Thai Journal of Mathematics
Special Issue: The 17th IMT-GT ICMSA 2021
Pages 146–156 (2022)
http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Daily Temperature Predictive Model and Trend Analysis in Yangon, Myanmar Over 25 Years

Marinee Uma[1],Salang Musikasuwan[1,*], Rattikan Saelim[1],Santhana Chaimontree[1] and Kriangsak Damchoom[1]

[1] *Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Pattani, Thailand 94000*
e-mail: 6420320004@email.psu.ac.th (M. Uma); salang.m@psu.ac.th (S. Musikasuwan); rattikan.s@psu.ac.th (R. Saelim); santhana.c@psu.ac.th (S. Chaimontree); kriangsak.d@psu.ac.th (K. Damchoom)

Abstract Myanmar is rich in minerals and precious stones. However, agriculture, forestry, and fishing together were the main contributors to Myanmars economy. Like most other developing countries, the productions of these three main contributors are robustly dependent on the climate. The aims of this study were (i) to analyze the trend of temperature change in Yangon city, Myanmar in the last 25 years and (ii) to investigate the appropriate models for the temperature prediction in Yangon, Myanmar. Daily temperature from 1996 to 2020 was considered. It was assumed in general that the current temperature would be affected by a number of temperature of some times earlier. Hence the autoregressive integrated moving average (ARIMA) model was taken into account and found that ARIMA(4,0,0) was the best model in terms of the smallest values of the Akaike Information Criterion (AIC) and Root Mean Square Error (RMSE). Moreover, the multiple linear regression (MLR) and neural network models both taking time-lag temperatures as the independent variables has confirmed the result by the ARIMA model. In fact, the MLR suggested that the last 1 up to 4 days temperatures have significantly contributed to the current daily temperature, while the neural network model slightly outperforms the other two models measured by the RMSE. Furthermore, a broad picture of the temperature in Myanmar's capital has not altered considerably in the last 25 years, according to simple linear regression.

MSC: 49K35; 47H10; 20M12

Keywords: daily temperature; pattern and trend analysis; temperature change; autoregressive integrated moving average; linear regression; neural network

Submission date: 31.01.2022 / Acceptance date: 31.03.2022

1. Introduction

Temperature is one of the main climate variables. Temperature can provide effects in many fields such as chemical reaction, human health, human activity, animal behavior, etc. There were many studies confirmed that temperature has effects on crop yields [1–4]

*Corresponding author.

Published by The Mathematical Association of Thailand.

and cattle production [5]. On one hand, crop needs temperature at a certain range to survive. On the other hand, extremely high and low temperatures can have a greater effect on plants growth and yields [6]. In fact, a study on temperature thresholds [1] for several food crops such as wheat, broccoli, peanut, tomato, soybean, etc., has shown and confirmed that extreme temperature has negative effects on crop productions. Moreover, the increasing temperature and precipitation variability increase the risk to yield [3].

Myanmar is one of those developing countries whose economy is partially relying on agriculture. It is reported by Food and Agriculture Organization (FAO) of the United Nation [7] that the agriculture sector in Myanmar contributes up to 37.8 percent of gross domestic product (GDP). According to FAO, the livestock, and fisheries sectors alone account for more than 7% of the national GDP. It is hence of great important to learn about how temperature change in Myanmar. There are some studies focus on temperature change in Myanmar [8–10]. Recently, Wang (2018) has studied factors affecting land surface temperature (LST) change in Yangon, Myanmar [10]. The study found that the patterns of land use have potential impact on LST change there.

This paper aimed to develop models for predicting daily temperature in Myanmar from the station data (1996-2020) provided. Firstly, the seasonal pattern was achieved as a primary exploration. Then, the simple linear regression was fitted with the seasonally adjusted daily temperature to get the tendency. After that, as in general time series analysis, the ARIMA model was considered for prediction. Finally, as a validation for the ARIMA, the time series was fitted by a multiple regression model and a neural network, taking time-lag variables as the predictors for the current daily temperature.

2. DATA AND METHODS

2.1. STUDY AREA

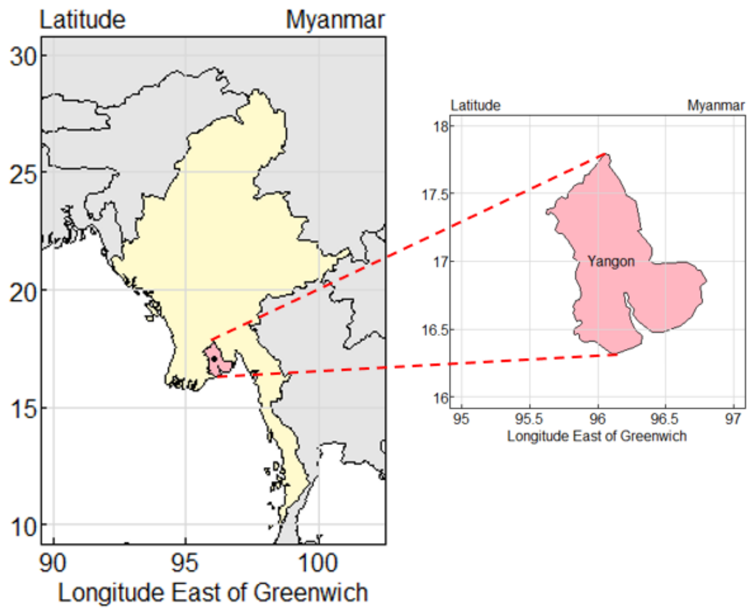

FIGURE 1. The location of Yangon city in Myanmar map

Yangon, also known as Rangoon, is Myanmar's main city and is located in the lower Myanmar region at latitude 16.80 and longitude 96.15, about 23 meters above sea level. The city of Yangon has a total size of 598.8 square kilometers. Myanmar's capital was Yangon until 2006, when the military regime relocated administrative functions to Naypyidaw, a purpose-built capital city in northern Myanmar. Yangon is Myanmar's most populated metropolis and most important commercial center, with a population of about 7 million people [11]. The location of Yangon is presented in Figure 1. It has a tropical monsoon climate and features a lengthy wet season from May through October where a substantial amount of rainfall is received, and a dry season from November through April, where little rainfall is seen. Average daily temperature used in this study focused on only Yangon area.

2.2. DATA COLLECTION

From January 1996 to May 2020, the daily temperature of Myanmar was gathered from https://academic.udayton.edu/kissock/http/Weather/ [12] (totally 8901 observations). This website's data comes from the National Climatic Data Center and is free to use for research and non-commercial purposes.

2.3. SEASONAL PATTERNS ANALYSIS

The seasonal patterns have been analyzed to the effect of the season in the region after all missing data has been imputed. Figure 5 depicts the daily temperature observations in Myanmar as a scattered plot (grey dots). A temperature time series is well known to have a seasonal influence, and the season is repeated annually. The linear trend was obtained using simple linear regression to gain a clear image of how daily temperature change (the tendency) over the last 25 years. The seasonal pattern is obtained using a method similar to that employed in [13–15]. Specifically, the seasonal pattern of dally temperature is determined by averaging temperatures from the same day of the year over a long period of time (25 years).

2.4. METHODS

To analyze the long-term trend of observation data, simple linear regression was applied in this study. The simple linear regression model takes the following form (Equation 2.1):

$$\hat{y}_t = \beta_0 + \beta_1 t + \epsilon \tag{2.1}$$

where $\hat{y}_t$ is the daily average temperature, β_0 is the intercept, β_1 is the regression coefficient, t is the time (in day) and ϵ is the error term.

Stationary variables, whose statistical features such as mean and variance are constant over time, can be used to create time series models. The prediction model was fitted using the Autoregressive Integrated Moving Average (ARIMA) model. According to the literature, the ARIMA model is one of the most effective methods for forecasting time series. In particular, the method includes fitting an appropriate model, estimating the parameters, and evaluating the model [13]. Equation 2.2. shows the ARIMA(4,0,0) for a particular example of $p = 4$:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + e_t \tag{2.2}$$

where, $\hat{y}_t$ is the seasonal adjusted temperature at observation t; ϕ_1, ϕ_2, ϕ_3 and ϕ_4 are constants y_{t-1}, y_{t-2}, y_{t-3}, and y_{t-4} are the four-order lag phase $(t-1, t-2, t-3, t-4)$ of seasonally adjusted data; and e_t is the value not explained by the predictors.

Multiple linear regression analysis is a statistical method that aims to create a suitable regression model for predicting the value of the dependent variable which, in this case, is the daily average temperature. Based on the values of independent variables or predictors, the multiple linear regression model has the following Equation 2.3:

$$\hat{y}_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + ... + b_k y_{t-k} \tag{2.3}$$

where, $\hat{y}_t$ is the dependent variable (predicted temperature at time t)
y_{t-i} is the temperature at time $t - i$, $i = 1, 2, ..., k$
b_0 is the intercept (constant term)
b_i is the slope coefficients for temperature at time $t - i$, $i = 1, 2, ..., k$.

An Artificial Neural Network (ANN), often known as a Neural Network (NN), is a network of nodes with input, output, and hidden layers. Each node, or artificial neuron, is connected to the others and has a weight and threshold linked with it. If a node's output exceeds a certain threshold value, the node is activated, and data is sent to the next layer of the network. Otherwise, no data is sent on to the network's next layer. A neural network is a sophisticated model that may be used for both classification and regression. Both classification and regression problems have been successfully solved using NN. Figure 2 depicts the architecture of NN.

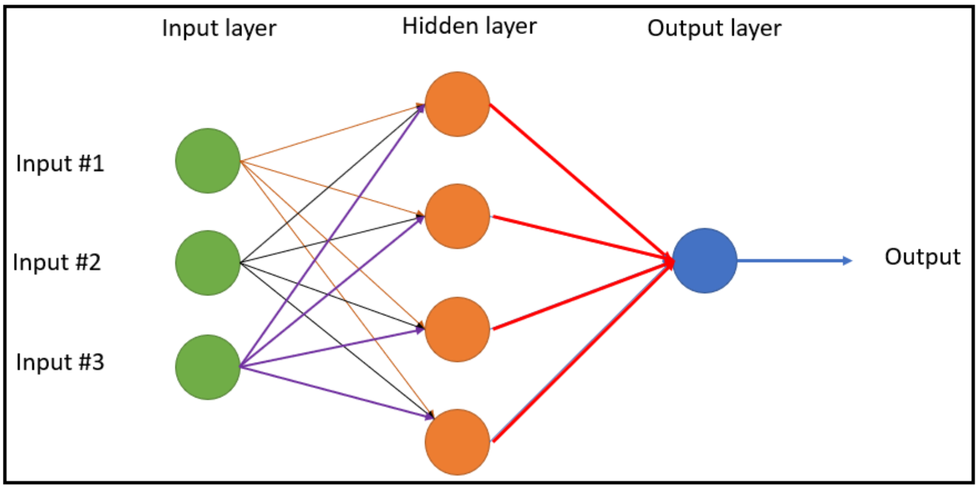

FIGURE 2. The architecture of Neural Network (adopted from [16])

After obtaining the appropriate models, the predicted value is calculated from the training and testing datasets and those models are evaluated by using RMSE as shown in Equation 2.4:

$$RMSE = \sqrt{\sum_{t=1}^{n} \frac{(y_t - \hat{y}_t)^2}{n}}, \tag{2.4}$$

where n is the number of observation data, t is time, y_t is the observation at time t, and $\hat{y}_t$ is the predicted value.

3. RESULTS

From obtained data, there were 165 missing observation (value equal to -99 in data records). Mean substitution has been applied for dealing with missing data. Figure 3 showed the plots of daily average temperature of collected data.

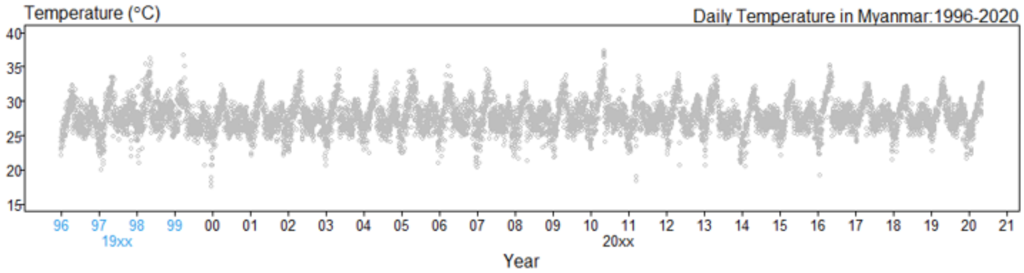

FIGURE 3. The plots of daily average temperature in Yangon city during 1996-2020

Before developing the predictive model, the data were split into two sets with 70%:30% for training and testing datasets, respectively. The training dataset was used for developing the predictive model, while the testing dataset was used for evaluating the model. The plots of data observation for training and testing datasets can be shown in Figure 4.

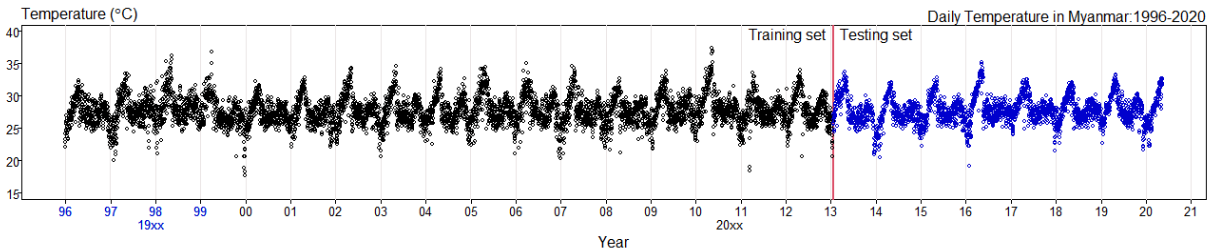

FIGURE 4. Time-series plots of daily average temperature for training dataset (black) and testing dataset (blue)

3.1. SEASONAL PATTERNS ANALYSIS

The resulting seasonal pattern is illustrated by the red curve in Figure 5. After that, by subtracting the seasonal effect from the original observation and adjusting by adding back the average value of the original series, the seasonally adjusted daily temperature is achieved as shown in Figure 6 (grey dots).

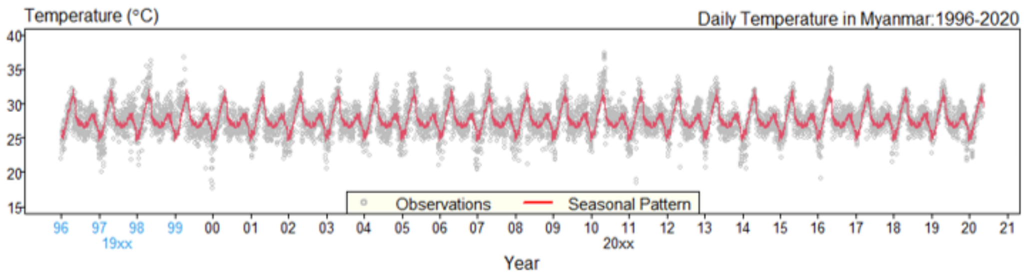

FIGURE 5. The plots of daily average temperature and its seasonal patterns in Yangon city during 1996-2020

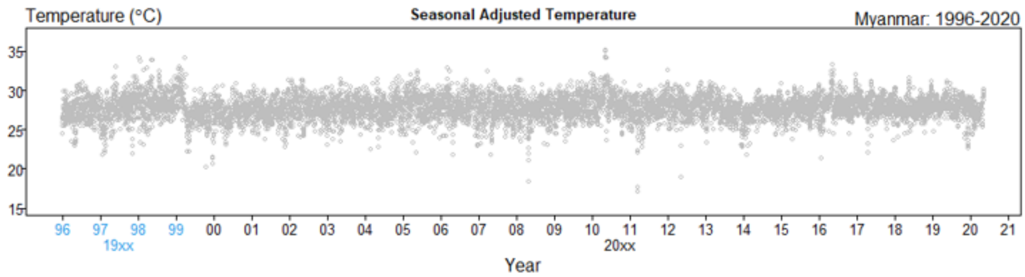

FIGURE 6. The plots of daily average temperature with season adjusted in Yangon city during 1996-2020

3.2. TEMPERATURE TREND ANALYSIS

The long-term trend of daily average temperature in Yangon had been analyzed by using simple linear regression. Figure 7 showed the plots of the trend (red line) and original data observations (grey dot) for the whole dataset.

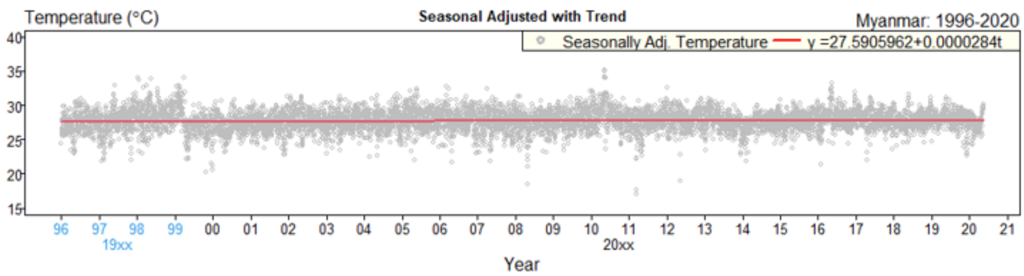

FIGURE 7. The plots of the trend (red line) and seasonally adjusted daily temperature (grey dot) in Yangon city during 1996-2020

To analyze the trend of daily average temperature, the simple linear regression has been fitted with the whole dataset. The obtained formula can be shown in Equation 3.1:

$$\hat{y}_t = 27.5905962 + 0.0000284t. \tag{3.1}$$

From the Equation 3.1, it can be observed that the temperature in Yangon have been slightly increased around 0.0103 degree Celsius per year or 1.03 degree Celsius per century.

3.3. DAILY TEMPERATURE PREDICTING MODELS

ARIMA

After the imputation for the missing values and checking the stationary of the data, it was found that all data were met and satisfied the assumptions of using ARIMA model. The training dataset has been used to create the ARIMA model by considering the graph of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to determine the q and p parameters of ARIMA model as shown in Figure 8.

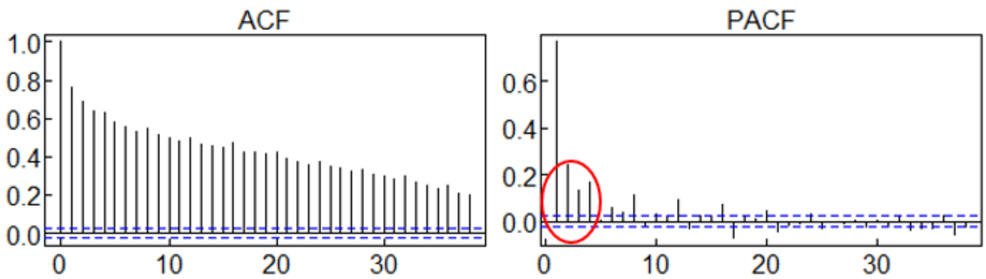

FIGURE 8. The plots of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

From Figure 8, the ACF is gradually decreasing while the PACF appears to have 4 distinct consecutive spikes, these suggest that the parameters q and p would be 0 and 4, respectively. Notice that the original daily temperature is stationary and its mean and variance are approximately constants over the whole period. Consequently, the time series is ready to be analyzed and hence the parameter $d = 0$. Now, four ARIMA(p, d, q) models for $p = 1, 2, 3$, and 4 with zeros d and q are considered.

TABLE 1. Results obtained from ARIMA analysis on training and testing dataset.

		RMSE	
Model	AIC	Training Dataset	Testing Dataset
ARIMA(1,0,0)	21987.63	1.412	0.498
ARIMA(2,0,0)	21625.24	1.372	0.830
ARIMA(3,0,0)	21518.57	1.359	1.035
ARIMA(4,0,0)	21346.83	1.341	1.101

The AIC and RMSE of the four models indicate that ARIMA(4,0,0) gives the best performance as shown in Table 1. The performance is best in the sense that it produces the smallest AIC and RMSE for the training data while the RMSE for the testing data are not much different among the four models. Consequently, ARIMA(4,0,0) can be formulated into linear equation as shown in Eqation 3.2:

$$\hat{y_t} = 3.5174 + 0.5251y_{t-1} + 0.1359y_{t-2} + 0.0416y_{t-3} + 0.1704y_{t-4} \tag{3.2}$$

where $\hat{y_t}$ is the estimate of today temperature and y_{t-k} for $k = 1, 2, 3, 4$ are the temperature of the previous k days.

To evaluate the performance of ARIMA(4,0,0), Equation 3.2 has been used to calculate the desired predicted values for both training and testing datasets. Figure 9 showed the time series plots of predicted values for both training (red dots) and testing (blue dots) datasets against original observations (grey dots).

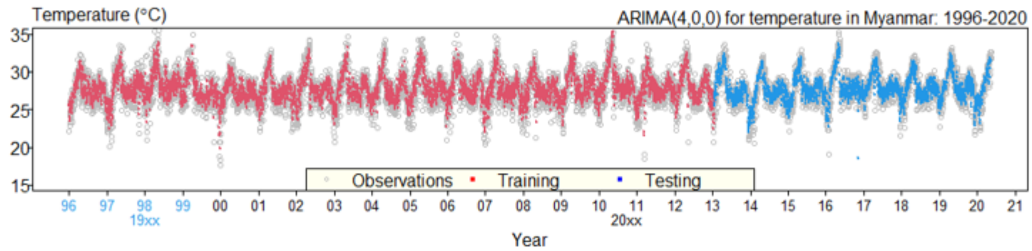

FIGURE 9. The time series plots of ARIMA(4,0,0) predicted values against the original observations.

Multiple Linear Regression

Now, when fitting the multiple regression to the training dataset by taking the current temperature as the outcome variable and the 4 consecutive time-lag temperatures as the predictors, it turned out that the 4 predictors are highly significantly explaining the outcome variable. In this study, it was found that the 4 predictors were the average temperature at time-lag 1, 2, 3, and 4 of day earlier. Obtained MRL model can be finally written in formula form as shown in Equation 3.3:

$$\hat{y_t} = 3.5245 + 0.5249y_{t-1} + 0.1357y_{t-2} + 0.0417y_{t-3} + 0.1704y_{t-4}. \tag{3.3}$$

To evaluate MRL model, Equation 3.3 has been used for calculating the predicted values for both training and testing datasets. Figure 10 showed the time series plots for predicted values gained from MRL model for both training (red dots) and testing (green dots) datasets against original observation data (grey dots).

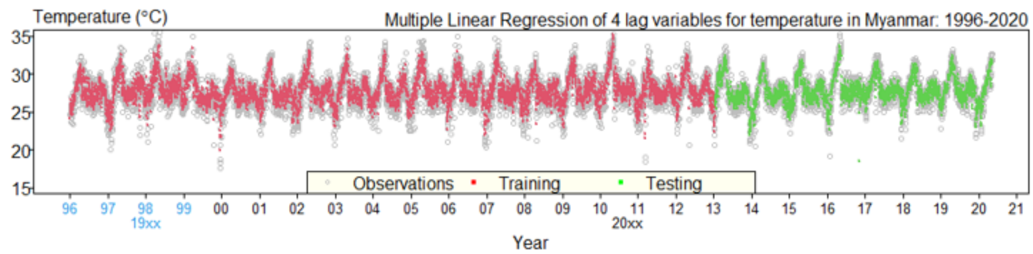

FIGURE 10. The time series plots of MLR model predicted values against the original observations.

Neural Network

The neural network model is applied to predict the daily temperature by using the library *neuralnet* in R program. With continuous outcome variable, the model with two hidden layers of 2 and 1 nodes give the fitted temperature for both training and testing sets as illustrated in Figure 11.

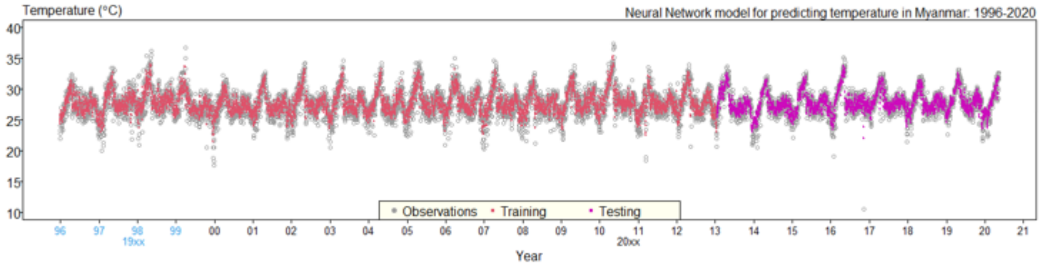

Figure 11. The time series plot of neural network model predicted values against the original observation

3.4. Model performance comparison

Once ARIMA, MLR, and neural network models have been developed, the root mean square errors of each model have been calculated for evaluating the models. The obtained RMSE gained from each model for both training and testing datasets can be shown in Table 2.

Table 2. Root mean square error of each model

Model	RMSE	
	Training Dataset	Testing Dataset
ARIMA(4,0,0)	1.341	1.101
MLR	1.340	1.066
Neural Network	1.337	1.059

4. DISCUSSION AND CONCLUSION

This study aimed to analyze the trend of temperature change in Yangon city, Myanmar over 25 years, and investigate the appropriate models for the temperature prediction in Yangon, Myanmar. Daily average temperature from 1995 to 2020 was used to conduct the study. It was assumed in general that the current temperature would be affected by the temperature of a number of times earlier.

From the analysis of seasonal patterns as shown in Figure 5, it was found that the daily average temperatures in Yangon, Myanmar had clear seasonal patterns associated with 3 seasons: winter starts from late October to mid-February, which was influenced by the dry northeast monsoon, summer starts from mid-February to mid-May which was influenced by the dry monsoon season and the rainy season starts from mid-May to late October, which is influenced by the southwest monsoon.

For analyzing trend of daily temperature, simple linear regression was used to analyze trend of temperature change over 25 years in Yangon city, Myanmar. From the analysis, the temperature in Yangon city has been slightly increased (p-value<.001) aroualnd 0.0103 degree Celsius per year or 1.030 degree Celsius per century.

From the results obtained, ARIMA(4,0,0) and MLR (with 4 consecutively time-lags of earlier days) models were fitted and chosen as suitable predictive models. Both models performed pretty good based on the RMSE values as shown in Table 2. The coefficients of ARIMA(4,0,0) are pretty close to the parameters obtained by MLR. Moreover, the neural network model also confirmed the result by the other two models. In fact, the neural network performs slightly better, in term of the relative small values of RMSE, than the other two models.

In conclusion, all three models can be used as a predictive model when we applied with daily average temperature in Yangon city, Myanmar. However, since this study assessed only daily average temperature in Yangon city, so the results may be different when other climate variables are taken into account.

Acknowledgements

This work was supported by Faculty of Science and Technology, and Graduate School, Prince of Songkla University, Pattani Campus, Thailand. We would like to thank the referees for their comments and suggestions on the manuscript.

References

[1] Q. Luo, Temperature thresholds and crop production: a review, Climatic Change 109 (3) (2011) 583-98.

[2] T. Parthasarathi, G. Velu, P. Jeyakumar, Impact of crop heat units on growth and developmental physiology of future crop production: a review, Journal of Crop Science and Technology 2 (1) (2013) 1-8.

[3] J.R. Porter, M.A. Semenov, Crop responses to climatic variation, Philosophical Transactions of the Royal Society B: Biological Sciences 360 (1463) (2005) 2021-35.

[4] P.K. Thornton, P.G. Jones, G. Alagarswamy, J. Andresen, Spatial variation of crop yield response to climate change in East Africa, Global Environmental Change 19 (1) (2009) 54-65.

[5] X. Wang, H. Gao, K.G. Gebremedhin, B.S. Bjerg, J. Van Os, C.B. Tucker, G.A. Zhang, Predictive model of equivalent temperature index for dairy cattle (ETIC), Journal of Thermal Biology 76 (2018) 165-70.

[6] C. Zhao, B. Liu, S. Piao, X. Wang, D.B. Lobell, Y. Huang, M. Huang, Y. Yao, S. Bassu, P. Ciais, J.L. Durand, Temperature increase reduces global yields of major crops in four independent estimates, Proceedings of the National Academy of Sciences 114 (35) (2017) 9326-31.

[7] Food and Agriculture Organization of the United Nations. 2021. FAO in Myanmar. Available online: https://www.fao.org/myanmar/fao-in-myanmar/myanmar/en/ [October 16, 2021]

[8] K.K. Sein, A. Chidthaisong, K.L. Oo, Observed trends and changes in temperature and precipitation extreme indices over Myanmar, Atmosphere 9 (12) (2018) 1-21.

[9] M. Nagashima, T. Armbruster, T. Hainschwang, A temperature-dependent structure study of gem-quality hibonite from Myanmar, Mineralogical Magazine 74 (5) (2010) 871-85.

[10] Y.C. Wang, B.K. Hu, S.W. Myint, C.C. Feng, W.T. Chow, P.F. Passy, Patterns of land change and their potential impacts on land surface temperature change in Yangon, Myanmar, Science of the Total Environment 643 (2018) 738-50.

[11] Wikipedia. 2021. Yangon. Available online: https://en.wikipedia.org/wiki/Yangon [October 15, 2021]

[12] University of Dayton. 2020. Environmental protection agency average daily temperature archive. Available online: http://academic.udayton.edu/kissock/http/Weather/default.htm [August 15, 2021]

[13] R. Anokye, E. Acheampong, I. Owusu, I.E. Obeng, Time series analysis of malaria in Kumasi: using ARIMA models to forecast future incidence, Cogent Social Sciences 4 (1) (2018) 1-13.

[14] J. Hassan, ARIMA and regression models for prediction of daily and monthly clearness index, Renewable Energy 68 (2014) 421-7.

[15] R. Saelim, S. Musikasuwan, N. Chetae, Land surface temperature changes in Songkhla, Thailand from 2001 to 2018, Naresuan University Journal: Science and Technology (NUJST) 28 (3) (2020) 39-45.

[16] L. Fernandez-Arjona, A neural network model for solvency calculations in life insurance, Annals of Actuarial Science 15 (2) (2021) 259-75.