International Journal of Social Science and Education Research Studies

ISSN(print): 2770-2782, ISSN(online): 2770-2790

Volume 04 Issue 03 March 2024

DOI: https://doi.org/10.55677/ijssers/V04I3Y2024-09, Impact Factor: 6.759

Page No : 234-243

Mathematical Modeling of Dengue Cases in Region XI Philippines

Louie Resti S. Rellon[1], Rosalie M. Baclay[2], Milanie E. Ogayre[3]

[1,2,3] University of Mindanao, Davao City, Philippines

Published Online: March 20, 2024

ABSTRACT

Dengue was recognized by the World Health Organization (WHO) in 2023 as a prominent arthropod-borne viral illness and poses a significant global health threat. This study focuses on Region XI in the Philippines, where dengue is alarmingly prevalent, and in forecasting dengue cases, it employs the ARIMA model. The study, spanning from January 2019 to August 2023, integrates univariate exploratory data analysis and statistical methods to understand the patterns and trends of dengue cases. The ARIMA model, specifically (0,0,1) with a nonzero mean, the candidate model, was identified and utilized from RStudio through coefficient tests of AIC values. The forecasting results predict a decline in dengue cases over the next 16 months. The normality test of residuals and Q-Q plot affirm the model's reliability. A significant test using the two-sample t-test demonstrates a substantial difference between actual and forecasted values. In conclusion, this study provides crucial insights for public health planning, community intervention, and future research. The ARIMA model's successful application emphasizes the need for refined dengue control and mitigation strategies in Region XI. The results underscore the urgency of collective efforts to minimize dengue transmission and address the challenges posed by this prevalent and impactful disease.

KEYWORDS:

Dengue cases, ARIMA modeling, forecasting, univariate, Region XI

INTRODUCTION

Battling dengue is a relentless struggle as individuals grapple with debilitating symptoms, from excruciating joint pain to severe headaches and relentless fatigue. In 2023, the World Health Organization (WHO) identified dengue as a prevalent arboviral illness transmitted by female Aedes mosquitoes. Caused by the dengue virus with four serotypes, it rapidly spreads across tropical regions. Indications encompass fever, headache, joint discomfort, rash, and slight bleeding, persisting for almost one week, following an incubation period of 4 to 10 days post-infection.

Following a bite from an Aedes mosquito, individuals typically undergo a swift increase in temperature, redness, acute joint discomfort, nausea, muscle pain, vomiting, headache, and similar symptoms. When a mosquito bites a person, the virus is transferred through the mosquito's needle, having resided in the mosquito's body for several days. Subsequently, the virus reproduces and proliferates within the human body (Stolerman & Maia, 2019).

Corresponding Author: Louie Resti S. Rellon

****Cite this Article: Louie Resti S. Rellon, Rosalie M. Baclay, Milanie E. Ogayre (2024). Mathematical Modeling of Dengue Cases in Region XI Philippines. International Journal of Social Science and Education Research Studies, 4(3), 234-243***

Approximately 50 to 500 million individuals globally are believed to contract dengue annually (Sutriyawan et al., 2022). Moreover, an annual death toll of 10,000 to 20,000 individuals is reported, with approximately two and a half billion people at risk of infection (Abualamah et al., 2021).

The Philippines has consistently ranked as the top country in global dengue cases. It includes historical events such as the initial recorded dengue epidemic in Southeast Asia, which took place in Manila in 1954, as well as the highest-ever recorded global contribution to dengue cases in 2019, amounting to 437,563 cases (Ong E, Obeles, Ong B, & Tantengco 2022). In terms of regions, Central Luzon, accounting for 13% with 6,641 cases; Central Visayas, representing 12% with 6,361 cases; and the Zamboanga Peninsula, contributing 9% with 4,767 cases, were the leading contributors (Montemayor, 2022). Furthermore, Department of Health (DOH) (2020) reported that the Davao Region is among the 17 regions in the country where dengue is prevalent. Philippine Statistics Authority (PSA) (2021) stated that this region, Davao City, ranks as the third most populous city in the Philippines, boasting an almost two-million-strong population. Furthermore, it holds the distinction of being the largest city in terms of land area. Given that the disease predominantly spreads in densely populated urban settings, Davao City is the region's focal point for dengue incidence.

In 2019, Davao City documented 4,495 instances of dengue, constituting half of the total cases in the region. In the initial six months of 2021, the city recorded 1,431 dengue cases, marking a 66% surge compared to the previous year. A dengue alert is typically issued ahead of the rainy season, and hotspots are identified when cases cluster in the same area for at least two consecutive weeks. According to a 2023 study by the World Health Organization (WHO), regions with elevated temperatures, such as the Philippines, experience heightened dengue transmission rates due to the persistence of stagnant water, serving as potential breeding grounds for mosquitoes. The escalating cases indicate various underlying scenarios and causes that are challenging to pinpoint.

However, the current strategies for dengue control and mitigation may need to be more effective in addressing the expanding issue.

The ARIMA method, alternatively referred to as the Box-Jenkins methodology, is a widely used technique for time series forecasting. It generates predictions by considering the temporal correlations and the stochastic characteristics inherent in the analyzed data (Dagoumas, 2021). Forecasting methods like ARIMA rely on the presumption that the time series can be rendered nearly stationary. Consistent statistical properties characterized a constant mean and constant variance known as a stationary time series. ARIMA model has a non-seasonal ARIMA model and a seasonal ARIMA model. Singh, Sundram, Rajendran, Law, Aris, Ibrahim, and Gil (2020) claim that ARIMA models are a versatile and empirical approach capable of providing dependable forecasts, especially in scenarios with restricted data.

Furthermore, determining a suitable ARIMA model involves three key stages: model identification, parameter estimation, and model validation. The autocorrelation function (ACF) and partial autocorrelation function (PACF) are utilized as plots of the original time series data to assess stationarity and provide specifications for the p, d, and q parameter terms. The general configuration of an ARIMA model is expressed as p, d, and q, where p represents the autoregressive (AR) component order, d signifies the degree of differencing applied to the original time series through integration (I), and q indicates the order of the moving average (MA) component. Moreover, in their study, Thiruchelvam, Dass, and Asirvadam (2021) employ ensemble ARIMA models to analyze how neighboring regions impact the occurrence of dengue cases. The investigation compares two models —Autoregressive Integrated Moving Average (ARIMA) and Ensemble ARIMA—utilizing the Akaike Information Criterion (AIC) to assess their performance. Significantly, Ensemble ARIMA models demonstrate a superior fit compared to basic ARIMA models, attributed to their integration of neighboring effects from the seven districts of Selangor. The AIC values for ensemble ARIMA models are notably lower when compared to their traditional ARIMA counterparts.

Researchers developed a monthly ARIMA model to predict the trend of dengue cases in Region XI due to rising occurrences. The findings may serve as a guide to address the dengue cases; specifically, The Department of Health (DOH) Region XI will be able to use the information to facilitate the planning of public health interventions. The community, particularly in Region XI, will provide critical information for early intervention and resource allocation to mitigate the impact of outbreaks. Future Researchers can utilize this research as a local reference for their future use.

Researchers acquired dengue case data from the Department of Health – Region XI, focusing on a 16-month forecast (September 2023 to December 2024) using the ARIMA model. The study underscores the importance of collaborative efforts to reduce dengue transmission in Region XI.

METHOD

Research Design

Reich and Subrahmanian (2022) highlight the necessity of employing diverse perspectives to examine the subject thoroughly. It helps create a complete picture of the topic. Also, if the research design is too complicated, many studies fail to achieve their goals or explain how they did things. The researchers used univariate exploratory data analysis. Exploratory data analysis is the comprehensive process of visually and statistically summarizing a dataset, encompassing tasks like data cleaning and univariate analysis, to understand patterns and relationships to inform subsequent analysis or decision-making (Daele & Janssenswillen, 2023) and analyzing and interpreting the properties of a single variable. Understanding the distribution, trends, and summary statistics related to that variable is the only goal of this kind of study.

Research Data

The secondary data sets on monthly dengue cases were obtained from the proper authority from the DOH – Region XI. The coverage of these data began from January 2019 to August 2023.

Research Locale and Data Gathering Procedures

This study focuses on Region XI (Davao Region), which is in the southeastern part of Mindanao. Monthly data on dengue cases was manually organized, categorized, and coded, covering January 2019 to August 2023. Subsequently, the data was brought into R Studio for thorough examination and analysis. The obtained results are employed to unveil significant advancements in this research.

Statistical Treatment of the Data

This study employed the ARIMA model to predict the occurrences of dengue cases in Region XI from September 2023 to December 2024. Dhamodharavadhani and Rathipriya (2020) applied the ARIMA time series model to predict the dengue incidence rate in Tamil Nadu. This choice contributes to more accurate forecasting, potentially leading to better results in predicting the occurrence of dengue. Hence, the ARIMA model is a more popular and extensively employed statistical method for time-series forecasting.

The ARIMA model was scrutinized using the Box-Jenkins approach, deemed suitable for an extended forecasting period. Specifically, the model encompasses three types of parameters—autoregressive parameters (p), the number of differencing phases (d), and moving average parameters (q).

The foundation of the Box-Jenkins approach for modeling time series is encapsulated in Figure 1, comprising three phases: identification, estimation, and application

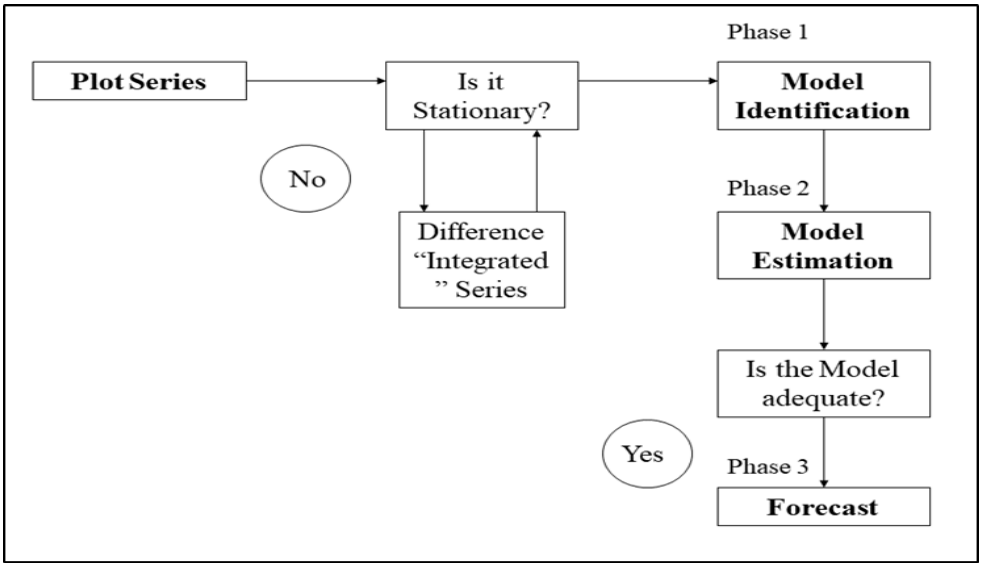

Figure 1. The Box Jenkins model building process.

RESULTS AND DISCUSSION

Model identification

The first step is plotting the data from Table 1 *(see page 9)* to determine whether the time series is stationary or non-stationary. In Figure 2 (*see page 9*), the dengue cases in Region XI, a total of 49,900, were observed from January 2019 to August 2023, given that the collected data points represent monthly cases related to the onset of the illness. Furthermore, the rate exhibited the highest number of cases, reaching its peak at 2606 in August 2019. Dengue cases have been relatively increasing in the region. The autocorrelation and partial autocorrelation functions were utilized in RStudio to determine the model associated with the observed dengue cases in Region XI and assess the stationarity of the data. Upon implementing the autocorrelation function, Figure 3 reveals a notable fluctuation in the ACF plot at lags 1, 2, 3, 4, 5, 6, and 7. In Figure 4, the partial autocorrelation function indicated one spike at lag 1. As observed, Figure 3 and Figure 4 (*see page 10*) indicate that the autocorrelation and partial autocorrelation function plots are not stationary; however, furtheranalysis is required to determine the specific trends and patterns in the data contributing to the observed non-stationarity.

Table 1. The actual data from the Department of Health XI, from January 2019- August 2023

MORBIDITY MONTH	**2019**	**2020**	**2021**	**2022**	**2023**
JANUARY	1716	840	371	427	1125
FEBRUARY	1457	890	484	436	1052
MARCH	1075	528	382	526	1057
APRIL	785	262	394	667	973
MAY	626	310	542	829	1547
JUNE	1245	252	695	1011	1814
JULY	2221	401	650	880	2183

AUGUST	2606	562	411	926	2596
SEPTEMBER	1604	544	279	919	
OCTOBER	1223	471	302	885	
NOVEMBER	1154	415	378	1189	
DECEMBER	923	337	442	1081	
TOTAL	16635	5812	5330	9776	12347

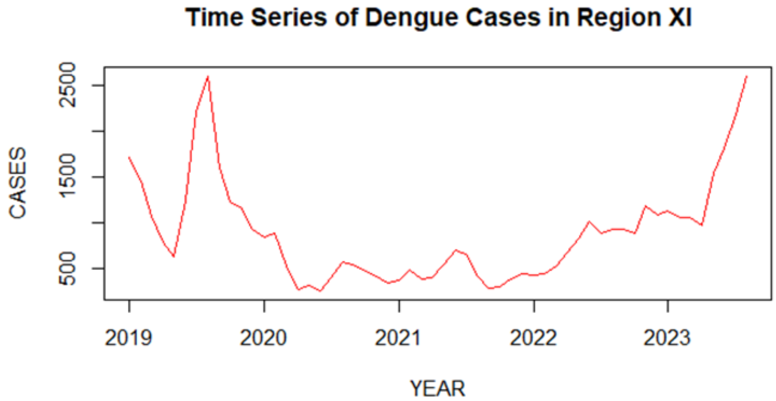

Figure 2. Dengue Cases in Region XI from (2019 to 2023) are represented graphically and acquired from the Department of Health – Region XI.

Figure 2 displays the monthly dengue cases in Region XI. Based on the plot, the series is an increasing trend from January 2019 to August 2023. The peaks and troughs are inconsistent since the monthly dengue cases keep fluctuating from high to low growth and vice versa. Additionally, it can beobserved that the graph lacks a consistent mean and variance over the period due to its recurring upward trends. It implies that the series may not be stationary.

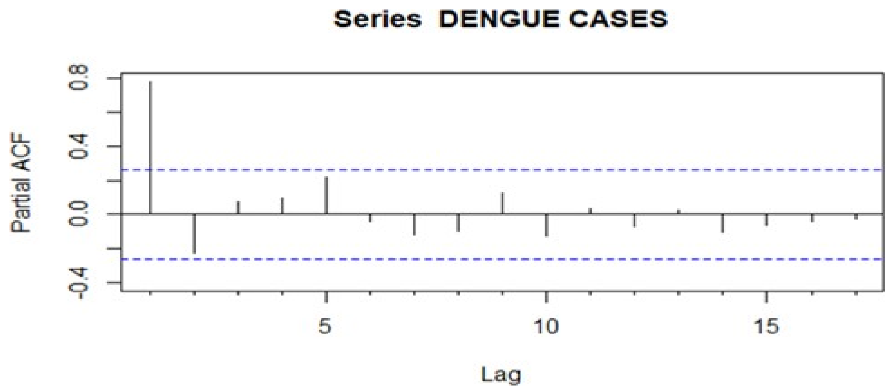

Figure 3. The plot represents the autocorrelation function.

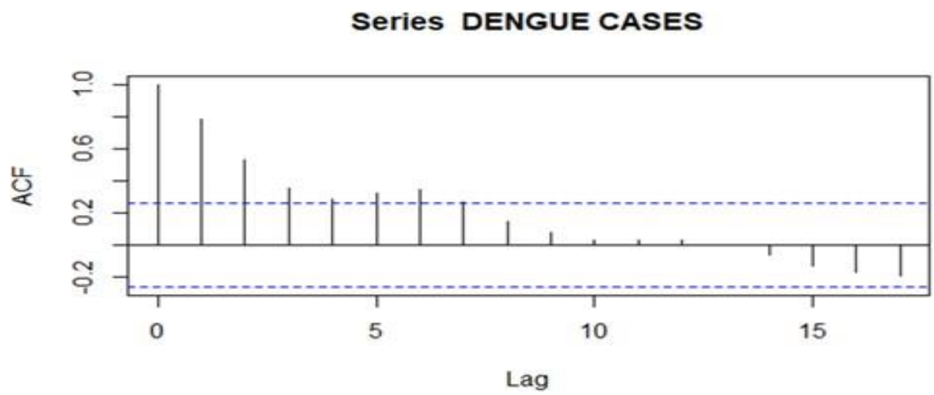

Figure 4. The plot represents the partial autocorrelation function.

The researchers operated the ADF test to verify if the data was stationary, as shown in Table 2. As shown in Region XI, the p-value estimated in RStudio indicates that it exceeds the 95% significance level for the monthly cases. The data must be modified to demonstrate stationary.

Table 2. The generated result of the ADF test

Data:DengueCases	t-Statistic	p-value	Lag order
AugmentedDickey-FullerTest	-0.32253	0.9865	3
Thesignificancelevel is setat:{95%}			

In transforming the data to become stationary, the command *=diff(log)* and *plot.ts* were operated in RStudio; see the graphical representation for Figure 5.

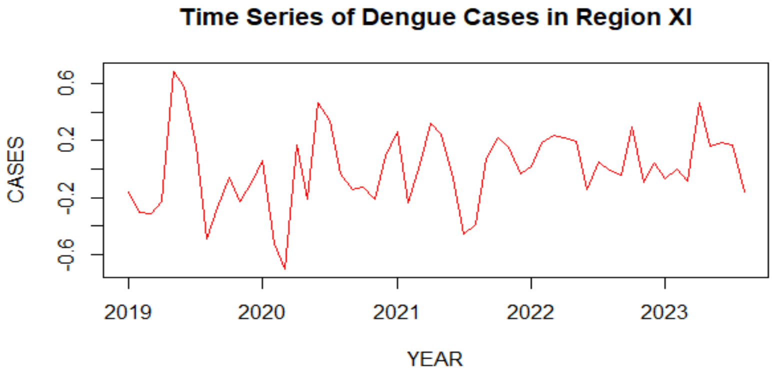

Figure 5. Dengue Case events that occurred in Region XI from (2019 to 2023) show stationarity performed from R Studio.

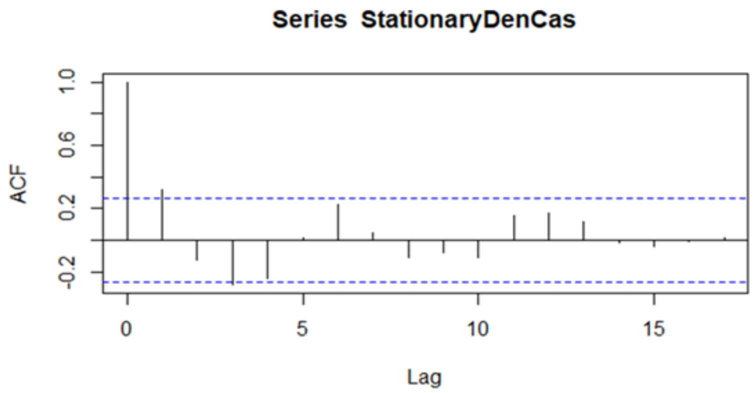

Figure 6. The autocorrelation plot shows stationarity performed from RStudio

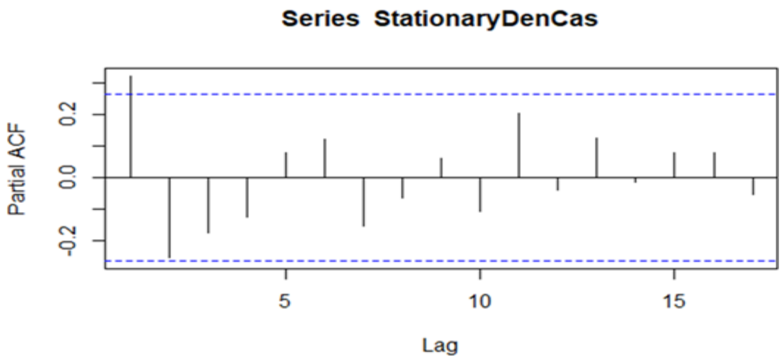

Figure 7. The partial autocorrelation plot indicating stationarity was generated using RStudio.

Model Coefficient Test

In the coefficient test, the models were identified and evaluated using the Akaike Information Criterion (AIC). The AIC served as a measure to gauge the performance of candidate forecasting models in comparing actual data against predictions. The model with the lowest AIC was selected as the optimal forecasting model.

Table 3. The candidate archetypal list, as well as the corresponding AIC figures.

MODEL	AIC
ARIMA(2,0,2)withnon-zeromean	11.54357
ARIMA(0,0,0)withnon-zeromean	17.05817
ARIMA(1,0,0)withnon-zeromean	13.01076
ARIMA(0,0,1)withnon-zeromean	11.23674
ARIMA(0,0,0)withzeromean	15.10013
ARIMA(1,0,1)withnon-zeromean	13.2335
ARIMA(0,0,2)withnon-zeromean	13.22946
ARIMA(1,0,2)withnon-zeromean	13.82953
ARIMA(0,0,1)withzeromean	**9.263068**
ARIMA(1,0,1)withzeromean	11.25948
ARIMA(0,0,2)withzeromean	11.25506
ARIMA(1,0,0)withzeromean	11.03243
ARIMA(1,0,2)withzeromean	11.87688

Among the candidate models obtained from RStudio using the *auto.arima* command, approximations were made based on the list of fitting models. It suggests that the optimal candidate model was ARIMA with a nonzero mean (0,0,1), having an AIC value of 9.263068. The model's projected parameters were also assessed for the definite dengue case instances in Region XI.

Table 4. The results of ARIMA with a nonzero mean (0,0,1) and its quantities.

	ma1	intercept
ARIMA(0,0,1)withanon-zero mean	0.3855	0.0076
Standarderrors	0.118	0.0471

The model coefficients reveal a moving average term (ma1) with a coefficient of 0.3855, signifying the weight assigned to the most recent observation, and an intercept term of 0.0076 representing the baseline value. The standard errors for these coefficients are 0.1180 and 0.0471, respectively, indicating the associated uncertainty. The estimated residual variance (sigma^2) is 0.06421, measuring the variability of residuals. The log-likelihood is -2.62, indicating how well the model explains the data, and the AIC is 11.24, reflecting the model's goodness of fit relative to others. In the context of dengue cases, this suggests that the model captures short-term fluctuations in the data. Soto-Ramírez, Odium, Foxe, Flynn, & Tester (2020) expounded that the equation for an ARIMA (0,0,1) model can be represented as follows:

$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$

Y_t the observed value at time (t),
$\mu$the mean of the time series,
$\varepsilon_t$the error term at time (t),
θ_1 the coefficient of the first lagged error term ε_{t-1}.

Using the equation, we formulated:

$Y_t = 0.0076 + \varepsilon_t + 0.3855\varepsilon_{t-1}$.

This formulation now represents our ARIMA model.

Model Forecasting using ARIMA

From the actual data, Dengue cases tested phenomena were predicted using ARIMA with a nonzero mean (0, 0, 1) using the RStudio, showing a slowly increasing trend from January 2019 up until the forecasted month of December 2024. Also, the researcher examined the normality of the distribution of the forecast errors of the ARIMA model with a nonzero mean (0, 0, 1), exhibiting constant variance and a mean of zero, which can be employed for simulation and predicting the confirmed cases observed in the region during 2019. Figure 7 displays a graph of the actual vs. predicted values of the dengue cases as seen in Region XI.

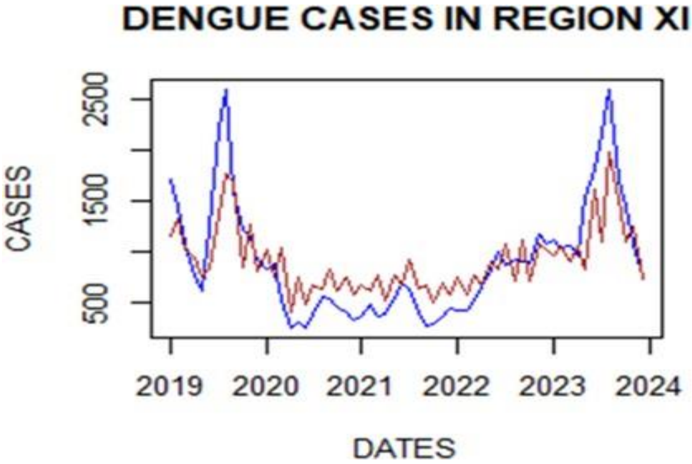

Figure 8. Plotting the Dengue cases in Region XI showing actual values (blue) vs forecasted values (brown).

Figure 9 shows the forecasted data for the next 16 months, from September 2023 to December 2024.

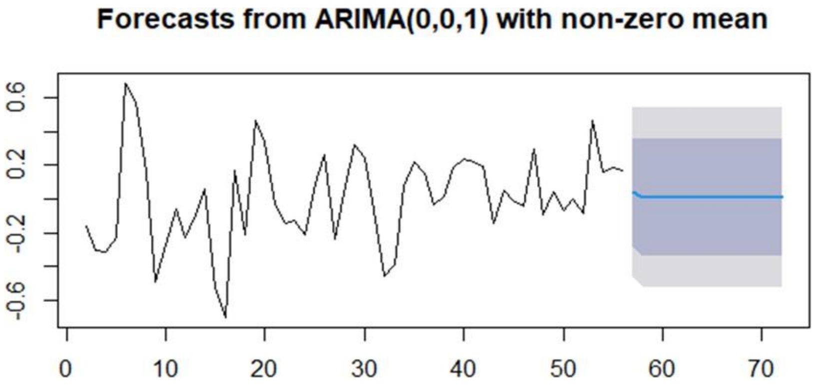

Figure 9. Plot forecast of ARIMA with a nonzero mean (0,0,1).

It predicted that in the next 16 months, the transition of dengue cases will decrease and become constant forecast, and a forecasted decrease in dengue cases is generally seen as a positive sign. It could indicate that preventive measures, public health interventions, or natural factors are contributing to a decline in the spread of the disease. On the other hand, a constant forecast for dengue cases could have important public health implications. The forecasted stability aligns with public health goals it may signal success in control efforts.

Figure 9 shows the Normal Q-Q plot, which is crucial in forecasting dengue cases as it visually assesses the distributional characteristics of historical data, ensuring the reliability of predictive models and facilitating the identification of deviations from normality that may impact the accuracy of forecasts.

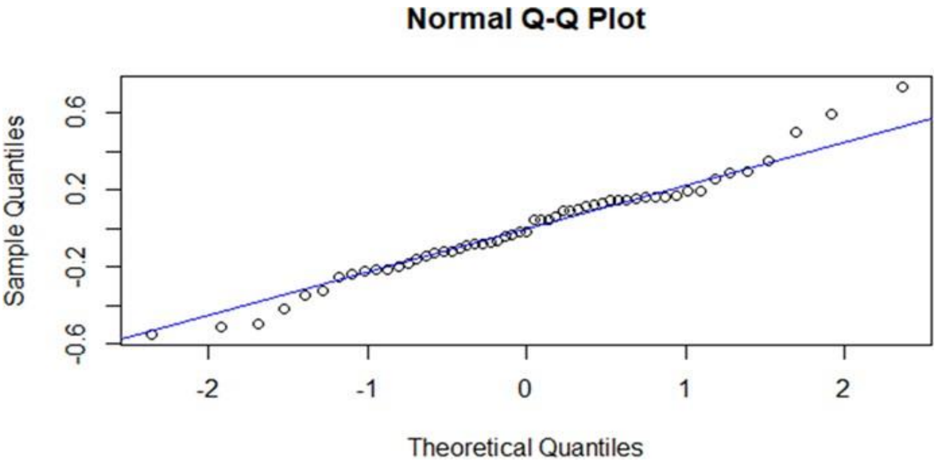

Figure 10. Testing normality (quantile-quantile) plot of residuals using ARIMA with a nonzero mean (0,0,1).

Close to points in a normal quantile-quantile (QQ) plot fall along the expected straight line, which suggests that a normal distribution well approximates the data. In other words, the observed distribution of the dataclosely matches the theoretical distribution of a normal distribution. This alignment indicates that statistical methods and models assuming normality may be appropriate for the dataset, enhancing the reliability of analyses and predictions based on that data.

Table 5. Forecast of Dengue cases in Region XI over the next sixteen months using ARIMA(0,0,1)

FORECASTEDVALUES

PointForecast	Low95	High95
2595.959	1089.4942	4102.423
2595.959	419.4187	4772.499
2595.959	-128.2883	5320.206
2595.959	-619.7807	5811.698
2595.959	-1080.4963	6272.414
2595.959	-1523.5040	6715.421
2595.959	-1956.6590	7148.576
2595.959	-2385.1976	7577.115
2595.959	-2812.8859	8004.803
2595.959	-3242.5992	8434.517
2595.959	-3676.6429	8868.560
2595.959	-4116.9427	9308.860
2595.959	-4565.1646	9757.082
2595.959	-5022.7928	10214.710
2595.959	-5491.1843	10683.102
2595.959	-5971.6063	11163.524

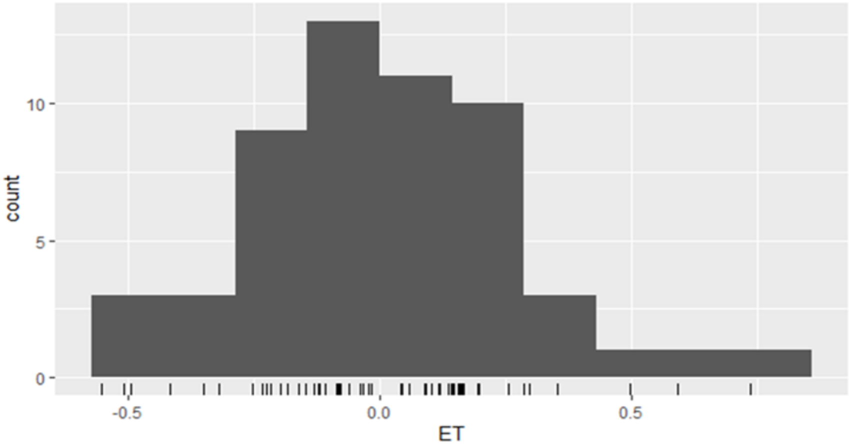

Figure 11. Histogram of the residuals using ARIMA with a nonzero mean (0,0,1)

The result of residuals in forecasting dengue cases suggests that the distribution of errors closely approximates a normal distribution, though there may be slight deviations. The nearly symmetrical and unimodal nature of the distribution, resembling the shape of a bell curve, indicates random and unbiased errors in the forecasting model.

CONCLUSIONS

The study comprehensively examines the dengue situation in Region XI, Philippines, utilizing the ARIMA time series forecasting method. Dengue, a widespread arboviral illness, poses a significant public health challenge globally, with the Philippines, notably Region XI, experiencing a high burden of cases. The findings highlight the increasing trend of dengue cases in Region XI from January 2019 to August 2023, emphasizing the need for effective control measures. The ARIMA model, specifically ARIMA (0,0,1) with a nonzero mean, was employed to forecast dengue cases for the next 16 months (September 2023 to December 2024).

The model identified a slowly increasing trend in dengue cases, but the forecast indicated a potential decrease and stabilization in the coming months. It could be interpreted as a positive sign, suggesting that preventive measures and public health interventions contribute to a decline in the spread of the disease. The study also addresses the issue of model identification, emphasizing the importance of transforming non-stationary data into a stationary form for accurate forecasting. The ADF test and differencing techniques were applied to achieve stationarity in the time series data.

In conclusion, the study underscores the significance of collective efforts in minimizing the transmission of dengue cases. It highlights the potential effectiveness of the ARIMA model in forecasting and addressing the growing problem of dengue in the region.

REFERENCES

1. Abualamah, W. A., Akbar, N. A., Banni, H. S., & Bafail, M. A. (2021). Forecasting the morbidity and mortality of dengue fever in KSA: A time series analysis (2006–2016). Journal of Taibah University Medical Sciences, 16, 448–455.
2. Daele, S. V., & Janssenswillen, G. (2023). Identifying the Steps in an Exploratory Data Analysis: A Process-Oriented Approach. In M. Montali, A. Senderovich, & M. Weidlich (Eds.), Process Mining Workshops. ICPM 2022. Lecture Notes in Business Information Processing, 468. Springer, Cham. https://doi.org/10.1007/978-3-031-27815-0_38
3. Department of Health. (2020). DOH Reports Steady Decline in Dengue Cases. Retrieved on 31 Jul 2021 from https://doh.gov.ph/doh-press-release/doh-reportssteady-decline-in-dengue-cases.
4. Dagoumas, A. (2021). Mathematical Modelling of Contemporary Electricity Markets. Elsevier Academic Press. https://www.everand.com/book/492741322/Mathematical-Modelling-of-Contemporary-Electricity-Markets
5. Dhamodharavadhani, S., & Rathipriya, R. (2020). Forecasting Dengue Incidence Rate in Tamil Nadu Using ARIMA Time Series Model. In Machine Learning for Healthcare (1st ed., pp. 16). Chapman and Hall/CRC. https://doi.org/10.1201/9780429330131
6. Montemayor, M. T (2022, July 6). DOH Logs 51,622 Dengue Cases, 58% in 2021. Govt. Philippines on July 4, 2022. Retrieved from https://reliefweb.int/report/philippines/doh-logs-51622-dengue-cases-58-2021
7. Ong, E. P., Obeles, A. J. T., Ong, B. A. G., & Tantengco, O. A. G. (2022). Perspectives and lessons from the Philippines' decades-long battle with dengue. The Lancet Regional Health Western Pacific, 24, 100505. DOI: 10.1016/j.lanwpc.2022.100505
8. Philippine Statistics Authority. (2021). Highlights of the Philippine Population 2020 Census of

Population and Housing (2020 CPH). Retrieved on 16 Sep 2021 from https://psa.gov.ph/content/highlights-philippine-population-2020-census-population-and-housing-2020-cph

9. Reich, Y., & Subrahmanian, E. (2022). Documenting design research by structured multilevel analysis: Supporting the diversity of the design research community of practice. Design Science, 8, E3. doi:10.1017/dsj.2021.28
10. Singh, S., Sundram, B. M., Rajendran, K., Law, K. B., Aris, T., Ibrahim, H., ... Gil, B. S. (2020). Forecasting monthly confirmed dengue cases in Malaysia using ARIMA models. Journal of Infection in Developing Countries, 14(9), 971-976.
11. Soto-Ramírez, N., Odeku, J., Foxe, C., et al. (2020). Applying time series modeling to assess the dynamics and forecast monthly reports of abuse, neglect, and/or exploitation involving a vulnerable adult. Arch Public Health, 78, 53. https://doi.org/10.1186/s13690-020-00431-0
12. Stolerman, L. M., Maia, P. D., & Kutz, J. N. (2019). Forecasting dengue fever in Brazil: An assessment of climate conditions. PLOS ONE, 14, e0220106. https://doi.org/10.1371/journal.pone.0220106
13. Sutriyawan, A., Herdianti, H., Cakranegara, P. A., Lolan, Y. P., & Sinaga, Y. (2022). Predictive Index Using Receiver Operating Characteristic and Trend Analysis of Dengue Hemorrhagic Fever Incidence. Open Access Macedonian Journal of Medical Sciences, 10, 681–687.
14. Thiruchelvam, L., Dass, S. C., Asirvadam, V. S., et al. (2021). Determine neighboring region spatial effect on dengue cases using ensemble ARIMA models. Scientific Reports, 11, 5873. https://doi.org/10.1038/s41598-021-84176-y
15. World Health Organization. (2023, March 17). Dengue and severe dengue. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue