MALAYSIAN JOURNAL OF SCIENCE
MJS
ISSN 1394 - 3065

FAST-UTHM Global Nexus Conference 2024

Special issue

Malaysian Journal of Science 45 (Special Issue 1): 95 -103 (June 2026)

https://mjs.um.edu.my

Investigation Public Holiday is Factor Affecting Dengue Cases by Using Arima and Arimax Model in Selangor and Sarawak from 2015 to 2019

Nur Dayana Abdul Rahman[1a],Loshini Thiruchelvam[2a*],Nur Balqishanis Zainal Abidin[3b,f],Sarat Chandra Dass[4c],Balvinder Singh Gill[5d],Nur Amalina Mat Jan[6a] and Vijanth Sagayan Asirvadam[7e]

Abstract: This study investigated the relationship between dengue cases and human movement, as represented by public holidays in Malaysia for two states, Selangor and Sarawak, using a weekly dataset from 2015 to 2019. This study developed dengue prediction models by employing Autoregressive Integrated Moving Average (ARIMA) and ARIMA with public holiday as the exogenous variable (ARIMAX) time series models. The Box Jenkins approach for model development with Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) selection criteria are applied to select the optimum model by using the training dataset. This study, however, found the positive effect of human movement during public holidays is only apparent in three study areas; Klang and Sabak Bernam districts in Selangor state and Sibu division in Sarawak. The AIC and BIC values for Klang, Sabak Bernam, and Sibu were (-573.51, -556.89), (-162.93, -146.32), and (-446.64, -420.06), respectively. Despite this, the difference in the AIC and BIC values between ARIMA and ARIMAX models in other study areas is not large, and the study further found that the ARIMAX model of each study area did not fail to perform comparably well, with some study areas showing better accuracy scores when compared to their counterpart ARIMA model in the testing dataset. This concludes that public holidays, specifically the effect of human movement during that period, are significant factors for predicting dengue cases.

Keywords: Box Jenkins, dengue, Malaysia, prediction models, ARIMAX

1. Introduction

Dengue is a viral disease, transmitted by Aedes mosquito vectors and occurs in the tropical and subtropical regions of the world, including our country, Malaysia. In Malaysia, Selangor has been recognized as the most urbanized state, with several well-known districts such as Petaling and Klang, known to have the highest cases of dengue in previous years. Locations that can hold still water, such as containers, water tanks, or pots, are the types of places where mosquitoes can breed in large numbers. There are six stages in the process of dengue infection, as illustrated in Figure 1 below (Thiruchelvam et al., 2017). The metamorphosis of mosquitoes' life cycles consists of eggs that develop into larvae, then pupae, and finally into adult mosquitoes (Rueda, 2008). This process occurs over one to three weeks during the first stage of the dengue infection.

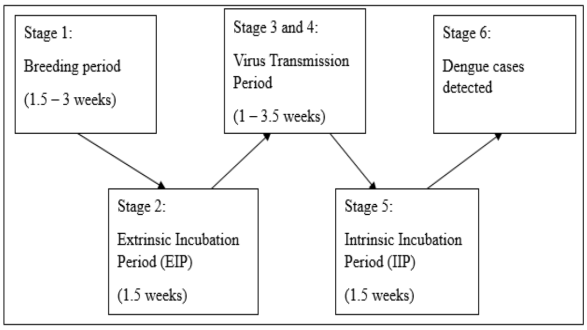

Figure 1. Stages of dengue infection process.

Dengue virus can be transmitted only by female Aedes mosquitoes after they bite infectious humans who have the dengue virus (Pliego Pliego et al., 2017), and it takes on average around two weeks for those female Aedes mosquitoes to activate the dengue virus inside their bodies (Chan & Johansson, 2012; Thiruchelvam et al., 2018). This process occurs in stage two and is known as the Extrinsic Incubation Period (EIP). Stages three and four are the transmission periods of the dengue virus that are transmitted by mosquitoes and the infected human, respectively, and usually occur around one to four weeks. Here, an infected female Aedes mosquito will bite an uninfected human, causing that human to acquire the dengue virus. Meanwhile, in stage four, the infected human in stage three will transmit the dengue virus

Authors information:
[a]Department of Physical and Mathematical Science, Faculty of Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, MALAYSIA. E-mail: dynraz@1utar.my[1], loshini@utar.edu.my[2], amalinaj@utar.edu.my[6],
[b]Faculty of Computing & Informatics (FCI), Multimedia Universiti, Persiaran Multimedia, 63100 Cyberjaya, Selangor. E-mail: nurbalqishanis@mmu.edu.my[3]
[c]School of Mathematical and Computer Sciences, Heriot–Watt University Malaysia, Putrajaya. E-mail: s.dass@hw.ac.uk[4]
[d]National Public Health Laboratory (MKAK), Sungai Buloh. Ministry of Health Malaysia. E-mail: drbsgill@moh.gov.my[5]
[e]Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, 31750 Seri Iskandar, Perak, Malaysia. E-mail: vijanth_sagayan@utp.edu.my[7]
[f]Centre for Advanced Analysics, CoE Artificial Intelligence, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia

*Corresponding Author: loshini@utar.edu.my

Received: May 25, 2025
Accepted: July 30, 2025
Published: June 30, 2026

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

to a new uninfected female mosquito which bites them during a blood meal (Cheong et al., 2013). Thus, there is a greater chance for other healthy humans to be infected too in the future. As for stage five, it is the Intrinsic Incubation Period (IIP), which is the time taken for the infected human to have the dengue virus activated in their body and takes about two weeks to occur (Chan & Johansson, 2012; Thiruchelvam et al., 2018).

In the last stage, the infected human will have onset of symptoms of dengue, such as fever, rashes, and vomiting (Thiruchelvam et al., 2017), after which he or she will seek a doctor's consultation. About a week is needed for the clinical test results to confirm if the person has dengue fever. Thus, on average, it takes about one to two weeks from when the infected person starts having dengue symptoms to clinical confirmation of dengue cases.

Considering the time period between the occurrences of virus transmission and dengue cases being detected clinically to be on average two to three weeks, this study seeks to investigate if there is a possibility of higher dengue cases being detected when there is a public holiday during the two to three weeks before the virus transmission period. Here, public holidays are used as a proxy for human movement, with the hypothesis that when there is a public holiday, more people might travel into the specified district, causing more human-mosquito contact and thus leading to more spread of infectious Aedes mosquito bites.

Statistical approaches using Time Series Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) models are among the most commonly used methods in epidemiological research for predicting infectious diseases, such as dengue (Mahmud et al., 2021). Moreover, ARIMAX is a significant and recognized statistical model that allows exogenous variables to improve the model's forecasting accuracy (Islam et al., 2023). Thus, with the above rising issues, this study aims to determine if travel during public holidays increases the spread of dengue cases within districts in Selangor by using ARIMAX models. Additionally, several divisions from Sarawak state will also be investigated since these two states differ widely in terms of ethnicity.

2. Materials and methods

2.1 Study region

This study focuses on two states in Malaysia: Selangor in Peninsular Malaysia, located in the west of Malaysia, and Sarawak on Borneo Island, located in the east of Malaysia. Based on the Disease Control Division under the Ministry of Health Malaysia (MOH), Malaysian Space Agency (MYSA), and Ministry of Science, Technology and Innovation Malaysia (MOSTI) (2023), Selangor recorded the highest dengue cases with a total of 47,953 cases from December 2019 until October 2023. Another state, Sarawak, is taken as a comparison because it may have differences in culture and travel behaviour compared to Selangor due to significant differences in ethnicity between these states. Selangor consists of Malay and Bumiputera as the largest population, followed by Chinese, Indians, and some minority groups (Selangor Kawasanku, 2023), whereas Sarawak consists of Iban as the largest population, followed by Chinese, Malay, Bidayuh, Melanau, and a number of other local ethnic groups too (Badaruddin et al., 2024).

Figure 2 presents nine districts in Selangor: Gombak, Hulu Langat, Hulu Selangor, Klang, Kuala Langat, Kuala Selangor, Petaling, Sabak Bernam, and Sepang. According to the Selangor Travel (2024) website, the highest population district in Selangor is Petaling, followed by Hulu Langat and Klang. Petaling has the largest population in this state because of its urban centre, modern facilities, healthcare institutions, international colleges, and its public transportation and infrastructure are well developed, making it easy to access any place. Petaling is the most progressive district due to its rapid urbanisation. Hulu Langat's main economic activity is agriculture, and it is popular for its natural surroundings. As for Klang, it was the capital city of Selangor from 1875 until 1880.

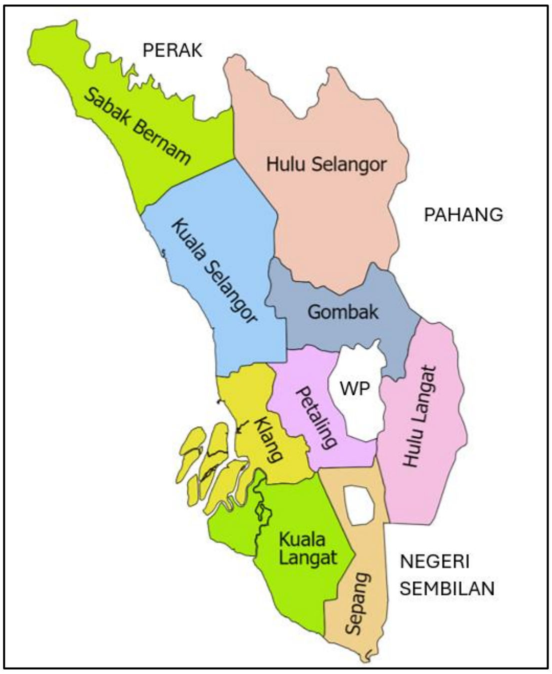

Figure 2. Districts in Selangor.

Table 1 presents the estimated area, population, and population density for nine districts in Selangor. As shown in Table 1, the most densely populated district is Petaling, followed by Klang and Hulu Langat, whereas the least dense district is Hulu Selangor, followed by Kuala Selangor and Kuala Langat.

Meanwhile, Sarawak consists of more than ten divisions, as illustrated in Figure 3. In this study, division is used as the unit of analysis, with only three divisions being examined, which are Kuching, Miri, and Sibu. These three divisions are selected due to the rapid expansion of developing cities along with the increasing population, where most of the forested area is used for urban development (Harvie et al., 2020). Based on the Sarawak Government (2024) website, Kuching division comprises three districts: Kuching, Lundu, and Bau districts, while Miri division includes Miri and Marudi districts. There are three districts in Sibu division: Sibu, Kanowit, and Selangau districts. Table 2 provides the values of estimated areas, population, and population density

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

for these divisions. As shown in Table 2, Kuching is the most densely populated division, whereas Miri and Sibu have relatively lower population densities compared to Kuching.

Table 1: Estimated Area and Population of Selangor District.

District	Estimated Area (km²)	Population	Population Density (people/km²)
Gombak	650	942,400	1,449.85
Hulu Langat	840	1,400,461	1,667.21
Hulu Selangor	1,756	243,029	138.40
Klang	626	1,088,942	1,739.52
Kuala Langat	858	307,449	383.33
Kuala Selangor	1,195	281,711	235.74
Petaling	484	2,298,130	4,748.20
Sabak Bernam	997	107,057	107.38
Sepang	600	325,244	542.07

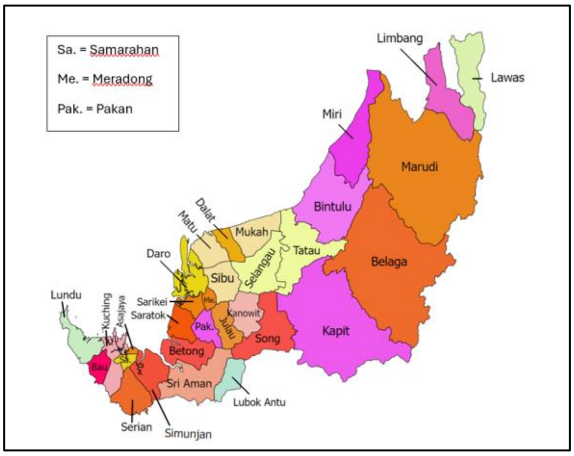

Figure 3. Districts in Sarawak.

Table 2: Estimated Area and Population of Sarawak District.

District	Estimated Area (km²)	Population	Population Density (people/km²)
Kuching	4,560	812,900	178.27
Miri	26,777	433,800	16.20
Sibu	8,278	349,700	42.24

2.2 Public Holidays

Table 3 shows a list of public holidays during the study period from 2015 to 2019 for the Selangor and Sarawak states. There are two types of holidays: national and state holidays. The national holidays include Chinese New Year, school holidays, Eid Fitri, and Eid Adha, while the state holidays are Gawai (orange-colored) and Deepavali (green-colored), celebrated by Sarawak and Selangor states, respectively.

Before including public holiday data in ARIMAX models, the data are transformed into binary variables. For example, the weeks with any public holidays listed in Table 3 are considered as '1' while the weeks that have no holidays are listed as '0'.

2.3 Data Description

Data on dengue cases were obtained from the Ministry of Health (MoH) Malaysia and represent a weekly dataset from 2015 until 2019, spanning a total of 260 weeks. The range and average of dengue cases for each study region are presented in Table 4. The data were then cleansed and normalized before being imported into the R programming language to develop Autoregressive Integrated Moving Average (ARIMA) and ARIMAX models. As this study uses a two-week lag of dengue cases, the data for 2015 started at week three, while other years began at week one. The data were then divided into two parts for training and testing the models.

3. Methodology

$$Y_t' = \phi_1 Y_{t-1}' + \phi_2 Y_{t-2}' \dots + \phi_p Y_{t-p}' + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (1)$$

3.1 ARIMA and ARIMAX models for dengue prediction

$$Y_t' = \sum_{t=1}^{5} \gamma_i A_i + \sum_{l=1}^{n_b} \sum_{i=a}^{m_l} \beta_i X_{t-i}'^{(l)} + \sum_{i=a}^{m_l} \beta_i W_{t-i}^{(l)} + \phi_1 Y_{t-1}' + \phi_2 Y_{t-2}' + \cdots + \phi_p Y_{t-p}' + \varepsilon_t - \theta_1 \varepsilon_{t-1} + -\theta_2 \varepsilon_{t-2} - \cdots -\theta_q \varepsilon_{t-q} \quad (2)$$

Both ARIMA and ARIMAX models can be represented as the following equations:
and

where Y'$_t$ is the stationarized dengue observation at time t. A_i is the population density at time i with $i = 1,2, \dots 5$ since the study uses a five-year timeline. Coefficients of γ_i are the respective values of population density to be estimated. $X_{t-i}'^{(l)}$ refers to the climate parameters to be included. The n_b in the climate parameter refers to the number of climate parameters being included. Variable $W_{t-i}^{(l)}$ is the public holiday binary information, denoted as 1 for presence and 0 for absence of public holidays during time $(t - \mathrm{i})$ and ε_t is the error value of the model at time t. The l and m_l in both climate and public holiday variables refer to lag order in the ARIMAX model and the maximum lag order considered, respectively, with a being the initial incorporated lagged-time.

The vectors $\beta_i = a, \dots, m_l$, $\underline{\phi} = (\phi_1, \phi_2, \dots, \phi_\mathrm{p})$ and $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_\mathrm{q})$ are the unknown parameters to be estimated, with each $\phi_\mathrm{i} \in \Re$ in (1) and (2) representing the correlation for previous dengue cases and $\theta_\mathrm{i} \in \Re$ in (1) and (2) representing the correlation for moving average of previous error terms to be included. The parameters are estimated using the Maximum Likelihood (ML) approach, and residual diagnostics are carried out on the obtained models to ensure the errors are white-noise.

There are two phases in this section, where Phase I involves the development of the ARIMA model and Phase II involves the inclusion of public holiday information, in the form of binary data, into the finalized ARIMA model of Phase I, forming the ARIMAX model. Both phases consist of three stages: model identification, model estimation, and diagnostic check. The detailed steps of these three stages are illustrated in Figures 4 and 5.

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

Table 3: List of Public Holidays in Selangor and Sarawak Based on Year (2015-2019).

List of public holidays in Selangor and Sarawak based on year				
2015	2016	2017	2018	2019
Week 07 Chinese New Year	Week 6 Chinese New Year	Week 4 Chinese New Year	Week 7 Chinese New Year	Week 6 Chinese New Year
Week 11 School holidays (1st term)	Week 11 School holidays (1st term)	Week 12 School holidays (1st term)	Week 12 School holidays (1st term)	Week 13 School holidays (1st term)
Week 22 Mid-year school holidays and Gawai	Week 22 Mid-year school holidays and Gawai	Week 22 Mid-year school holidays and Gawai	Week 22 Gawai	Week 22 Mid-year school holiday, Gawai and Eidul Fitri
Week 23 Mid-year school holidays	Week 23 Mid-year school holidays	Week 23 Mid-year school holidays	Week 24 Mid-year school holidays and Eidul Fitri	Week 23 Mid-year school holidays
Week 28 Eidul Fitri	Week 27 Eidul Fitri	Week 26 Eidul Fitri	Week 25 Mid-year school holidays	Week 33 School holidays (2nd term) and Eid Adha
Week 38 School holidays (2nd term) and Eid Adha	Week 37 School holidays (2nd Term) and Eid Adha	Week 35 School holidays (2nd Term) and Eid Adha	Week 34 School holidays (2nd Term) and Eid Adha	Week 44 Deepavali
Week 45 Deepavali	Week 44 Deepavali	Week 42 Deepavali	Week 45 Deepavali	Week 48-52 School holidays (last term)
Week 47-52 School holidays (last term)	Week 48-52 School holidays (last term)	Week 48-52 School holidays (last term)	Week 48-52 School holidays (last term)	

Table 4: Range of Dengue Cases for Selangor Districts and Sarawak Division.

District/Division	Range of dengue cases (2015-2019)	Average dengue cases for five years (2015-2019)
Gombak	[34 – 349]	135
Hulu Langat	[80 – 527]	243
Hulu Selangor	[03 – 77]	22
Klang	[70 – 615]	188
Kuala Langat	[04 – 67]	25
Kuala Selangor	[01 – 54]	16
Petaling	[116 – 1025]	391
Sabak Bernam	[01 – 19]	5
Sepang	[10 – 123]	41
Kuching	[01 – 19]	6
Miri	[10 – 123]	3
Sibu	[01 – 40]	11

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

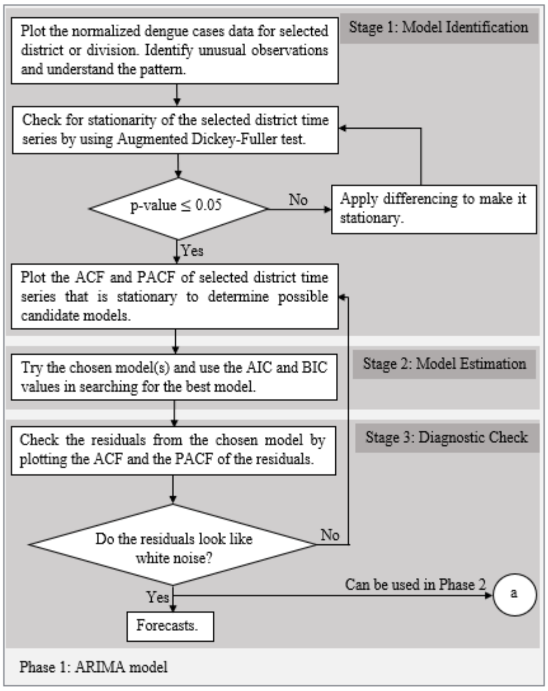

Figure 4. Flowchart of ARIMA model

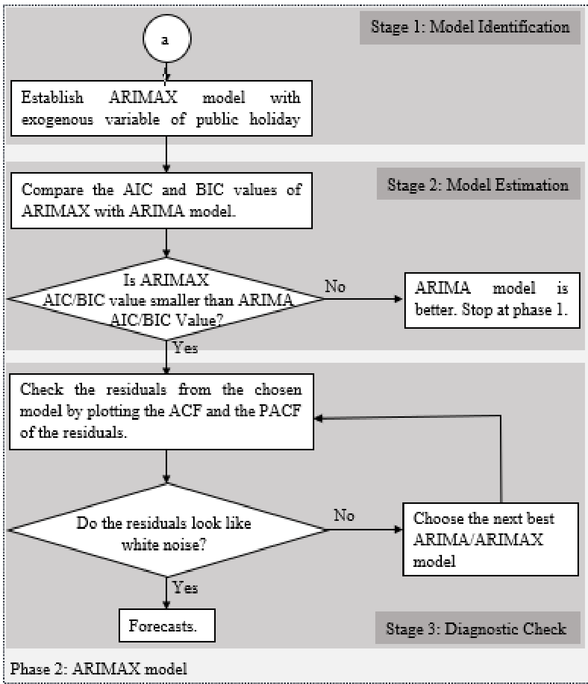

Figure 5. Flowchart of ARIMAX model.

3.2 Measurement Tools

This study employed three measurement tools to select the optimal model: the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Root Mean Square Error (RMSE).

The RMSE formulation with n is as the sample size is formulated as follows:

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(Predicted_i - Actual_i)^2}{n}} \quad (3)$$

4. Results

4.1 ARIMA Model Development

Sabak Bernam district is used here as an example to illustrate the development of a dengue prediction model, as can be seen in Figures 6-8. Figure 6 shows a time series of the weekly dengue cases for Sabak Bernam from the year 2015 until 2018. Although there are some high case numbers, it is not seen as observable peaks considering that the time series were within the 0-1 range of normalization. To confirm further, the study looks at the plots of Autocorrelation (ACF) and partial autocorrelation (PACF). Since the spikes managed to decay and fall within the confidence interval bands, it can be concluded that the time series is stationary. This is further confirmed through the Augmented Dickey-Fuller test, which found a p-value of 0.01, rejecting the null hypothesis and confirming that the time series is stationary. Anything within or close to the blue line is statistically close to zero and anything outside the blue line is statistically non-zero (Yasmin & Moniruzzaman, 2024).

Based on Figure 6, there are two significant spikes (at lag one and two) observed in the PACF plot. Thus, the candidate models would be ARIMA (1,0,0) and ARIMA (2,0,0). In choosing the best model fit under this condition, the pair of AIC and BIC ensures avoiding overfitting and underfitting of the model simultaneously. Next, RMSE is used to select the model that gives the best prediction accuracy for the testing data. Thus, applying AIC, BIC, and RMSE ensures selection of a model that has a good fit together with accurate prediction.

Table 5 shows the candidate ARIMA models for Sabak Bernam. Based on Table 5, ARIMA (2,0,0) is selected as the model with the best fit and also having the highest prediction accuracy as the pair of AIC and BIC while the RMSE value of training data is among the lowest. Figure 8 shows that the lags of ACF and PACF for the selected best ARIMA model are within the threshold line, which indicates the residuals are white noise. Hence, the ARIMA(2,0,0) model is the model set for predicting dengue cases in Sabak Bernam district. Figure 9 shows the plotted Sabak Bernam data with forecasted ARIMA models. In choosing the best model fit, two aspects are given importance: first, to ensure the selected model is not too complex, that is, accepting all the proposed variables/parameters into the model and causing an overfitting issue, in which the model tends to learn the current dataset too much. This will cause the selected model not to perform well in a new dataset. This can be addressed through BIC, which tends to select simpler models and avoid overfitting. On the other hand, problems with selecting a too simple model without sufficient information provided through the inclusion of parameters mean that the model cannot display the actual or existing relationship, and thus is unable to perform well with the available dataset itself. This situation can be overcome by introducing the AIC, which safeguards against underfitting error.

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

In summary, applying AIC and BIC model selection criteria simultaneously will ensure selecting the best model fit, with both overfitting and underfitting issues being taken care of. Next, RMSE is used on the testing dataset (for data of the year 2019) to select the model which gives the best prediction accuracy. Thus, applying AIC, BIC, and RMSE ensures selection of a model that has good fit together with accurate prediction.

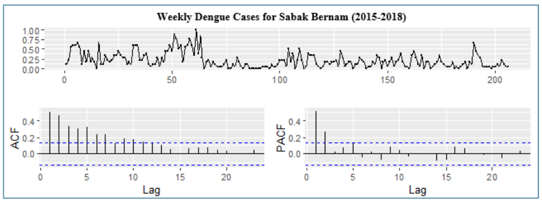

Figure 6. Time series of Sabak Bernam district from the year 2015 until 2018.

```
        Augmented Dickey-Fuller Test

data:  trainsb
Dickey-Fuller = -4.0569, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Figure 7. Result of Augmented Dickey-Fuller test for Sabak Bernam.

Table 5: Parameters for Candidate ARIMA Models in Sabak Bernam.

Candidate Model	AIC value	BIC Value	RMSE
ARIMA(1,0,0)	-141.0769	-131.0933	0.1674164
ARIMA(2,0,0)	-154.76	-141.4485	0.1696358

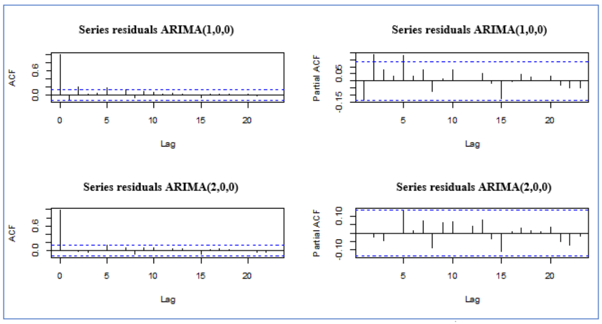

Figure 8. The residuals for each ARIMA model.

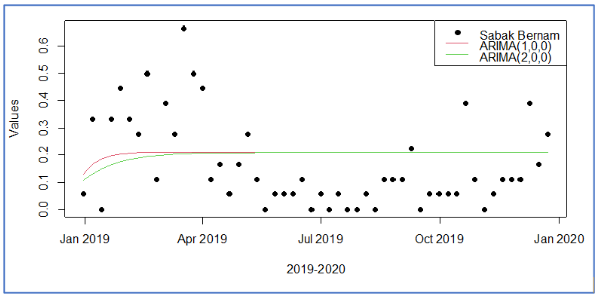

Figure 9. Plotted Sabak Bernam data with forecasted ARIMA models.

4.2 ARIMAX Model Development

The ARIMAX model is an extension of the ARIMA model where X represents the exogenous variable, which in this study is the public holidays, measured on a weekly basis. PH2, PH3, and PH2,3 represent 2-week lagged, 3-week lagged, and 2- and 3-week lagged variables, respectively. Sabak Bernam district is used again to illustrate the ARIMAX model development.

Based on Table 6, it can be seen that ARIMAX(2,0,0,PH2) is the best model as it has the smallest AIC and BIC values compared to other models. After selecting the most fitting and accurate model, a diagnostic check is carried out on the selected model, as shown in Figure 10. The residuals of ARIMAX(2,0,0,PH_2) are found to be white noise. Thus, the model can be finalized.

Figure 11 shows the forecasted values and range of the ARIMAX(2,0,0,PH_2) model for 52 weeks. The model is seen to forecast dengue cases to some acceptable extent. Prior Table 7 highlighted in orange in Table 7, and the bold models show the best model with the lowest AIC and BIC values for each district and division's training dataset. Although only three study areas showed positive results with the inclusion of public holiday variables in the training dataset, the ARIMAX models of each study area did not fail to perform comparably well, with some study areas shown to have better accuracy compared to their counterpart ARIMA model in the testing dataset. This indicates that public holiday, specifically the effect of human movement during that period, is a meaningful factor for predicting dengue cases.

Table 6. Parameters for ARIMAX models in Sabak Bernam.

Candidate Model	AIC value	BIC Value	RMSE
ARIMAX(1,0,0,PH_2)	-152.28	-138.97	0.1806147
ARIMAX(2,0,0,PH_2)	-163.10	-146.46	0.1894346
ARIMAX(1,0,0,PH_3)	-140.94	-127.63	0.1618569
ARIMAX(2,0,0,PH_3)	-153.12	-136.48	0.1674543
ARIMAX(1,0,0,$PH_{2,3}$)	-141.86	-128.54	0.1711214
ARIMAX(2,0,0,$PH_{2,3}$)	-155.51	-138.87	0.1782661

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

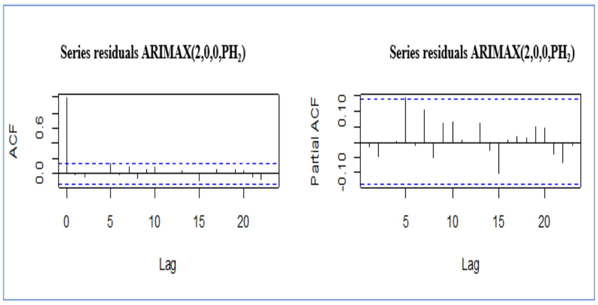

Figure 10. Residuals for Sabak Bernam ARIMAX(2,0,0,PH$_2$) model.

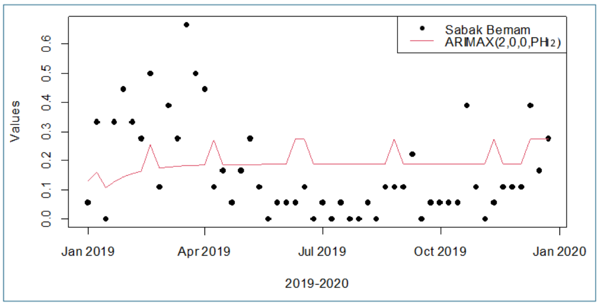

Figure 11. Plotted Sabak Bernam with forecasts of ARIMAX(2,0,0,PH$_2$) model

5. Discussion

Klang and Sabak Bernam districts in Selangor and Sibu division in Sarawak demonstrate better prediction as ARIMAX models in their training dataset, concluding that incorporating public holiday information provides additional data and is not considered redundant in dengue prediction for those study areas. Besides, all the ARIMAX models performed comparably well, with some study areas showing better accuracy, measured through the RMSE values, compared to their counterpart ARIMA models in testing data, thus indicating that ARIMAX models can make accurate predictions as well. The obtained results explain the significant impact of human movements and contacts during public holidays as a major contributing factor to the transmission and spread of dengue in both these study areas. The surplus of human movement, especially during public holidays, was observed in those study areas.

These findings align with prior studies (Thiruchelvam et al., 2021; Ramadona et al., 2023; Syukri & Wardiah, 2023), which show the pattern of dengue cases in a study area is also influenced by the pattern of dengue in its neighboring regions. This is believed to occur due to human movement involving dengue-infected people who might or might not have symptoms yet (Shahid et al., 2020; Hossain et al., 2020). Since once a mosquito bites the infected person, there is always a chance for the virus to be transmitted to other people, spreading the virus to the community through several cycles of mosquito bites (Hossain, 2022). Thus, population movement is considered a possible factor in dengue transmission. Additionally, a study by Siriyasatien et al.

Table 7. AIC, BIC and RMSE Values for All Districts

Division	Model	AIC value	BIC Value	RMSE
Gombak	**ARIMA(1,0,2)**	**-427.22**	**-410.58**	0.3254
	ARIMAX(1,0,2,PH$_3$)	-426.04	-406.08	**0.2716**
Hulu Langat	**ARIMA(2,0,2)**	**-342.84**	**-322.87**	0.2176
	ARIMAX(2,0,2,PH$_2$)	-341.17	-317.87	**0.1828**
Hulu Selangor	**ARIMA(4,0,2)**	**-337.46**	**-310.84**	**0.2537**
	ARIMAX(4,0,2,PH$_2$)	-336.13	-306.18	0.2983
Klang	ARIMA(2,1,2)	-573.73	-557.12	0.3477
	ARIMAX(2,1,2,PH$_2$)	**-574.07**	**-554.13**	**0.3271**
Kuala Langat	**ARIMA(2,1,0)**	**-255.18**	**-245.21**	**0.1707**
	ARIMAX(2,1,0,PH$_2$)	-253.62	-240.33	0.1780
Kuala Selangor	**ARIMA(4,1,2)**	**-319.6**	**-296.34**	**0.1959**
	ARIMAX(4,1,2,PH$_3$)	-319.16	-292.58	0.2039
Petaling	**ARIMA(2,1,3)**	**-532.94**	**-512.99**	**0.1374**
	ARIMAX(2,1,3,PH$_2$)	-531.83	-508.57	0.1452
Sabak Bernam	ARIMA(2,0,0)	-154.76	-141.45	**0.1696**
	ARIMAX(2,0,0,PH$_2$)	**-163.10**	**-146.46**	0.1894
Sepang	**ARIMA(0,1,1)**	**-248.89**	**-242.24**	**0.1732**
	ARIMAX(0,1,1,PH$_{2,3}$)	-246.89	-236.92	0.1734
Kuching	**ARIMA(0,1,1)**	**-542.26**	**-535.62**	**0.3475**
	ARIMAX(0,1,1,PH$_3$)	-540.48	-530.51	0.3491
Miri	**ARIMA(2,0,1)**	**-842.74**	**-826.1**	**0.3786**
	ARIMAX(2,0,1,PH$_3$)	-840.84	-820.88	0.3860
Sibu	ARIMA(5,1,2)	-445.37	-418.78	**0.0737**
	ARIMAX(4,1,2,PH$_2$)	**-446.64**	**-420.06**	0.0771

(2018) mentioned tourism as one of the reasons for people to travel, besides traveling for work or trade.

Our findings, especially for the rural district of Sabak Bernam, align with several previous studies conducted in countries such as Bangladesh, Singapore, and China (Aguiar et al., 2014; Hossain, 2022; Lin et al., 2016; Seah et al., 2021; Shahid et al., 2020; Sippy et al., 2019; Wu et al., 2018). For example, a study by Hossain (2022) concluded that the movement of migrants residing in the

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

capital city, Dhaka, to their hometowns during Eid al-Adha caused peak infections and new outbreaks in those cities. This conclusion was supported by a study by Hossain et al. (2020), which found that after a few weeks of the Eid holiday, most admitted dengue patients were reported from other cities, rather than the capital, Dhaka, which is the common case. This finding also aligns with another study by Rafi et al. (2020), which found that 40% of dengue patients in the northwestern city had a recent history of traveling to Dhaka. A study (Hay, 2013) highlighted that dengue cases increase during any holiday season or carnival, as the city's population is denser at that time, thereby increasing the chances of dengue infection.

6. Conclusion

This study was conducted to investigate the best models for predicting dengue cases in the districts and divisions of Selangor and Sarawak, in Malaysia, respectively, using ARIMA and ARIMAX models. This study has identified the ARIMAX model as a better model in two study districts, Klang and Sabak Bernam, in Selangor state, and also in Sibu division, in Sarawak, with the AIC and BIC values of (-573.51, -556.89), (-162.93, -146.32) and (-446.64, -420.06), respectively, and with the rest of the districts still having comparably good ARIMAX models too. Nonetheless, the difference between AIC and BIC values of ARIMA and ARIMAX models in other study areas is not that large. Furthermore, this study showed that the ARIMAX model of each study area performs comparably well, with some study areas showing better accuracy results when compared to their corresponding ARIMA model in the testing dataset. It can be hypothesized that dengue cases are influenced by human movement from the capital city to semi-urban and rural areas during public holidays in Malaysia.

Additionally, forecasting is a crucial early warning mechanism that can facilitate proactive planning and response for clinical and public health services. Based on these findings, recommendations can be addressed to the government, especially the Ministry of Health (MoH), so that they can take timely action in preventing and spreading awareness about dengue through media such as Twitter, radio, and television, especially whenever cases increase, for example reminding the public of steps such as wearing long clothes, applying mosquito repellents, and avoiding being outdoors early in the morning and evening before dusk, which are known as the peak biting periods of Aedes mosquitoes. Moreover, health services, especially government health services, can prepare themselves to combat dengue.

7. References

Aguiar, M., Rocha, F., Pessanha, J. E. M., Mateus, L., & Stollenwerk, N. (2014). Carnival or football, is there a real risk for acquiring dengue fever in Brazil during holidays seasons? Scientific Reports, 5. https://doi.org/10.1038/srep08462

Chan, M., & Johansson, M. A. (2012). The Incubation Periods of Dengue Viruses. PLoS ONE, 7(11). https://doi.org/10.1371/journal.pone.0050972

Cheong, Y. L., Burkart, K., Leitão, P. J., & Lakes, T. (2013). Assessing weather effects on dengue disease in Malaysia. International Journal of Environmental Research and Public Health, 10(12), 6319–6334. https://doi.org/10.3390/ijerph10126319

Harvie, S., Nor Aliza, A. R., Lela, S., & Razitasham, S. (2020). Detection of dengue virus serotype 2 (DENV-2) in population of Aedes mosquitoes from Sibu and Miri divisions of Sarawak using reverse transcription polymerase chain reaction (RT-PCR) and semi-nested PCR. In Tropical Biomedicine (Vol. 37, Issue 2).

Hossain, M. S. (2022). Megacity-centric mass mobility during Eid holidays: a unique concern for infectious disease transmission in Bangladesh. In Tropical Medicine and Health (Vol. 50, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s41182-022-00417-4

Hossain, M. S., Siddiqee, M. H., Siddiqi, U. R., Raheem, E., Akter, R., & Hu, W. (2020). Dengue in a crowded megacity: Lessons learnt from 2019 outbreak in Dhaka, Bangladesh. PLoS Neglected Tropical Diseases, 14(8), 1–5. https://doi.org/10.1371/journal.pntd.0008349

Islam, M. A., Hasan, M. N., Tiwari, A., Raju, M. A. W., Jannat, F., Sangkham, S., Shammas, M. I., Sharma, P., Bhattacharya, P., & Kumar, M. (2023). Correlation of Dengue and Meteorological Factors in Bangladesh: A Public Health Concern. International Journal of Environmental Research and Public Health, 20(6). https://doi.org/10.3390/ijerph20065152

Lin, H., Liu, T., Song, T., Lin, L., Xiao, J., Lin, J., He, J., Zhong, H., Hu, W., Deng, A., Peng, Z., Ma, W., & Zhang, Y. (2016). Community Involvement in Dengue Outbreak Control: An Integrated Rigorous Intervention Strategy. PLoS Neglected Tropical Diseases, 10(8). https://doi.org/10.1371/journal.pntd.0004919

Mahmud, N., Syuhada Muhammat Pazil, N., Jamaluddin, H., Aqilah Ali, N., Teknologi MARA, U., & Perlis Kampus Arau, C. (2021). Prediction of Dengue Outbreak: A Comparison Between ARIMA and Holt-Winters Methods ARTICLE HISTORY ABSTRACT. ESTEEM Academic Journal, 17, 101–111.

MOH, D. C. D. M. of H., MYSA, M. S. A., & MOSTI, M. of S. T. and I. (2023). Dengue Announcement by State in Malaysia for the year 2023. https://idengue.mysa.gov.my/ide_v3/index.php

Pliego Pliego, E., Velázquez-Castro, J., & Fraguela Collar, A. (2017). Seasonality on the life cycle of Aedes aegypti mosquito and its statistical relation with dengue outbreaks. Applied Mathematical Modelling, 50, 484–496. https://doi.org/10.1016/j.apm.2017.06.003

Rafi, A., Mousumi, A. N., Ahmed, R., Chowdhury, R. H., Wadood, A., & Hossain, G. (2020). Dengue epidemic in a non-endemic zone of Bangladesh: Clinical and laboratory profiles of patients. PLoS Neglected Tropical Diseases, 14(10), 1–14. https://doi.org/10.1371/journal.pntd.0008567

Rueda, L. M. (2008). Global diversity of mosquitoes (Insecta: Diptera: Culicidae) in freshwater. In Hydrobiologia (Vol. 595, Issue 1, pp. 477–487). https://doi.org/10.1007/s10750-007-9037-x

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)

Sarawak Government. (n.d.). Retrieved February 15, 2024, from https://sarawak.gov.my/web/home/article_view/240/175/?id=158#175

Seah, A., Aik, J., Ng, L. C., & Tam, C. C. (2021). The effects of maximum ambient temperature and heatwaves on dengue infections in the tropical city-state of Singapore – A time series analysis. Science of the Total Environment, 775. https://doi.org/10.1016/j.scitotenv.2021.145117

Selangor.travel | Your Smart Online Travel Guide In Selangor. (n.d.). Retrieved January 24, 2024, from https://selangor.travel/

Shahid, F., Ony, S. H., Albi, T. R., Chellappan, S., Vashistha, A., & Alim Al Islam, A. B. M. (2020). Learning from Tweets: Opportunities and Challenges to Inform Policy Making during Dengue Epidemic. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW1). https://doi.org/10.1145/3392875

Sippy, R., Herrera, D., Gaus, D., Gangnon, R. E., Patz, J. A., & Osorio, J. E. (2019). Seasonal patterns of dengue fever in rural Ecuador: 2009-2016. PLoS Neglected Tropical Diseases, 13(5). https://doi.org/10.1371/journal.pntd.0007360

Siriyasatien, P., Chadsuthi, S., Jampachaisri, K., & Kesorn, K. (2018). Dengue epidemics prediction: A survey of the state-of-the-art based on data science processes. In IEEE Access (Vol. 6, pp. 53757–53795). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2018.2871241

Thiruchelvam, L., Asirvadam, V. S., Dass, S. C., Daud, H., & Gill, B. S. (2017). K-Step Ahead Prediction Models for Dengue Occurrences.

Thiruchelvam, L., Dass, S. C., Asirvadam, V. S., Daud, H., & Gill, B. S. (2021). Determine neighboring region spatial effect on dengue cases using ensemble ARIMA models. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-84176-y

Thiruchelvam, L., Dass, S. C., Zaki, R., Yahya, A., & Asirvadam, V. S. (2018). Correlation analysis of air pollutant index levels and dengue cases across five different zones in Selangor, Malaysia. Geospatial Health, 13(1), 102–109. https://doi.org/10.4081/gh.2018.613

Wu, X., Lang, L., Ma, W., Song, T., Kang, M., He, J., Zhang, Y., Lu, L., Lin, H., & Ling, L. (2018). Non-linear effects of mean temperature and relative humidity on dengue incidence in Guangzhou, China. Science of the Total Environment, 628–629, 766–771. https://doi.org/10.1016/j.scitotenv.2018.02.136

Yasmin, S., & Moniruzzaman, Md. (2024). Forecasting of area, production, and yield of jute in Bangladesh using Box-Jenkins ARIMA model. Journal of Agriculture and Food Research, 16, 101203. https://doi.org/10.1016/j.jafr.2024.101203

DOI: https://doi.org/10.22452/mjs.vol45sp1.12
Malaysian Journal of Science 45 (Special Issue 1): 95-103 (June 2026)