DAVAO RESEARCH JOURNAL

P-ISSN 2244-4432
E-ISSN 2984-7125

Time Series Analysis of Mean Temperature using SARIMA: An example from Davao Oriental, Philippines

Larry M. ICHON Jr.[1*], Jerd M. DELA GENTE[1]

[1]Faculty of Computing, Data Science, Engineering, and Technology Davao Oriental State University City of Mati, Davao Oriental, 8200 Philippines. *ORCID Larry M. Ichon Jr.: https://orcid.org/0009-0008-0164-9465* , Jerd M. Dela Gente*: https://orcid.org/0001-0001-9711-9515 *Corresponding author: larryjr.ichon@dorsu.edu.ph*

ABSTRACT. In many practical disciplines, time series analysis, and forecasting—a technique that predicts future values by analyzing past values—play a substantial role. In this paper, the researchers analyze the monthly mean temperature in Davao Oriental from 2010 to 2022 using the SARIMA (Seasonal Auto-Regressive Integrated Moving Average) technique. Data from January 2010 to May 2020 were used as the training data set, while data from June 2020 to December 2022 were used as the testing data set. The presentation includes a thorough overview of model selection and forecasting precision. The findings demonstrate that the suggested research strategy achieves good forecasting accuracy. The analysis reveals that the best model which was satisfactory to describe was SARIMA (0,1,3) (2,0,0) [12], and in the month of May 2023, the temperature will be 28.28 ^{0}C. In subsequent work, the researchers hope to expand the number of possible grid search parameter combinations. This method may lead us to models with improved predictive ability. The length of the training set may also affect forecasting accuracy, in addition to the SARIMA model's parameters. A follow-up study is needed to investigate both hypotheses.

Submitted: 5 March 2023
Revised: 6 May 2023
Accepted: 11 May 2023
Published: 27 June 2023

https://davaoresearchjournal.ph

Keywords: *climate change, Davao Oriental, forecasting, mean temperature, time series*

How to cite: Ichon, L.M., Dela Gente, J.M. (2023). Time Series Analysis of Mean Temperature using SARIMA: An Example from Davao Oriental, Philippines. Davao Research Journal (DRJ), 14(1), 73-83, https://doi.org/10.59120/drj.v14i1.91

INTRODUCTION

A major environmental threat facing the entire world today is climate change. It is one of the growing issues concerning water resources, livelihoods, and forest diversity (Hatfield et al., 2018). In climatic studies, the analysis of long-term hydroclimatic data sets is crucial. Despite decades of research on public perceptions of climate change, much of it has focused on why some developed countries have seen a decrease in public concern. Why developing countries are more concerned about climate change is unknown. It's important to consider local contexts and factors that affect this increase in public opinion (Bollettino et al., 2020).

Climate change is associated with a worldwide increase in temperature. Since 1850, the temperature rise has become a common occurrence. The period between 1850 and 1900 is called the "preindustrial" period. There has been a rising acknowledgment of climate change, even though experts are divided on the connection between climate change and the frequency of cyclones or hurricanes (Singh, 2021). When an extreme weather or climate event happens, people often ask questions about the role of climate change, whether and how much the event can be linked to climate change, and whether the event is a sign of what is to come. Extreme event attribution is a field that tries to find answers to these questions and is seeing many new methods and approaches being made. Several groups are now doing in-depth analyses and have made many case studies. Some types of analyses, like the attribution of temperature and large-scale precipitation extremes, have been done so often and with such consistent results that they could be operationalized. The general method can now be standardized for these types of events, and only certain parts, like model evaluation, need to be changed on a case-by-case basis. Research teams have often done attribution studies in close to real-time. So that operational analyses can be compared, there needs to be a protocol for how they are designed and how they are put together. This paper is meant to be a place to start. It can also be used as a standard method for figuring out who did what in an academic setting (Philip et al., 2020).

The Philippines has long been one of the most climate-vulnerable countries in the world. Much of the country is vulnerable to the growing frequency and severity of natural catastrophes because of its position in the Western Pacific. In the study of Lagmay and Racoma (2018) entitled "Lessons from tropical storms Urduja and Vinta disasters in the Philippines," the analysis of these incidents reveals that the Pre-Disaster Risk Assessment (PDRA) alerts for both storms were largely ineffective because they were too general and broad, calling for forced evacuations in too many provinces.

The climate of Davao Oriental, which is characterized by tropical rainforest conditions, can be attributed to the region's elevation of 25.67m (84.22 ft) above sea level. The annual temperature in this province, which is 27.33 ^{0}C, is 0.11 ^{0}C, higher than the average for the Philippines. There are 111.55 rainy days per year in Davao Oriental, which accounts for 30.56% of the total time, and the city receives an average of 66.05 mm (2.6 inches) of precipitation. Weather predictions are becoming increasingly inaccurate due to global warming, and not just in Davao Oriental (Scher and Messori, 2019). Also, climate change will slow the growth of agricultural productivity. It is expected to change agricultural productivity by affecting cropping schedules, yield quality and levels, the spread of pests and diseases, livestock and fisheries production, and infrastructure (Gomez, 2018).

A common method for analyzing and predicting stationary and seasonal time series data is using autoregressive integrated moving averages (ARIMA) and seasonal ARIMA (SARIMA) models. Mehta (2022) used three forecasting techniques, SARIMA (Seasonal Auto Regressive Integrated Moving Average), ARMAX (ARMA with Exogenous Variables), and ARMA (Auto Regressive Moving Average), in his study entitled "Fuzzy Logic based Crop Yield Prediction using Temperature and

DOI: https://doi.org/10.59120/drj.v14i1.91

Rainfall parameters predicted through ARMA, SARIMA, and ARMAX models." Based on comparing the three models' performances, the best model is used to forecast temperature and rainfall, which are then used to forecast crop yield using a fuzzy logic model. By analyzing data from the previous twelve years (1994-2006), the Box Jenkins method was used to predict the temperatures and precipitation for the next five years. The results demonstrated that the seasonal ARIMA model provided accurate and satisfactory monthly predictions for precipitation and temperature parameters.

Another study that used SARIMA Model to forecast the monthly mean temperature was Kabbilawsh et al. (2020). They created SARIMA models using the Box-Jenkins methodology for two hydroclimatic variables: Minimum Temperature and Maximum Temperature of 4 (four) NEI stations. According to the study, SARIMA takes different forms at different stations. It demonstrates spatial variation between various stations. Second, it can be seen from the graphical evidence that the SARIMA produces better-forecasted values for temperature than for rainfall in terms of model performance.

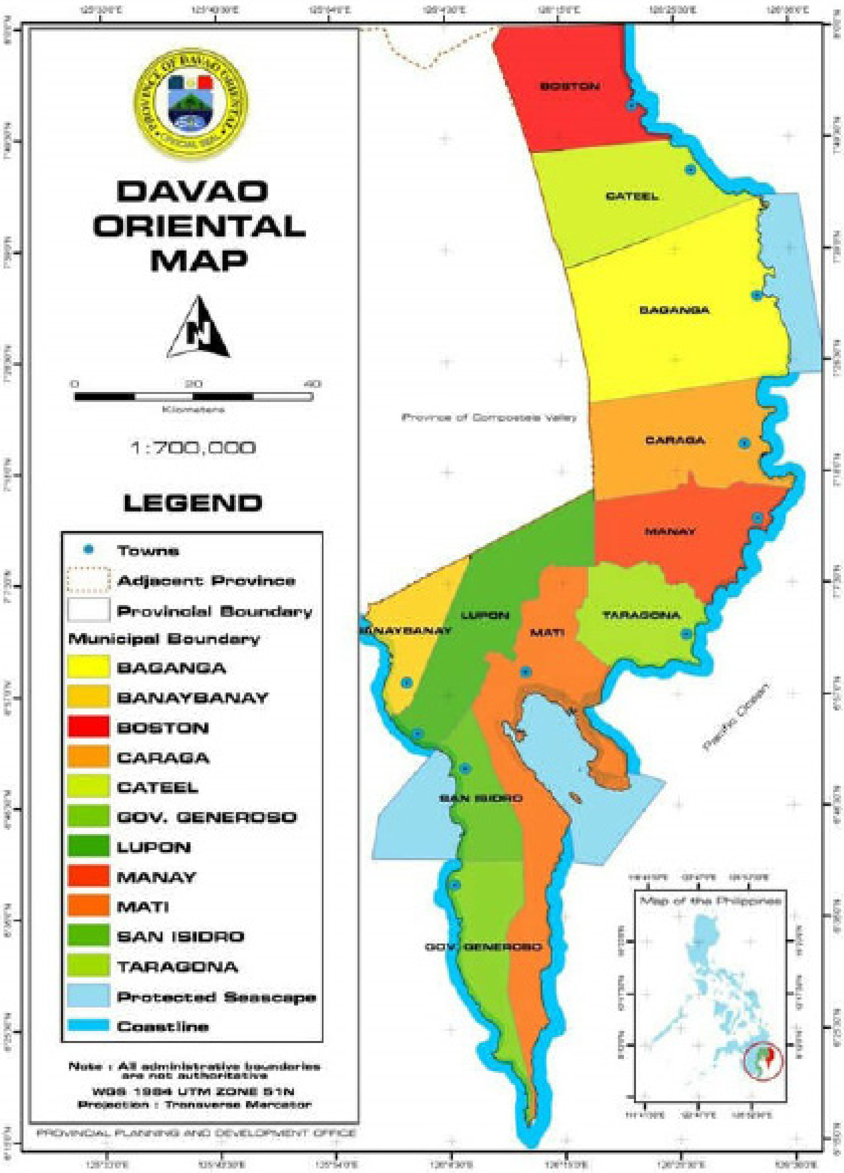

Figure 1. Area of the Study

The objective of this study is to forecast values of the mean temperature in the province of Davao Oriental using the data from January 2010 to December 2022. This study significantly impacts various aspects of human society and the natural environment, and it is important to monitor and manage its effects to ensure the health and well-being of people and the planet. The SARIMA model could help to determine possible future strategies in the respective field for the Davao Oriental and nearby provinces to address the problems brought by climate change. An example of some problems the region faces was stated in the study by Cabrera and Lee (2018), 95.91% of Davao Oriental is currently at low or moderate flood risk, but that should drop to 95.75%. High and very high flood-risk areas cover 3% of the province and are unlikely to change. 28 of 183 barangays (towns) are at high or very high flood risk, but only one will be added in the coming years. These barangays are mostly riverside or coastal. Decision-makers must act now to create a community-based disaster risk plan for the future.

The study of Sabino et al. (2020) showed that the mean minimum temperature went up by 0.74 ^{0}C (p 0.01) and the mean maximum temperature went down by 0.65 ^{0}C (p 0.01). Rainfall patterns showed a decreasing trend and showed significant changes in June (p 0.01), August, and December (p 0.05), which shows that climate change and variability happened because farming households experienced floods, landslides, and drought. The study showed that most farming households had "moderate to high vulnerability" to climate change and variability. As climate change brings new risks, the right adaptation strategies are needed to deal with current and future vulnerabilities. This requires a strong vulnerability assessment based on recent scientific progress and new strategies that are similar to this study.

Bollettino et al. (2020) found that 59.9% of respondents had never heard of climate change and did not feel very well informed about it. Only 11.7% of respondents had heard much about it and felt well informed. The rest thought they were "somewhat informed" (28.4%). When asked about the effects of climate change, responses at the national average were ranked as follows: Higher temperatures (46%); Seasons that change (41.6%); Heavier rains (23.4%); More changes in weather patterns (20.9%); More powerful tropical cyclones (13.2%); More tropical cyclones (7.2%); More flooding (6.2%); and A later start to the rainy season (4.4%).

MATERIALS AND METHOD

Data Collection

The study was conducted last April 2023 and was finished last May 2023 at Davao Oriental State University. The data used in the study were secondary and gathered monthly from Time and Date (*Past Weather in Province of Davao Oriental, Philippines — Yesterday or Further Back*, 1998). There were 156 collected data from January 2010 to December 2022.

Statistical Analysis

In this case, SARIMA (Seasonal Autoregressive Integrated Moving Average) was used to forecast future values using RStudio Version 1.2.5033. When non-seasonal and seasonal factors are included in a time series, a multiplicative seasonal ARIMA model can be applied. The SARIMA model requires stationary data. The unit root test can be used to determine whether or not the data is stationary. If not stationary, the data should be transformed into stationary data.

The first data preparation steps involved Box and Jenkins (1970) methodologies. Transform data to stabilize variance and difference data to obtain stationary series. Second, model selection. Determine potential models by analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF). Third, estimation. Evaluate parameters in possible models. Determine which is the best model by applying the appropriate criteria. Fourth, Diagnostic checking. Check the residuals' ACF and PACF. Determine whether or not residuals follow white noise. If it does not follow white noise, use the model selection criterion to choose another model. Lastly, use the model for forecasting if residuals follow white noise.

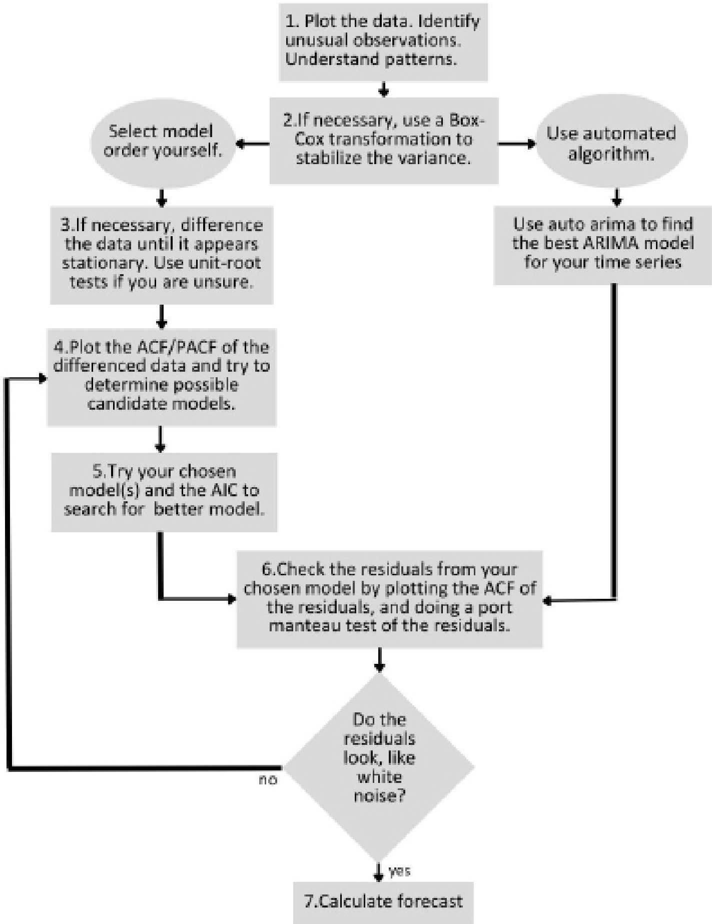

Figure 2. A process on SARIMA.

RESULTS AND DISCUSSION

Preliminary Analysis

Table 1 shows the descriptive statistics for mean temperature in Davao Oriental. The mean and standard deviation of the temperature were 28.13 ^{0}C and 0.73 ^{0}C, while the minimum and maximum values for the temperature were 26 ^{0}C and 30 ^{0}C, respectively.

Table 1. Descriptive statistics for the mean temperature data (January 2010-December 2022).

Variable	Mean	Std Dev	Minimum	Maximum
Temperature	28.13	0.73	26	30

Figure 3 shows that the mean temperature of Davao Oriental was non-monotonic, as observed by the lines going up and down. This behavior will be used to study the behavior of time series. In the study by Lu (2022) entitled *"Correlation of*

Climate Change Indicators with Health and Environmental Data in the Philippines", the results of the study showed that the number of people who died from air pollution, household air pollution, malaria, and other tropical diseases is going down in the Philippines. However, deaths from non-communicable diseases increased from 41.99 in 1990 to 55.00 in 2016. Conversely, the average temperature has a strong negative correlation with household air pollution, the number of malaria cases, and neglected tropical diseases. It also has a strong positive correlation with non-communicable diseases.

Analysis showed that natural landscapes increase the surface temperature, while other urban forms decrease the surface temperature. Even though both sites showed the same trend for every urban form, the exploratory regression analysis report showed that the adjusted R2 values were different for each set of independent variables for each study site. This shows that different urban forms have different effects on different places.

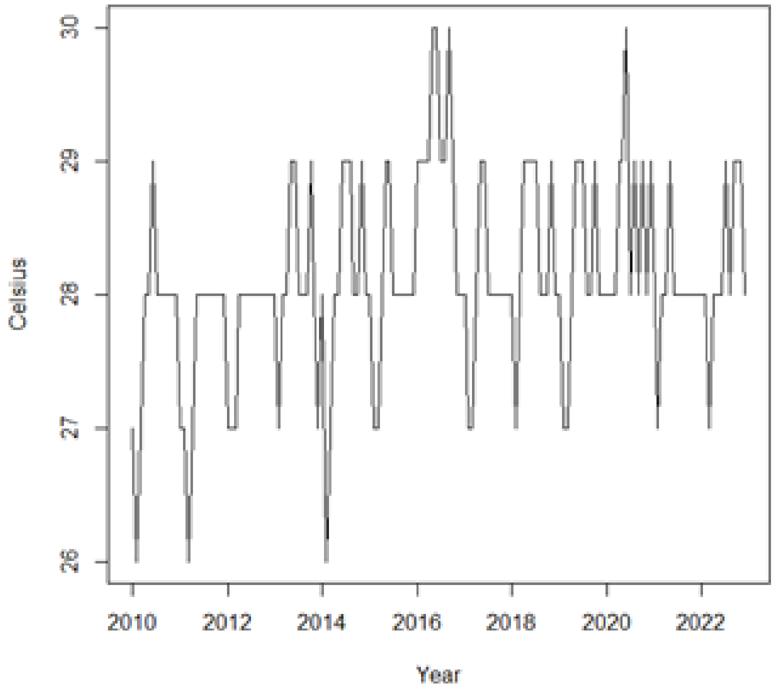

Figure 3. Time series plot of monthly mean temperature of Davao Oriental.

ARIMA (Box-Jenkins) Modelling

The data were divided into a training set (80%) consisting of 126 observations and test data (20%) comprising 30 observations. The data was checked to determine whether stationary or not by using KPSS (Kwiatkowski–Phillips–Schmidt–Shin) Test before it could be used to estimate and develop a model.

Table 2. Results after the first differencing.

	KPSS level	*P*-value	Stationary
Original Data	0.69	0.01	No
First Difference	0.02	0.10	Yes

Upon differencing the data, results showed that the *p*-value was 0.65, which is lower than our significant value of 0.05. We can say that the data is not stationary. Therefore, we need to transform the data before determining the potential models. After differencing, the KPSS level becomes stationary, as shown in Table 2. After transforming the data to stationary, testing the Auto correlational function and partial auto-correlational function will be tested to know the number of lags used to identify the models.

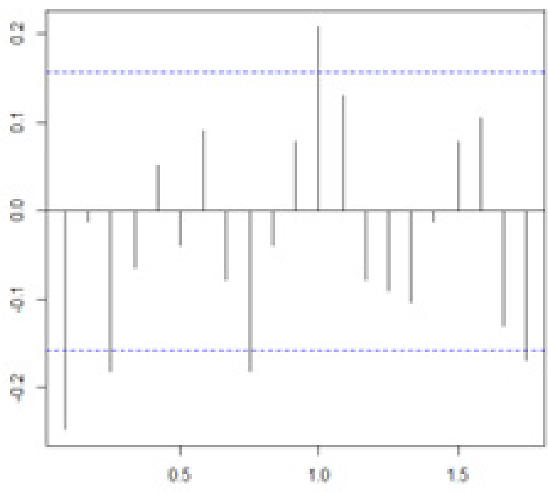

Figure 4. ACF plot after first difference.

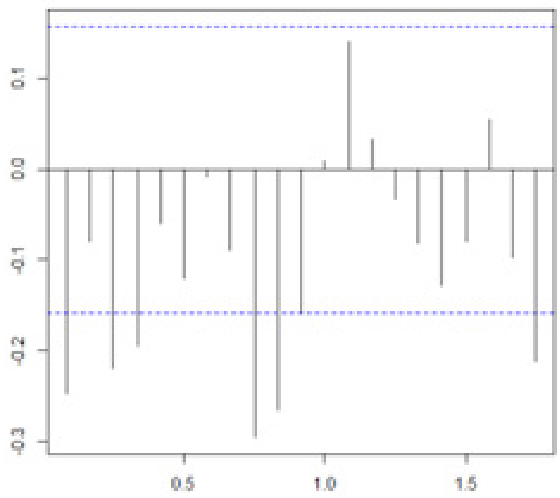

Figure 5. PACF plot after first difference.

After reaching the stationary level, exploring models from the SARIMA (p, d, q) (P, D, Q) family is necessary, where p and q represent the possible order of AR and MA components. P, D, and Q are seasonal Autoregressive orders, seasonal differencing, and seasonal Moving Average orders, respectively. For determining appropriate values for p and q, the plot of sample ACF and PACF of the mean temperature series (Figure 2) was analyzed. Figure 4 shows that the sample PACF has one substantial autocorrelation at lag 1. As a result, it is possible to hypothesize under the ARIMA model that the fitted AR order is 1.

Table 3. Parameters of the SARIMA models.

		Estimate	Standard Error	*P*-values
SARIMA (0,1,3) (1,0,0)	MA1	-0.40	0.09	0.00
	MA2	-0.36	0.12	0.00
	MA3	-0.17	0.09	0.07
	SAR1	0.33	0.09	0.00
SARIMA (0,1,4) (1,0,0)	MA1	-0.38	0.09	0.00
	MA2	-0.32	0.10	0.00
	MA3	-0.08	0.10	0.42
	MA4	-0.17	0.10	0.07
	SAR1	0.34	0.10	0.00
SARIMA (0,1,3) (2,0,0)	MA1	-0.45	0.10	0.00
	MA2	-0.36	0.12	0.00
	MA3	-0.13	0.10	0.16
	SAR1	0.27	0.10	0.00
	SAR2	0.18	0.10	0.07

By comparing the values presented in Table 3, it will be now possible to determine which forecasting model to employ. The smallest values indicate a better-fitting model. SARIMA (0,1,4) (1,0,0) in Table 3 displayed smaller values than other estimated models. SARIMA (0,1,3) (1,0,0) displayed the smallest values among the three models in standard errors. SARIMA (0,1,3) (1,0,0) presented the smallest p-values compared to the other three models, as seen in the table. Finally, Table 4 compared the models' accuracy metrics. It indicates that the RMSE, MAE, and MAPE for the validation set from SARIMA (0,1,3) (2,0,0) model deviates from the observed data shown in the table, which can be considered within an acceptable range.

Table 4. Results of model estimation.

	RMSE	MAE	MAPE
SARIMA (0,1,3) (1,0,0)	0.57	0.43	1.53
SARIMA (0,1,4) (1,0,0)	0.56	0.43	1.52
SARIMA (0,1,3) (2,0,0)	0.56	0.42	1.50

The comparison of the three fitted models using the Akaike Information Criterion (AIC), the Corrected Akaike Information Criterion (AICc), and the Bayesian Information Criterion (BIC) is presented in Table 5. It was discovered that, among the three models, SARIMA (0,1,4) (1,0,0) had the lowest AIC, AICc, and BIC values. However, by looking at the values of the p-values model, 3 has more significant values. In light of these comparisons, we can conclude that SARIMA (0,1,3) (2,0,0) will be used to predict future values for the Davao Oriental mean temperature.

In the study of Ayitey et al. (2021), the data spans from January 1960 to December 2020 and includes average monthly temperatures for Northern Ghana. The KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test was used to determine the data is stationary prior to estimation and model development. The model created during the process was SARIMA (1,0,0) (1,0,0) [12]. In this study, the same process was used before doing the model diagnostic.

Table 5. Accuracy measures of SARIMA models.

Model	AIC	AICc	BIC
SARIMA (0,1,3) (1,0,0)	225.61	226.12	239.71
SARIMA (0,1,4) (1,0,0)	224.30	225.02	241.22

Model Diagnostic

This study used the Ljung-Box Test to determine whether the residuals are independently distributed. The best-case scenario would be if the null hypothesis was not rejected. In other words, it is desirable to have a *p*-value of greater than 0.05, as this indicates that the residuals for our time series model are independent, which is a common assumption made when developing such models.

The result of Ljung-Box test shows 0.07 indicates higher than the alpha, 0.05. This means that the residuals are normally distributed and follow the white noise. As a result, we cannot conclude that the data values are dependent and must instead accept the test's null hypothesis.

Table 6. The P-value of the SARIMA (0,1,3) (2,0,0) with the statistical.

Statistical Test	*P*-value
Ljung-Box Test	0.07

Table 6 displays the results of the Ljung-Box Test, which indicates that the model is normally distributed (as depicted in Figure 5) and does not require further adjustment because the *p*-value is more significant than the alpha level of 0.05 (0.07). Figure 6 below shows the SARIMA (0,1,3) (2,0,0) fitted line plot of observed and forecasted monthly mean temperature in Davao Oriental.

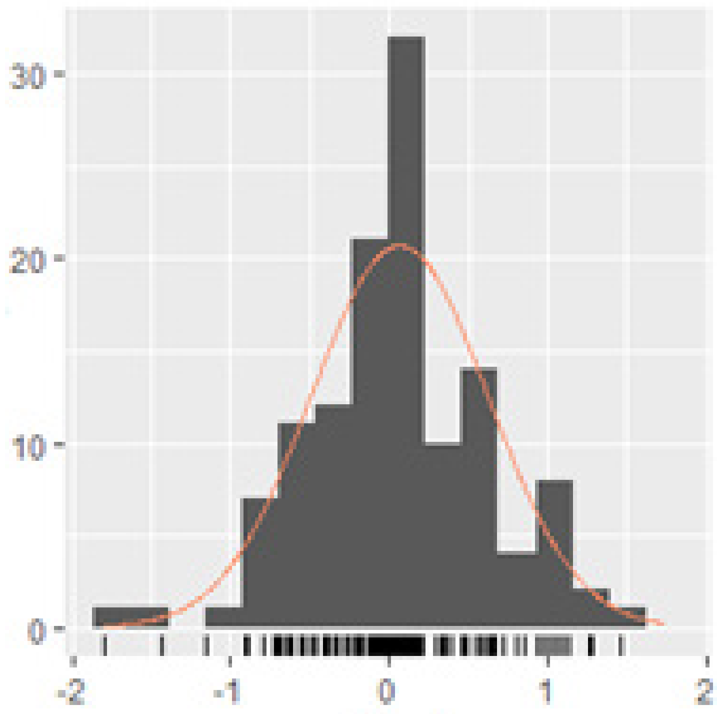

Figure 6. Normality plot of mean temperature using SARIMA (0,1,3) (2,0,0).

Using the model, table 7 and Figure 6 shows predicted values of mean temperature in Davao Oriental from January 2023 to May 2023 having 95% confidence interval. Results showed that almost the same meant temperature in the next five months from December 2023. In May, it has a higher mean temperature upon using the model.

Table 7. Forecasted values of mean temperature in Davao Oriental from January 2023 to May 2023 using SARIMA (0,1,3) (2,0,0) with a 95% confidence interval.

Month	Forecasted Value
January	28.03
February	27.86
March	27.90
April	28.14
May	28.28

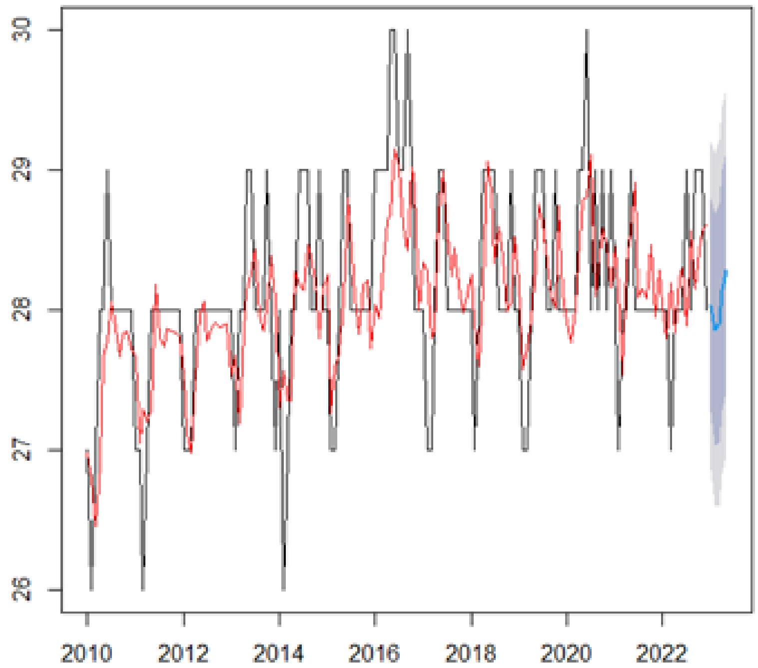

Figure 7. Original plot (black line), fitted values (red line), and forecast values (blue line).

In relation to the study of Weslati et al. (2023), the SARIMA (1,1,1)(0,1,1)12 model shows a steady trend of 470 mm of rain per year, on average. The most rain is likely to fall in January (60 mm) and December (55 mm), while the least is likely to fall in July (7 mm). However, the model's prediction is different from what has been seen in the past because it only accounts for irregular patterns of rainfall in places where the monthly rainfall is less than 100 mm. The model predicts that precipitation will act the same way for the next five years because it copies the trend from the first year and repeats it for the next four years. This is because SARIMA's one-period-ahead prediction method treats data that was predicted in the past as historical data.

CONCLUSION

Autocorrelation, partial autocorrel ation, and Kwiatkowski–Phillips–Schmidt–Shin test all showed that the set of mean temperatures in Davao Oriental between 2010 and 2022 was stationary. The seasonality in the mean temperature time series could only be removed by performing a single seasonal differentiation.

Following the model-building process outlined, several candidate models were developed using the Box Jenkins method, and AIC, AICc, and BIC values were calculated for each model. The ultimate winning model was SARIMA (0,1,3) (2,0,0) [12]. When checked, residuals were found to be uncorrelated because they followed the same pattern as random white noise and the residuals also had a normal distribution. The model was tested and resulted in a potential model for forecasting. The results show predicted values of mean temperature in Davao Oriental from January 2023 to May 2023, having a 95% confidence interval. Results showed that there is a fluctuation in mean temperature for the next five months from December 2022.

ACKNOWLEDGEMENT

I would like to express my gratitude to all of the lecturers and supervisors who gave their best to help me reach the objective and to urge me to keep moving forward. This work would not have been possible. Their passion, expertise, and meticulous attention to detail have inspired me and kept my work on track. To my friends and classmates, my sincere gratitude for taking the time to assist me in creating my project and for being there for me whenever I needed it.

REFERENCES

Ayitey, E., Kangah, J., & Twenefour, F. B. (2021). Sarima modeling of monthly temperature in the northern part of Ghana. *Asian Journal of Probability and Statistics*, 37–45. https://doi.org/10.9734/ajpas/2021/v12i330287

Bollettino, V., Alcayna-Stevens, T., Sharma, M., Dy, P., Pham, P., & Vinck, P. (2020). Public perception of climate change and disaster preparedness: Evidence from the Philippines. *Climate Risk Management*, 30, 100250. https://doi.org/10.1016/j.crm.2020.100250

Gomez, N. (2018). Climate change and adaptation on selected crops in Southern Philippines. *International Journal of Climate Change Strategies and Management*, 7(3), 290–305. https://doi.org/10.1108/ijccsm-03-2013-0014

Kabbilawsh, P., Sathish Kumar, D., & Chithra, N. R. (2020). Trend analysis and SARIMA forecasting of mean maximum and mean minimum monthly temperature for the state of Kerala, India. *Acta Geophysica*, 68(4), 1161–1174. https://doi.org/10.1007/s11600-020-00462-9

Lagmay, M., & Racoma, B. A. (2018). Lessons from tropical storms Urduja and Vinta disasters in the Philippines. *Disaster Prevention and Management: An International Journal*, 28(2), 154–170. https://doi.org/10.1108/dpm-03-2018-0077

Lu, J. L. (2022). Correlation of Climate Change Indicators with Health and Environmental Data in the Philippines. *Acta Medica Philippina*, 56(1). https://doi.org/10.47895/amp.v56i1.3976

Past Weather in Province of Davao Oriental, Philippines — Yesterday or Further Back. (1998). Past Weather in Province of Davao Oriental, Philippines — Yesterday or Further Back. https://www.timeanddate.com/weather/@1715342/historic

Sabino, L., Pulhin, J., Dizon, J., Espaldon, M. V., & Cruz, R. V. (2020). Climate variability, change and the impacts on livelihood vulnerability of farming households in Koronadal, South Cotabato, Philippines. *Journal of Environmental Science and Management*, 42–59. https://doi.org/10.47125/jesam/2020_sp2/04

Scher, S., & Messori, G. (2019). How global warming changes the difficulty of synoptic weather forecasting. *Geophysical Research Letters*, 46(5), 2931–2939. https://doi.org/10.1029/2018gl081856

Singh, S. (2021). Global climate change. *Elsevier*. Amsterdam, Netherlands, ISBN- 9780128230978, p. 3.

Takemi, T., & Yamasaki, S. (2020). Sensitivity of the intensity and structure of tropical cyclones to tropospheric stability conditions. *Atmosphere*, 11(4), 411. https://doi.org/10.3390/atmos11040411

Weslati, O., Bouaziz, M., & Serbaji, M. M. (2023). Precipitation forecasting and monitoring in degraded land: A study case in Zaghouan. *Land*, 12(4), 738. https://doi.org/10.3390/land12040738