. IJCS, (2019). 7(1), 894-897.

[24] Athraa, A. K, Noor, I. A. W, Ishak, A. Jasronita, J, and Ahmed N. A. Advanced Wind Speed Prediction Model Based on a Combination of Weibull Distribution and an Artificial Neural Network. Energies, 2017, 10, 1744.

[25] G E P. BOX, G M. JENKINS, "Time series analysis for casting and control", book, (1976).

[26] Aladag, C.H., "A New Architecture Selection Method Based On Tabu Search for Artificial Neural Networks. Expert Systems with Application ", (2011). 38, 3287–3293.

[27] E. Egrioglu, C.H. Aladag, U. Yolcu, V.R. Uslu, and M.A. Basaran, "A new approach based on artificial neural networks for high order ultivariate fuzzy time series", Expert Systems with Applications, (2009). 36(7), 10589-10594.

[28] C. G. Justus, Winds and System Performance. Franklin Institute Press, Philadelphia (1978).