Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 16, 2023, page 57 – 66

Time Series Forecasting on Weather in Malaysia Using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN)

Chia Yi Hwee, Muhammad Fauzee Hamdan*
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
*Corresponding author: mfauzee@utm.my

Abstract
Accurate weather forecasting plays a crucial role in various domains, including agriculture, transportation, and disaster management. This paper presents a comparative study that focuses on temperature prediction using Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN). The study employs historical weather data, specifically a dataset comprising temperature records, to investigate the performance of these methods. The ARIMA model is utilized to capture the inherent time-series patterns present in the temperature data. By leveraging past observations, the ARIMA model predicts future temperature patterns. Subsequently, an ANN model is developed, employing a multi-layered neural network architecture to learn and predict temperature patterns. The study demonstrates that both the ARIMA and ANN methods can provide reasonably accurate temperature forecasts. However, the ARIMA model consistently outperforms the ANN model in terms of forecasting accuracy, as indicated by lower values of mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE).

Keywords: Weather forecasting; Autoregressive Integrated Moving Average; Artificial Neural Network; Multi-layered neural network

1. Introduction

Weather forecasting is the prediction of the weather through application of the principles of physics, supplemented by a variety of statistical and empirical techniques. In addition to predictions of atmospheric phenomena themselves, weather forecasting includes predictions of changes on Earth's surface caused by atmospheric conditions such as snow and ice cover, storm tides, and floods. Weather forecasting is the use of science and technology to predict the condition of the weather for a given area. It is one of the most difficult issues the world over. Based on the study done by other research, it is possible for us to estimate the weather by utilizing predictive analysis. Analysis of various data mining procedures is needed before applying.

Mathur et al. [1] stated that weather forecasting is one of the most important types of forecasting because agriculture sectors as well as many industries are largely dependent on the conditions of weather. The present technology uses numerical model-based weather forecasting which involves many calculations made very rapidly using supercomputers. There is a scope of improvement in the accuracy of forecasts made by these models when considering the benefits that it has. The idea of incorporating human intelligence together with numerical models can result in enormous improvement in accuracy of weather forecasting [2].

This study aims to forecast the daily temperature in Malaysia based on the temperature by using time series analysis. The weather forecasting can be done using statistical methods like autoregressive

(AR), moving average (MA), autoregressive moving average (ARMA), Autoregressive integrated moving average (ARIMA) and artificial neural network (ANN) model. The results are evaluated through four performance measures: Mean Absolute Error (MAE), Mean Square Error (MSE),
Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) to check the accuracy of forecasting performance.

2. Literature Review

Autoregressive Integrated Moving Average (ARIMA)

Darji et al. [3] states that the weather forecasting using data mining technique can be done by statistical method and numerical weather prediction models. The weather forecasting can be done using a statistical method like AR, MA, ARMA, ARIMA and multiple regression. Each method has its limitations. AR model regresses against the past value of the series. MA model uses past errors as an explanatory variable. The AR model is suitable only for linear correlated data and is not appropriate for nonlinear data. AR and MA can be combined to form a general and useful class of time series model known as the ARMA model; ARIMA gives better results than ARMA. The details of this methodology are explained in further sections. It is usually in statistical technique when the data is responsibly extended and the correlation between past observations is stable.

Over the years, many intelligent time series models have been developed in the literature to improve the accuracy and efficiency of time series forecasting. One of the most widely used and recognized statistical forecasting time series models is the Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model is well-known for notable forecasting accuracy and efficiency in representing various types of time series with simplicity as well as the associated Box-Jenkins methodology for optimal model construction [4]. The basic assumption made in implementing this model is to assume the time series is linear and follows a statistical distribution, such as the normal distribution [5]. For seasonal time series forecasting, Box and Jenkins [6] proposed a quite successful variation of the ARIMA model called the Seasonal Autoregressive Integrated Moving Average (SARIMA) model.

Artificial Neural Network (ANN)

Moreover, Hayati et al. [7] show that how ANN can be used for forecasting weather for the city of Iran. They utilize ANN for one day ahead prediction of weather parameter as temperature of city of Iran. Their study was based on the most common neural network model Multilayer Perceptron (MLP) which is trained and tested using ten years past metrological data. To improve the accuracy of prediction, they split data into four seasons' which are spring, summer, fall and winter and then for each season one network is presented. They showed that MLP network with this structure has minimum error between exact and predicted values at each day and has a good performance, reasonable prediction accuracy and minimum prediction error in general. The forecasting reliability was evaluated by computing the mean absolute error between the exact and predicted values. The result shows that this network can be an important tool for temperature forecasting.

Kadu et al. [8] proposed a model of temperature prediction which uses new wireless technology for data gathering with the combination of statistical software. They proposed approach of artificial neural network that uses analysis of data and learn from it for future predictions of temperature, with the combination of wireless technology and statistical software. Abhishek et al. [9] presented a survey that using ANN approach for weather forecasting yields good results and can be considered as an alternative to traditional metrological approaches. The study describes the capabilities of ANN in predicting several weather phenomenon's such as temperature, thunderstorms, rainfall and concluded that major architecture like BPN and MLP are suitable to predict weather phenomenon.

3. Methodology

Data Source

The historical data of the daily temperature from January 2000 until December 2021 in Malaysia are retrieved from the website of Visual Crossing Corporation. At first, the data from January 2000 to December 2020 is analysed by using the appropriate forecast method. The temperature of weather for 150 days onwards starting from January 2021 in Malaysia will be forecasted. Next, the comparison between original dataset and the forecasting values calculated by the method chosen will be investigated. The evaluation for forecasting performance measurement such as MAE, MSE, RMSE and MAPE will be evaluated. The method with higher accuracy which provide the low error measurement will be chosen as the optimal method for weather forecasting.

Box-Jenkins Method

In an autoregression model, we forecast the variable of interest using a linear combination of past values of the variable. The term autoregression indicates that it is a regression of the variable against itself. Thus, an autoregressive model of order p can be written as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$

where ε_t is white noise. This is like a multiple regression but with lagged values of y_t as predictors. We refer to this as an AR (p) or ARMA $(p, 0)$ model, an autoregressive model of order p.

Rather than using past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model.

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

where ε_t is white noise. We refer to this as an MA (q) or ARMA $(0, q)$ model, a moving average model of order q. Of course, we do not observe the values of ε_t, so it is not really a regression in the usual sense.

These two concepts can be combined; not surprisingly, the resulting model is then called autoregressive moving average model, ARMA (p, q).

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (3)$$

Obviously, this model nests the two individual ones, by setting $p = 0$ reduces it to the moving average model MA (q), whereas $q = 0$ blanks out the MA part and reduces into AR (p) model.

The Box-Jenkins method refers to a set of procedures for identifying, fitting, estimating, and checking ARIMA models with time series data. Forecast follow directly from the form of fitted model. The time series is stationary when the mean, variance and auto-covariance do not vary over time and when the series is fluctuation around the means by plotting the series. The stationarity can be observed by Autocorrelation Function (ACF) plot and the Augmented Dicker-Fuller (ADF) test is used to support the decision making after examining the pattern of ACF.

Table 1 Summary of the behaviour of ACF and PACF

Properties	$\boldsymbol{AR\ (p)}$	$\boldsymbol{MA\ (q)}$	$\boldsymbol{ARMA\ (p, q)}$
ACF	Decay	Cuts after the q lag	Decay
PACF	Cuts after the p lag	Decay	Decay

The Akaike Information Criterion (AIC) can be used to assess the best model selection for the Box-Jenkins model. The good model with the least prediction errors should comprise the least value of AIC and BIC.

Table 2 Formula of AIC and BIC

Akaike Information Criterion (AIC)	Bayesian Information Criterion (BIC)
$AIC = ln\hat{\sigma}^2 + \frac{2}{m}r$	$BIC = ln\hat{\sigma}^2 + \frac{\ln m}{m}r$

Diagnostic checking is important for ensuring the residuals meet the major assumptions of adequate. A model is adequate if the residuals satisfy the properties of constant on variances, independent and normally distributed with zero means and variance. The modified Box-Pierce (Ljung-Box) Ch-Square (LBQ) statistics also tested to choose the model which satisfy all the adequate assumptions.

Artificial Neural Network

The back propagation algorithm is used in layered feed forward ANN. The idea of the back propagation algorithm is to reduce this error, until the ANN learns the training data. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers. The first layer is the input layer where to input data, the second layer is the hidden layer to process data and the last layer is the output layer to produce the result.

The technique of Min-Max normalization is used for data preprocessing. The process of Min-Max normalization involves rescaling the values of each feature so that they fall within the range of 0 to 1. The formula of the min-max normalization of this training dataset, Y_t with original input X_t is shown as below.

$$Y_t = \frac{X_t - 22.9}{31.4 - 22.9} = \frac{X_t - 22.9}{8.5} \tag{4}$$

Therefore, standardized data is used as the inputs to the MLP model. To rescale the standardized data to original scale, the data can be calculated by using the formula below.

$$X_t = 8.5\, Y_t + 22.9 \tag{5}$$

Forecasting Accuracy Performance

Model accuracy for forecasting in time series is evaluated by using MSE, RMSE, MAE and MAPE in this study. A good performance model should comprise the lowest error measurement.

$$MSE = \sum_{t=1}^{n} \frac{(y_t - F_t)^2}{n} \tag{6}$$

$$RMSE = \sqrt{\sum_{t=1}^{n} \frac{(y_t - F_t)^2}{n}} \tag{7}$$

$$MAE = \sum_{t=1}^{n} \frac{|y_t - F_t|}{n} \tag{8}$$

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - F_t}{y_t} \right| \tag{9}$$

4. Results and discussion

Time Series Analysis

The time series plot and trend analysis plot of the daily temperature of weather in Malaysia from January 2000 to December 2020 is visualized as Figures 1 and 2. Based on the Figure 1, we can observe that it has several fluctuations among the data. There exist several peaks of the daily temperature in Malaysia. As the time series of the daily temperature is not fluctuating around the mean temperature, the series is non-stationary. Based on the Figure 2, the data used is existing with increasing trend hence this indicates that it is not stationary. The steeper or slope of the trend line helps determine the magnitude and strength of the observed trend. The shallower upward slope shown in Figure 2 indicates that there is a weaker positive trend.

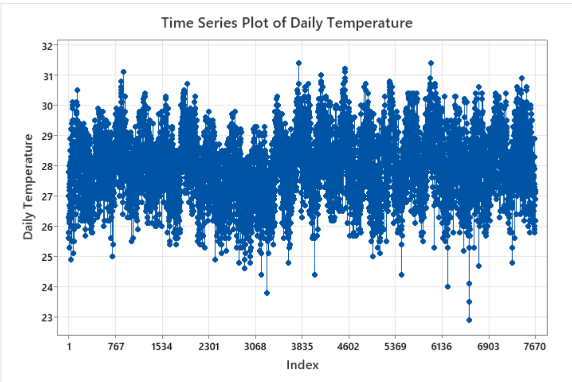

Figure 1 Time series plot of daily temperature

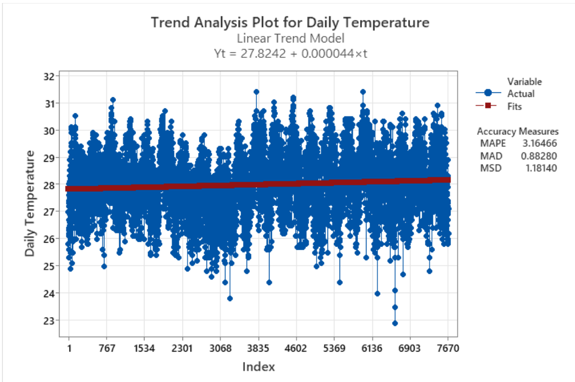

Figure 2 Trend analysis plot of daily temperature

The A'F test is a statistical test used to determine the presence of 61tationarityy of the data. The result of ADF test is conducted by Minitab and tabulated as Table 3. In ADF test, H_0 indicates that the time series is non-stationery and H_1 indicates the time series is stationary. The test statistic is equal to -8.0081 which is lower than the critical value of -2.8619, hence it has sufficient to reject null hypotheses. Therefore, the time series is stationary, thus time series is not required transformation such as differencing to employ ARIMA model.

Table 3 ADF test of daily temperature

Test Statistic	P-Value	Recommendation
-8.00810	0.000	Test statistic <= critical value of -2.86192. Significance level = 0.05 Reject null hypothesis. Data appears to be stationary, not supporting differencing.

However, the result of stationary from visualisation by using trend analysis plot and ACF plot should be prioritised compared with the result of stationary by using ADF test. Therefore, the dataset used for modelling from January 2000 to December 2020 is non-stationary. The differencing is required before modelled data. The ACF plot after differencing is visualized as Figure 4.7. The ACF plot shows that the values decay quickly. Therefore, this indicates that the series is stationary.

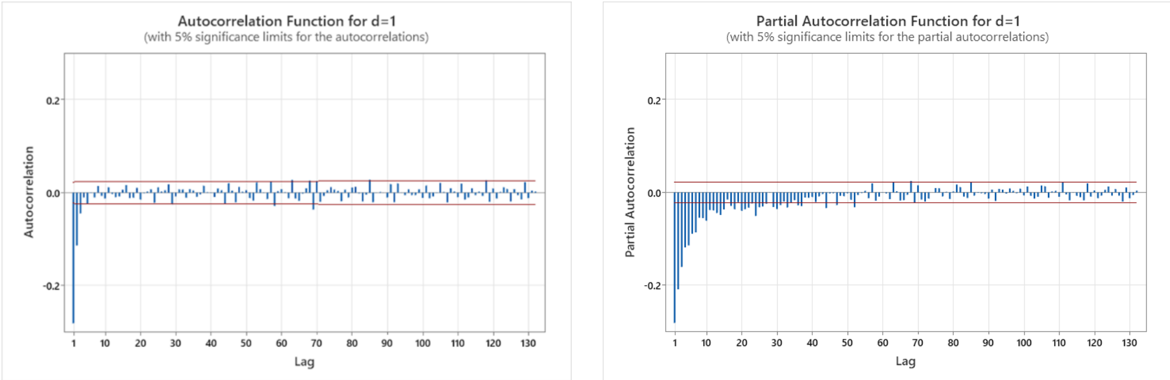

Figure 3 ACF and PACF plot of daily temperature in Malaysia

From the time series plot at Figure 1, it does not exist any repeating patterns or cycles that occur at regular intervals. The plot does not show any consistent fluctuations occurring within a specific time frame, such as daily, weekly, monthly, or yearly cycles. In addition, based on the ACF plot of the time series in Figure 3, it does not present any significant spikes at regular lag intervals. There are no oscillations occurred at lag=12 and lag=24. These properties shown that the time series does not exist seasonality from January 2000 to December 2020. In summary, the time series of the daily temperature from January 2000 to December 2020 is the stationary and non-seasonality data set.

Box-Jenkins Method

The transformation of differencing is required for this data set due to the ACF plot before differencing is decay extremely slowly. The value of $d = 1$ is obtained. Next, the order of p and q need to be determined, ARIMA $(p, 1, q)$. Therefore, the AIC value is calculated to choose the optimal order of values p and q. The tentative models with order combinations of $0 \leq p \leq 5$ and $0 \leq q \leq 5$ will be tested. The model with least AIC_c value is preferred to be chosen as the forecasting model in this study. The AIC_c values for each tentative models with different order combination are computed by using Minitab and the best forecast model with the least AIC_c value is ARIMA (3,1,5).

Next, the modified Box-Pierce (Ljung-Box) Chi-Square (LBQ) statistics is tested to choose the model which satisfy all the adequate assumptions. The model of ARIMA (3,1,5) is first to be tested and the result is illustrated as in Figure 4.

Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.8997	0.0718	-12.53	0.000
AR 2	0.4008	0.0370	10.82	0.000
AR 3	0.6985	0.0329	21.21	0.000
MA 1	-0.4364	0.0727	-6.00	0.000
MA 2	1.0146	0.0256	39.62	0.000
MA 3	0.7536	0.0466	16.17	0.000
MA 4	-0.3165	0.0271	-11.67	0.000
MA 5	-0.0949	0.0135	-7.01	0.000

Differencing: 1 Regular
Number of observations after differencing: 7670

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	9.82	22.22	30.22	42.34
DF	4	16	28	40
P-Value	0.043	0.136	0.353	0.370

Figure 4 Parameters estimation for model of ARIMA (3,1,5)

Based on the results in Figure 4, the t statistics is significant at $\alpha = 0.05$ as the p-value of all parameter estimators are smaller than 0.05. However, the LBQ statistics are not significant a $\alpha = 0.05$t as indicated by the p-values larger than 0.05. Therefore, all the tentative models with order combinations of $0 \leq p \leq 5$ and $0 \leq q \leq 5$ are tested before the selection of adequate model. The best model among these tentative models is selected based on the considerations of T-test, Q-test and also the values of AIC and BIC.

Table 4 Selective tentative models of ARIMA

Model	T-test	Q-test	AIC	BIC
$ARIMA$ (3,1,5)	√	×(lag 12<0.05)	**17774.5**	**17837.0**
$ARIMA$ (1,1,1)	√	×	17814.1	17841.9
$ARIMA$ (0,1,5)	√	×	17822.1	17870.7
$ARIMA$ (0,1,4)	√	×	17856.7	17898.4
$ARIMA$ (0,1,3)	√	×	17896.0	17930.7

Based on the Table 4, the t-test is the priority consideration to choose the best model. The value of AIC and BIC is then observed, the model of ARIMA (3,1,5) has the least value of both AIC and BIC. Although the Q-test for ARIMA (3,1,5) is not significant at $\alpha = 0.05$, by comparing between the tentative models, the model of ARIMA (3,1,5) is suggested to be an optimal and best model to use for forecasting the daily temperature in year 2021.

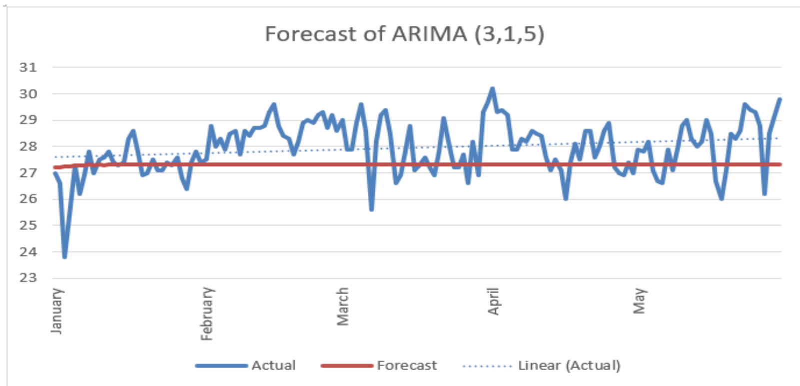

Figure 5 Forecast values of ARIMA (3,1,5)

Artificial Neural Network

Data preprocessing is a crucial step in preparing the training data for ANN. In this study, normalization for each raw input and output data set was calculated by the rescaling method, namely min-max normalization. The standardized data as visualised in Figure 6 is used as the inputs to the MLP model.

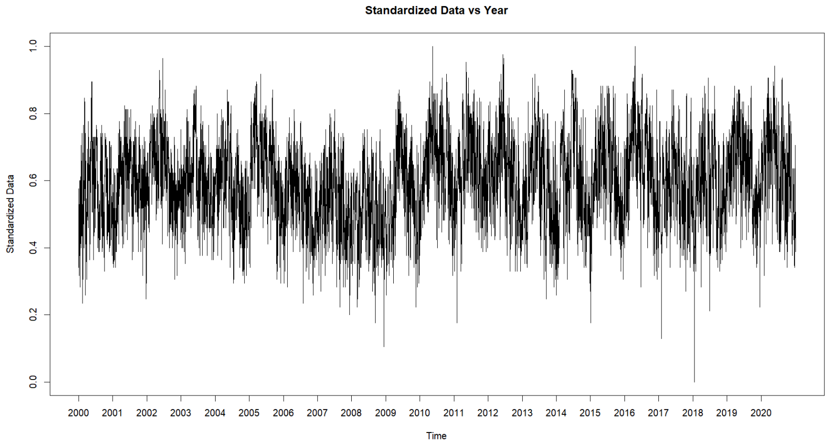

Figure 6 Standardized data of training set

The inputs combination of y_{t-1} to y_{t-p}, where p is the number of input nodes. The number of input nodes used is 1, 2, 3 and 4, while the number of hidden nodes used is 2, 4, 6 and 8 respectively. The best network architecture with the optimal number of hidden nodes can be chosen based on the least MSE values. The optimal complexity of ANN model is determined by a trial-and-error approach. The trials of 250, 500 and 1000 training repetitions are used to determine the optimal network architecture. By a trial-and-error approach, the ANN model of 4-8-1 has the minimal loss of MSE value of 0.008096, 0.008097 and 0.008045 which is lowest among all the tested architecture with 250 repetitions, 500 repetitions and 1000 repetitions respectively. Therefore, the network architecture of ANN with 4 input layer neurons, 8 hidden layer neurons and 1 output layer neuron (4-8-1) is chosen.

The MLP model has fitted the daily temperature of Malysia from January 2000 to December 2020 by using 4 input nodes, 8 hidden nodes and 1 output nodes. Since the value of MSE for 1000 repetitions is the lowest among other repetition numbers, the number of repetitions used to train the data is set to 1000. This indicates that the data is trained 1000 times to improve the accuracy of the modelling process.

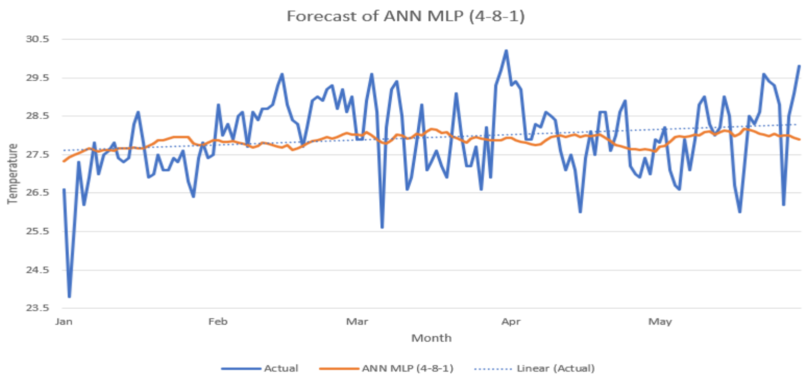

Figure 7 Forecast values of ANN MLP (4-8-1)

Forecasting Accuracy Performance

The performance of the forecast accuracy can be evaluated based on the time series plot of the actual data and predictions data from both methods shows in Figure 8. The error measurement is calculated to get the more accurate forecasting performance evaluation and the result is tabulated in Table 5.

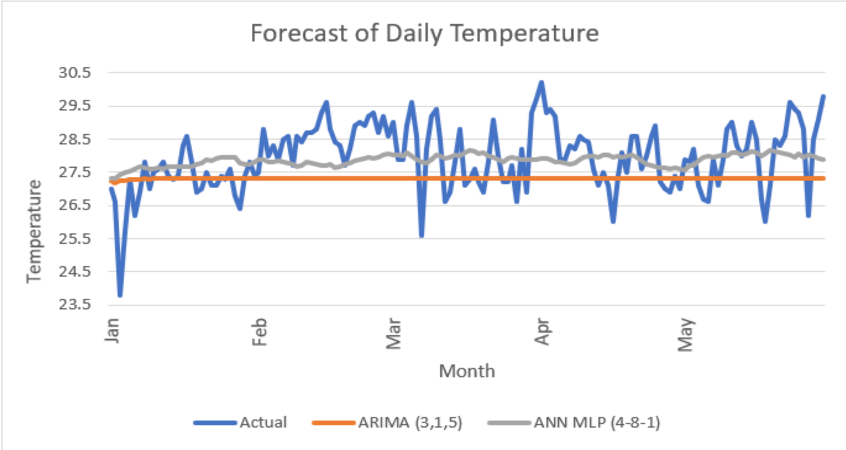

Figure 8 Forecast values of ARIMA (2,0,1) and ANN MLP (4-8-1) with actual data

Table 5 Forecast accuracy of employed models

Error Measurement	ARIMA (3,1,5)	ANN MLP (4-8-1)
MSE	1.364671	0.925550
RMSE	1.168192	0.962055
MAE	0.940573	0.780155
MAPE	2.749607	2.801578

The forecasting of the daily temperature for 150 days onwards starting from January 2021 by using both the model of ARIMA (2,0,1) and ANN MLP (4-8-1) are considered as highly accurate. The values of error measurement for MSE, RMSE, and MAE are lower than one for both ARIMA (2,0,1) and ANN MLP (4-8-1). The value of error measurement for MAPE is equal to 2.7496% and 2.8016% for ARIMA (2,0,1) and ANN MLP (4-8-1) respectively. The low error measurement observed in our analysis signifies the accuracy and precision of our forecasting model.

Conclusion

Overall, the evaluation criteria used to assess the forecast accuracy performance indicated that the ANN MLP (4-8-1) model outperformed the ARIMA (3,1,5) model. This implies that the ANN MLP model provided more precise and accurate forecasts compared to the ARIMA model. A low error measurement indicates that the predictions closely align with the true values, indicating the effectiveness of our model in capturing the underlying patterns and trends in the data. The selection of the ARIMA (2,0,1) model was based on its ability to capture the underlying patterns and dynamics in the historical temperature data, while the ANN MLP model was employed for training purposes. However, when it came to forecasting, the ANN MLP model demonstrated superior performance in terms of accuracy.

References

[1] Mathur, S., Kumar, A., & Chandra, M. (2017). A Feature Based Neural Network Model for Weather Forecasting, World Academy of Science, Engineering and Technology.

[2] Kumar, N., & Jha, G. K. (2013). Time Series ANN Approach for Weather Forecasting. International Journal of Control Theory and Computer Modeling, 3(1), 19–25. https://doi.org/10.5121/ijctcm.2013.3102

[3] Darji, M., Vipul Dabhi, & Prajapati, H. B. (2015, March 19). Rainfall forecasting using neural network: A survey. ResearchGate; unknown.

https://www.researchgate.net/publication/280451881_Rainfall_forecasting_using_neural_network_A_survey

[4] Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition. Procedia Computer Science, 48, 173–179. https://doi.org/10.1016/j.procs.2015.04.167

[5] Adhikari, R., & K, A. R. (2013). An Introductory Study on Time Series Modeling and Forecasting. ArXiv.org. https://doi.org/10.48550/arXiv.1302.6613

[6] Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis forecasting and control - rev. ed. Oakland, California, Holden-Day, 37 (2), 238 – 242

[7] Mohsen Hayati and Zahra Mohebi, “Temperature Forecasting Based on Neural Network Approach”,World Applied Sciences Journal ,Volume 2(6), pp 613-620,2007

[8] P.P. Kadu, Prof. K.P. Wagh, Dr. P.N. Chatur, “A Review on Efficient Temperature Prediction System Using Back Propagation Neural Network”, International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459), Volume 2, Issue 1, pp 52-55, January 2012

[9] Saxena A., Verma N., & Dr K. C. Tripathi, “A Review Study of Weather Forecasting Using Artificial Neural Network Approach”, International Journal of Engineering Research & Technology, ISSN: 2278-0181, Volume 2, Issue 11, November 2013