Short-Term Forecasting of Weekly Dengue Cases in Singapore Using ARIMA and Fourier-ARIMA Models

Sihan Chen *

School of Mathematical Science, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

* Corresponding Author Email: juicydiscount@outlook.com

Abstract. In this paper, a clear case-only forecasting model is developed on weekly basis. Dengue in Singapore based on official surveillance data between 2012-W01 and 2026.W17. The paper questions whether the number of cases routinely reported can be useful on their own. Short-term warning information in cases when external covariates are not available or are delayed. Dengue Fever and Dengue Haemorrhagic Fever cases were aggregated on a weekly basis, converted with log(1 + cases), tested to be stationary, and modelled using Autoregressive Integrated Moving Average (ARIMA). And Fourier-ARIMA specifications. Candidate models were compared through residual diagnostics, information criteria, and recursive one-week-ahead validation on the final 10% of the series. ARIMA(4,1,0)+Fourier(2) was selected because it balanced point accuracy, residual adequacy, and parsimony. In the 75-week test period, the model achieved MAE of 11.6 cases and RMSE of 16.0 cases, while 90% prediction interval coverage was 80%. The framework is therefore potentially useful as interpretable and timely surveillance decision support for early warning and preparedness, but it should be combined with field intelligence and future calibration work.

Keywords: Dengue, Singapore, ARIMA, Fourier-ARIMA, Public health surveillance.

1. Introduction

1.1. Research Background

Dengue is a mosquito-borne viral disease transmitted through Aedes mosquitoes. Its transmission is affected by environmental and public health conditions, including temperature, rainfall, mosquito activity, and vector-control measures [1]. In Singapore, dengue forecasting has been studied as a public health tool because early warning can support vector-control planning and the use of limited resources [1, 2].

For this paper, dengue forecasting is treated as a surveillance problem. The aim is not to explain every biological or environmental driver of transmission. The aim is to examine whether routine weekly dengue case notifications can provide a useful short-term warning signal. Earnest et al. showed that Singapore dengue notifications can be used in statistical surveillance forecasting, which supports the value of official case data as a forecasting input [3].

1.2. Existing Problem

Previous studies have used different forecasting approaches for dengue. Earnest et al. compared statistical models for Singapore dengue notifications [3]. Hii et al. developed a weather-based model using temperature and rainfall to forecast dengue incidence in Singapore [1]. Shi et al. developed real-time dengue forecast models for Singapore to support outbreak alerts and policy decisions [2]. These studies show that dengue forecasting can support surveillance and public health action.

However, two practical problems remain. First, many models depend on weather or other external predictors, which may not always be available at the same time as routine case notifications. Second, validation and reporting are not always sufficiently clear. Leung et al. found that some dengue outbreak prediction studies lacked validation or did not report methods and performance adequately [4]. These problems make it necessary to establish what can be achieved from official case counts alone before adding more complex inputs.

1.3. Research Objectives and Questions

This paper will have three objectives: to test the hypothesis that weekly case counts can be used to support one-week-ahead forecasting, to compare ARIMA and Fourier-ARIMA, and to interpret the result in terms of public health use.

The questions that are related are whether the next week is informed by recent cases, whether lightweight annual terms can be used to improve the ARIMA baseline and whether a transparent case-only model can be used to support early warning.

These questions relate statistical performance to operational interpretability as opposed to considering forecast error as the sole measure.

The remainder of the paper is structured around this principle: the Literature Review establishes the gap, the Methodology establishes the pipeline, the Results compare candidate models and the Discussion and Conclusion interpret public health value and limitations.

2. Literature Review

2.1. Epidemiological Background and Transmission Mechanism of Dengue

Aedes mosquitoes spread dengue and the environmental conditions influence its activity. In a Singapore dengue forecasting model, Hii et al. applied temperature and rainfall to support the importance of weather conditions in predicting dengue [1]. They also framed forecasting as a method of giving an earlier warning to control dengue.

Although transmission is complex, its effects are reflected in case surveillance data. Reported cases are not only administrative counts. They can also act as an indirect signal of changes in transmission intensity. Earnest et al. used Singapore dengue fever notifications to compare statistical forecasting models, showing that notification data can support surveillance forecasting [3].

2.2. Development of Infectious Disease Forecasting

Traditional epidemiological models, such as SIR and SEIR, describe transmission by modelling movement between susceptible, infected, exposed, and recovered groups. These models have strong explanatory value, but they require assumptions about transmission rates, recovery, contact patterns, and population structure.

Statistical forecasting methods follow a different approach. ARIMA, Seasonal Autoregressive Integrated Moving Average (SARIMA), and exponential smoothing use the time structure in past case data to forecast future observations. Earnest et al. compared statistical surveillance models for Singapore dengue notifications [3]. Hii et al. included autoregressive and seasonal components in a Singapore dengue model [1]. Dayama and Kameshwaran modelled weekly notified dengue incidence as an overall case-count series and showed that univariate time-series models can support short-term prediction and outbreak-severity classification [5].

2.3. Public Health Surveillance Setting

In public health practice, forecasting is not used only to reduce statistical error. It supports surveillance and risk management. Hii et al. stated that dengue forecasts can inform the timing and scale of vector-control operations and the use of limited resources [1]. Shi et al. framed real-time dengue forecasting in Singapore as an early-warning system for outbreak alerts and policy decision support [2].

A useful surveillance forecast should therefore be timely, interpretable, and easy to update. Public health users need to understand why a model is preferred, how it was validated, and how uncertain its forecast is. Transparent statistical forecasting remains useful when these requirements are as important as predictive accuracy.

2.4. Remaining Gaps

Existing studies show that dengue forecasting can support surveillance, but important gaps remain. Leung et al. reviewed 99 dengue outbreak prediction models from 64 studies and found that climate predictors were widely used, while validation and reporting were often incomplete [4]. Fang et al. similarly reviewed 23 machine-learning dengue studies and reported that 22 studies used meteorological predictors, whereas only 14 used historical dengue data [6]. This evidence shows that external covariates dominate much of the literature, leaving a practical need to test what can be achieved from routinely reported case counts alone.

Another concern is validation and reporting. An accuracy value alone cannot be used to judge a forecast; the model-development path, residual behavior, out-of-sample evaluation and uncertainty should also be reported. This paper thus employs explicit candidate rules and recursive out-of-sample validation as opposed to one fixed fit.

Machine learning, deep learning, and more general data streams are becoming more common in recent dengue forecasting studies. Machine-learning methods to predict temporal and spatiotemporal arbovirus were reviewed by Lima et al. , whereas Sylvestre et al. demonstrated that real-world data and machine-learning methods could be used to enhance dengue prediction and surveillance [7, 8]. Tian et al. applied meteorological predictors to a machine-learning comparison in Singapore, Zhao et al. suggested a deep-learning hybrid architecture to predict weekly dengue outbreaks in Singapore, and Finch et al. demonstrated that the combination of climate and serotype information can lead to better outbreak prediction in Singapore [9-11]. These papers show the worth of richer predictors and complex models. Nevertheless, Frasca et al. observed that lack of interpretability is one of the obstacles to the implementation of complex AI systems in health decision-making [12]. This is important since surveillance users must know model behavior, limitations and uncertainty prior to acting on forecasts.

Combined, the literature justifies a narrow need: an open case-only forecasting system with explicit model-selection criteria, time-based validation, and transparent public health interpretation. This paper fills this gap by presenting a reproducible workflow that does not use any external covariates and ensures that model screening, residual checks, and rolling validation are entirely traceable. No machine-learning models are considered here; they are mentioned solely to clarify why interpretability and reproducibility are still pertinent surveillance criteria.

3. Methodology

3.1. Study Design

This research addresses the issues identified in Chapter 2 by taking a case-only statistical forecasting approach to weekly dengue surveillance in Singapore. The framework prioritizes notifications that are readily available, validation and interpretability.

The framework predicts the number of total reported cases of dengue one epidemiological week in advance to monitor and prepare in the short term. It models temporal dependence on observed case counts, as opposed to individual infection risk, biological transmission parameters or intervention effects. Weather, mosquito density, serotype circulation, population movement and intervention activity are thus not included in the case-only specification.

The analysis is based on one pipeline: data collection, cleaning, target construction, transformation, stationarity assessment, model identification, candidate comparison and rolling validation. This chain maintains the modelling decisions to be traceable and independent of the performance results.

3.2. Data Source and Study Object

The data includes official weekly dengue surveillance data in Singapore between 2012-W01 and 2026-W17. The cleaned series has 748 epidemiological weeks. The 2026 data are incomplete and stop at week 17 of the present extract.

The data are obtained from more than one official source. 2012-2022 data are sourced from the data. gov. sg weekly infectious disease bulletin [13]. 2023 data are obtained from CDA weekly bulletin PDF files and 2024-2026 data are obtained from CDA weekly bulletin Excel files [14-17]. Given the different formats of the sources, all observations are converted into a single weekly format for analysis.

The forecasting target is the total weekly dengue cases (Dengue Fever (DF) and Dengue Haemorrhagic Fever (DHF). Separate DF and DHF models are not fitted, as the research focus lies on overall surveillance burden and DHF cases remain scarce: DHF accounts for only 507 out of 192,677 total cases, with no DHF occurrences recorded in 428 of the 748 monitored weeks. This finding is consistent with previous Singapore research on total notified dengue cases [5].

$$C_t = DF_t + DHF \tag{1}$$

Once the target has been defined, the resultant combined weekly series is subjected to data cleaning and validation. This will guarantee that subsequent transformation, autocorrelation and forecast comparisons are done using an evenly spaced and internally consistent series.

3.3. Data Cleaning and Validation

Since the records are in PDF, Excel and structured data formats, cleaning will standardize epidemiological week labels, weekly date ranges and chronological ordering prior to modelling.

The duplicate year-week records, missing weeks, negative counts and DF-plus-DHF totals are verified since ARIMA-type models need an evenly spaced and internally consistent series.

After the weekly framework and target values are proved, the analysis moves to transformation. The rationale behind the next step is to determine whether it is possible to stabilize outbreak-driven variance prior to stationarity and model identification.

3.4. Transformation of Weekly Case Counts

Raw dengue counts are non-negative and highly influenced by outbreak peaks. In large outbreaks, the number of cases can be significantly higher during a week compared to normal weeks, leading to a right-skewed distribution and a varying variance across series levels.

Three scales of modelling are taken into account: the raw count series, the logarithmic transformation log(1 + cases), and the square-root transformation. The 1 added in the logarithmic transformation enables weeks with zero cases to be transformed. Both the logarithmic and square-root transformations are considered as candidate variance-stabilizing transformations. They are compared based on distributional behavior, residual diagnostics and rolling one-week-ahead forecast performance.

$$Y_t = LOG(1 + C_t) \tag{2}$$

The final transformation is justified in the results section. The square-root scale is retained as a comparator, but the later diagnostics show stronger residual autocorrelation and worse out-of-sample RMSE than log(1 + cases). Therefore, subsequent modelling uses the logarithmic transformation.

3.5. Stationarity and Model Identification

The test of stationarity is done following transformation. The order is clear: transformation, testing stationarity, differencing and finally ACF/PACF interpretation.

The time plot and autocorrelation structure of the transformed level series is first examined. The Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are complementary tests for stationarity. First differencing is used when the transformed level is persistent. The checks are then repeated for the first-differenced transformed series.

$$\Delta Y_t = Y_t - Y_{t-1} \tag{3}$$

ACF and PACF are interpreted only after the differencing order has been established. They define tentative ARIMA orders, but no model is accepted without the candidate search and diagnostic checks described next.

3.6. Candidate Model Search

The non-seasonal ARIMA models are used to start the candidate search since the short-term dependence in the stationary transformed series is the immediate goal. The first ACF/PACF-based model is not the end of the search. An ARIMA specification-controlled neighborhood is considered to compare underfitting and unnecessary complexity.

The annual seasonal period of weekly dengue data is about 52 weeks. This pattern can be represented by SARIMA, although the terms of seasonal autoregressive and moving-average at a 52-week lag may be parameter-rich. Fourier harmonics are lighter since only a few sine-cosine pairs are needed to describe smooth annual motion and minimize overfitting risk; this justifies comparing Fourier-ARIMA to the non-seasonal ARIMA baseline [18].

$$\phi(B)(1-B)^d Y_t = \theta(B)\varepsilon_t \tag{4}$$

$$F_{\mathrm{t}} = \sum_{k=1}^{k}\left[a_{\mathrm{k}}\, sin\left(\frac{2\pi kt}{52}\right) + b_{\mathrm{k}}\, cos\left(\frac{2\pi kt}{52}\right)\right] \tag{5}$$

$$\phi(B)(1-B)^d Y_t = \theta(B)\varepsilon_t + F_t \tag{6}$$

Fourier-ARIMA is evaluated as a lightweight seasonal extension of the non-seasonal ARIMA baseline. The candidate set contains 26 ARIMA specifications and 5 Fourier-ARIMA specifications with K = 1 to K = 5. The same fitting, diagnostic, and validation procedures are applied to every candidate before hierarchical selection.

3.7. Model Selection Evidence

The selection of models is done in a predetermined hierarchy such that the ultimate choice can be replicated.

Step 1: Fit validity. Models which successfully complete estimation and report convergence are kept.

Step 2: Residual adequacy. A model retained should possess a Ljung-Box p-value of at least 0.05 and no more than two significant residual ACF lags out of the first 12 weekly lags. Models which do not pass this gate are ranked lower than models with good residuals.

Step 3: Metric ranks. AICc, BIC, MAE and RMSE are ranked in ascending order. AICc and BIC are penalized in-sample fit, whereas MAE and RMSE are one-week-ahead error on the original scale of case counts.

Step 4: Total score. The four ranks are added up and the smallest score is desired among models that have passed the residual gate. In case of ties, validation ranks are used followed by information-criteria ranks and lastly the fewer number of parameters.

Step 5: Supporting checks. RMSLE and bias are reported as sensitivity checks, but not as part of the combined score. Parameter estimates and coefficient significance are also provided to aid interpretation. They fail to override failure of the residual gate or the predefined rank order.

This hierarchy means that a model cannot be chosen simply because it has the smallest AICc, the smallest individual error measure or the most parameters. The Results section uses this process to the entire set of candidates and presents the table of coefficients and diagnostic evidence of the chosen model.

Performance is reported on the original weekly case-count scale. MAE is a summary of average absolute error, RMSE weights large misses more heavily, RMSLE checks relative error on low- and high-count weeks, and bias identifies systematic overprediction or underprediction.

$$MAE = (1/n)\sum_{I=1}^{n}|y_{\mathrm{i}} - \hat{y}_{\mathrm{i}}| \tag{7}$$

$$RMSE = \sqrt{[(1/n)\sum_{I=1}^{n}(y_{\mathrm{i}} - \hat{y}_{\mathrm{i}})^2]} \tag{8}$$

$$RMSLE = \sqrt{[(1/n)\sum_{I=1}^{n}(log(1 + y_i) - log(1 + \hat{y}_i))^2]} \quad (9)$$

$$Bias = (1/n)\sum_{I=1}^{n}(y_i - \hat{y}_i) \quad (10)$$

3.8. Rolling Validation and Forecast Evaluation

A recursive expanding-origin one-week-ahead procedure is used to evaluate the performance of forecasts. The initial training period consists of the first 90% of the ordered series, and the last 10% constitutes the evaluation period. The model parameters are estimated once during the initial training period.

The model generates a one-week-ahead forecast at every test week, based on observations that are available prior to the week. Once the forecast has been scored against the observed value, the observed value is added to the state of the model and then another forecast is generated. The information set thus increases by a week at every forecast origin with the estimated values of the coefficients being constant throughout the period of testing.

This process is favored over a single train-test forecast since it generates a series of true one-week-ahead errors over the entire test period. It also indicates routine surveillance, in which a new observation is available every week prior to the next forecast being made. This prevents future information leakage since no observation at or beyond week t is employed when predicting week t.

The one-week-ahead forecasts have a 90% prediction interval. The interval is useful in public health use because it demonstrates a reasonable upper limit instead of having one number that is definitive.

The combination of the rolling errors, interval coverage, residual diagnostics and hierarchical model ranking is used to determine whether the chosen model is appropriate as a transparent short-term surveillance tool.

4. Experiments and Results

4.1. Case Pattern and Transformation Evidence

The cleaned data had 748 consecutive epidemiological weeks (2012-W01 to 2026-W17) without any missing or duplicated year-week records. There was significant variation in weekly burden: the raw mean was 257.6 cases, variance 72,891.9, skewness 2.63 and the 99th percentile 1,451 cases.

The log(1 + cases) transformation minimized the skewness to 0.10 and diminished the impact of the biggest outbreaks without affecting the sequence and time of weekly changes. The square-root transformation enhanced the raw scale, but with greater residual autocorrelation remaining and worse rolling RMSE than with the logarithmic scale. Fig. 1 thus presents the raw and log-transformed series along with their distributions and further model identification is done using log(1 + cases).

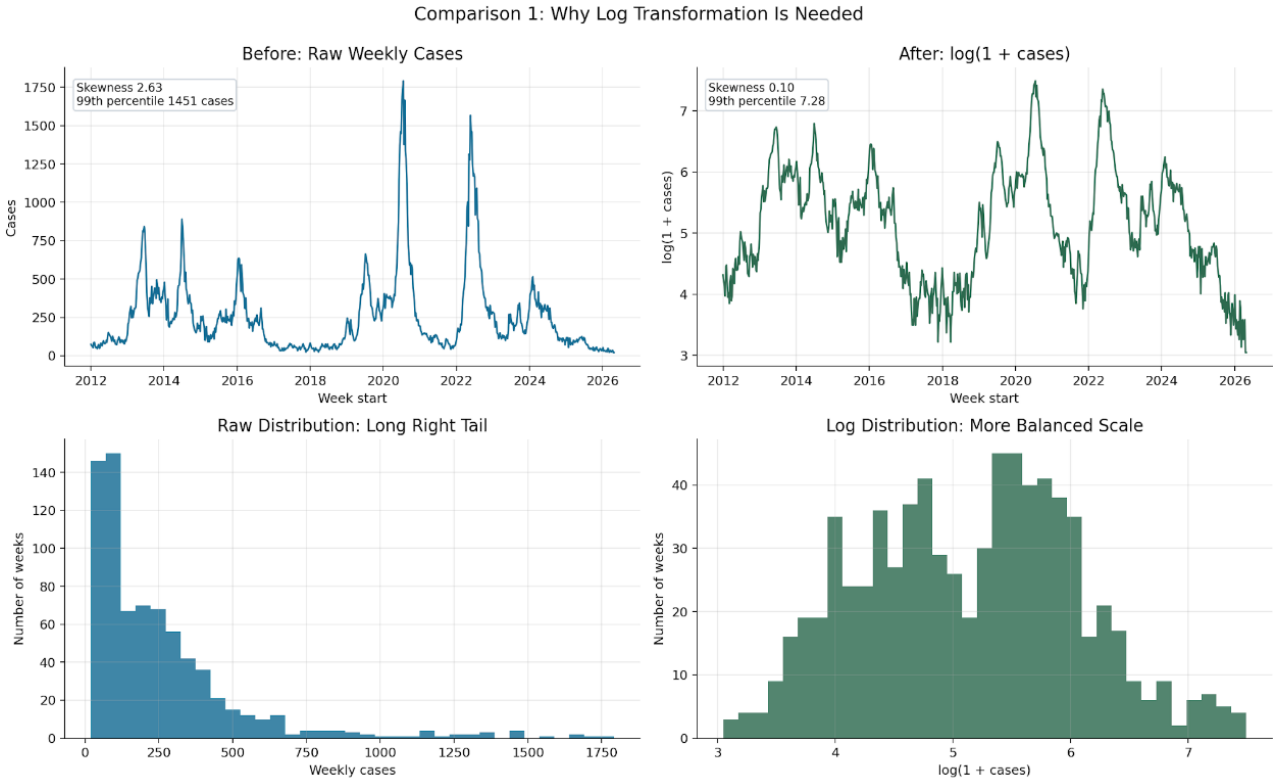


Fig. 1 Raw weekly dengue cases and log(1 + cases) for 748 consecutive epidemiological weeks from 2012-W01 to 2026-W17, with the corresponding empirical distributions **(**Photo credit: Original)

4.2. Stationarity and Initial Model Identification

The log-level series was not sufficiently stationary for direct ARIMA fitting. The augmented Dickey-Fuller test produced $p = 0.1128$, while the lag-1 autocorrelation remained 0.976. After one ordinary difference, the ADF test rejected a unit root with $p < 0.001$ and the KPSS test did not reject stationarity at the 5% level. The lag-1 autocorrelation decreased to -0.205 and the lag-52 autocorrelation was 0.005. These results supported an integration order of $d = 1$.

Seasonal differencing alone did not resolve the strong persistence: its ADF p-value was 0.0787 and lag-1 autocorrelation remained 0.975. Although combined ordinary and seasonal differencing also produced stationary test results, the strong negative lag-52 autocorrelation of -0.485 indicated a risk of unnecessary seasonal differencing. The ordinary-differenced log series was therefore used. Its short-lag ACF and PACF indicated an autoregressive starting structure, and ARIMA(1,1,0) was used as the first tentative model rather than as the final specification.

4.3. Candidate Model Comparison

The controlled search evaluated 26 non-seasonal ARIMA specifications. After the ARIMA structure had been assessed, five Fourier-ARIMA extensions with $K = 1$ to $K = 5$ were fitted using ARIMA (4,1,0) errors. All four models shown in Table 1 passed the residual adequacy gate, so the final comparison depended on the predefined balance of information criteria and rolling one-week-ahead errors.

Table 1. Leading models in the model-development comparison

Model	AICc	BIC	MAE / RMSE	Score
ARIMA(4,1,0)+Fourier(2)	-345.711	-304.458	37.866 / 65.743	14
ARIMA(4,1,0)	-330.746	-307.773	38.475 / 69.671	18
ARIMA(4,1,0)+Fourier(1)	-333.899	-301.775	37.585 / 67.991	21
ARIMA(0,1,3)	-330.712	-312.323	38.580 / 69.692	21

ARIMA(4,1,0)+Fourier(2) obtained the lowest combined rank score of 14. The non-seasonal ARIMA(4,1,0) baseline ranked second with a score of 18. Fourier(1) produced a slightly lower MAE than Fourier(2), but its RMSE and information-criterion balance were weaker. Fourier orders above K = 2 reduced AICc further but increased rolling forecast errors. This pattern indicated that additional harmonics improved in-sample fit without improving short-horizon prediction. The two-harmonic model was therefore selected as the best balance of accuracy, statistical adequacy, and parsimony.

The model-development MAE and RMSE in Table 1 were calculated over the common 332-week candidate-comparison period. They are used only to rank candidate specifications. The final held-out performance reported in Section 4.5 uses the separate 90/10 evaluation defined in Methodology Section 3.8.

4.4. Final Model Coefficients and Residual Adequacy

Table 2 presents the estimated coefficients of ARIMA(4,1,0)+Fourier(2). Four out of the eight non-variance parameters were significant at the 5% level: the first and second annual harmonic cosine terms, AR(1), and AR(3). The sine terms and AR(2) and AR(4) were not significant individually. The coefficient significance was used as an aid to the decision-making process instead of a rule to be automatically deleted since dropping terms would alter the predefined ARIMA error structure and Fourier pairs applied in the validated candidate comparison.

Table 2. Coefficient estimates for the selected ARIMA(4,1,0)+Fourier(2) model

Parameter	Component	Estimate (SE)	z	p
$\sin(2\pi t/52)$	Fourier 1 sine	-0.0215 (0.0777)	-0.277	0.7817
$\cos(2\pi t/52)$	Fourier 1 cosine	-0.2202 (0.0731)	-3.012	0.0026
$\sin(4\pi t/52)$	Fourier 2 sine	0.0731 (0.0376)	1.945	0.0517
$\cos(4\pi t/52)$	Fourier 2 cosine	0.1287 (0.0359)	3.582	<0.001
AR(1)	ARIMA error	-0.2397 (0.0320)	-7.489	<0.001
AR(2)	ARIMA error	-0.0112 (0.0334)	-0.335	0.7377
AR(3)	ARIMA error	0.1298 (0.0356)	3.643	<0.001
AR(4)	ARIMA error	0.0149 (0.0349)	0.428	0.6689
sigma^2	Variance	0.0359 (0.0017)	21.094	<0.001

Residual diagnostics supported the adequacy of the selected model. The residual mean on the log scale was -0.0017. Residual autocorrelation was -0.004 at lag 1 and 0.020 at lag 12, with no significant residual ACF lags among the first 12 weeks. The Ljung-Box p-values were 0.99993 at lag 12 and 0.99575 at lag 24. The remaining residuals were therefore consistent with random fluctuation around zero, and no material short-lag dependence remained.

4.5. Final One-Week-Ahead Test Performance

The final evaluation used the first 673 weeks for initial estimation and the last 75 weeks for recursive one-week-ahead testing. This 90/10 test is distinct from the longer candidate-development comparison in Section 4.3. Fig. 2 shows that the forecasts followed the broad decline and short-term movement of observed cases during the test period, although several sudden weekly changes were smoothed.

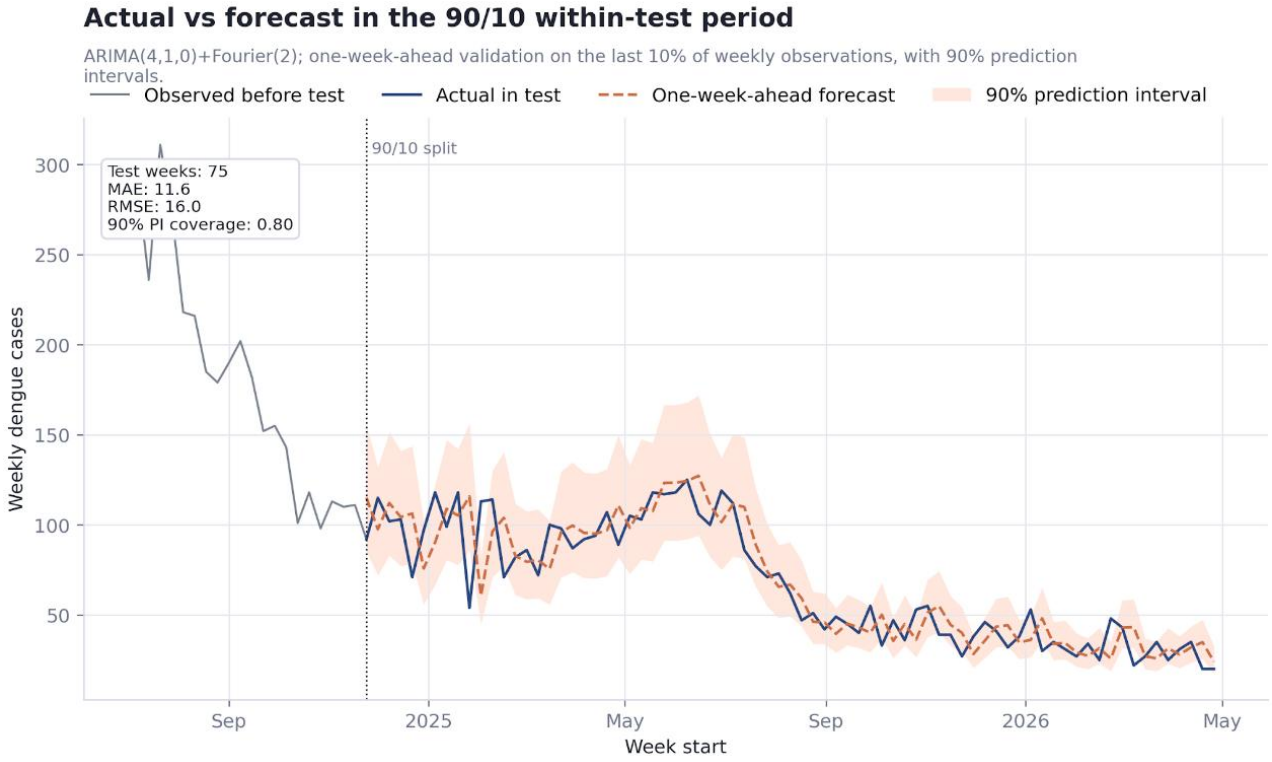


Fig. 2 Actual and recursive expanding-window one-week-ahead forecast dengue cases in the final 90/10 test period, covering 75 observation weeks, with 90% prediction intervals **(**Photo credit: Original)

On the original case-count scale, MAE was 11.6 cases and RMSE was 16.0 cases. Mean observed burden was 67.5 cases per week, compared with a mean forecast of 68.9. Bias, defined as actual minus forecast, was -1.4 cases, indicating slight average overprediction. The 90% prediction intervals covered 80% of test observations. Point accuracy was therefore informative for short-term monitoring, but interval calibration was below its nominal level and should not be described as fully calibrated.

4.6. Eight-Week Forecast

The fitted model was extended from the last observed week, 2026-W17, to 2026-W25 to examine the immediate eight-week forecast path. Point forecasts ranged from 21.2 to 26.5 cases and had a mean of 24.0 cases. Under the study's mean plus 1.5 standard deviations rule, the study monitoring threshold was 27.0 cases. This threshold is used as a simple monitoring benchmark informed by Singapore bulletin practice of comparing current weekly burden with historical baseline ranges; it is not treated as an official outbreak declaration rule. As shown in Fig. 3, no point forecast exceeded this monitoring threshold, and the highest point forecast was 26.5 cases in 2026-W25.

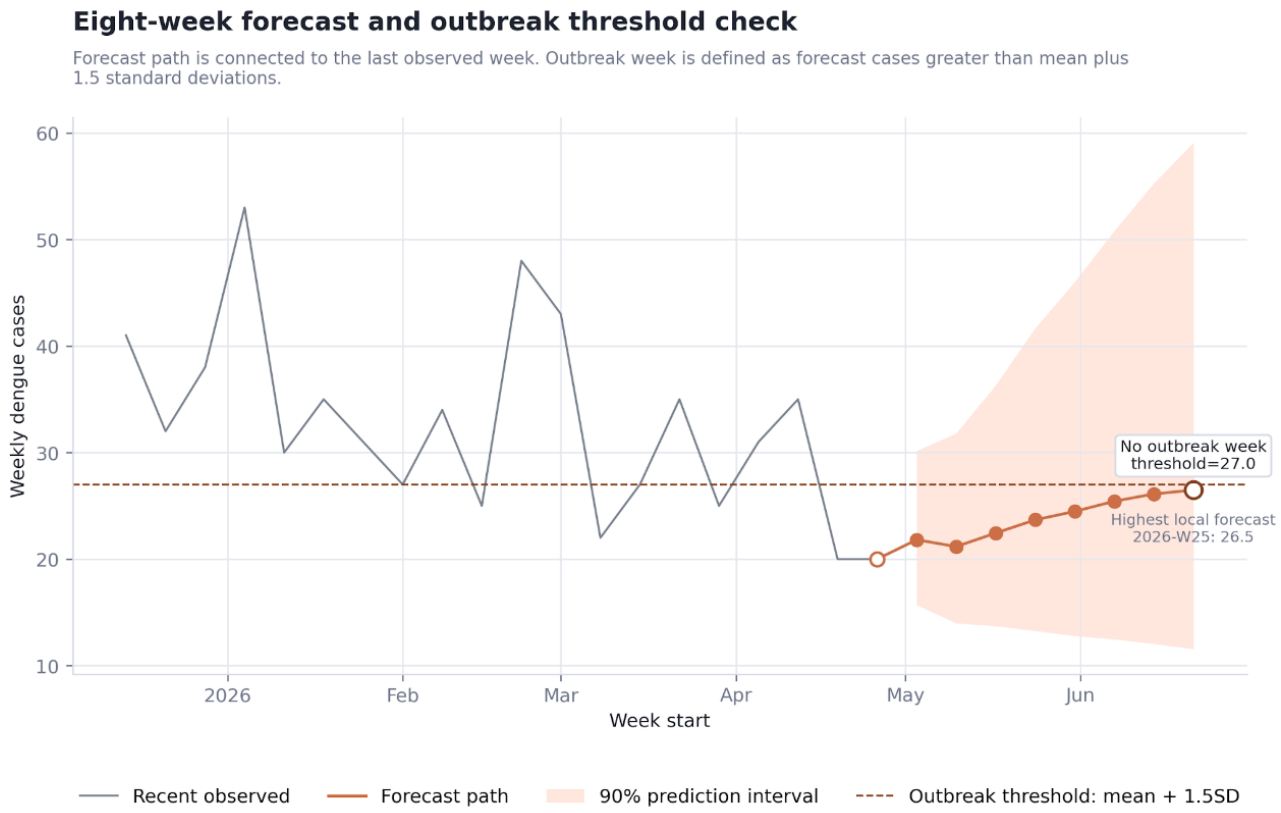


Fig. 3 Eight-week dengue forecast extrapolated from the last observed week, 2026-W17, to 2026-W25, with the study monitoring threshold and 90% prediction intervals **(**Photo credit: Original)

Forecast uncertainty increased with the horizon. The 90% interval was 15.7 to 30.2 cases for 2026-W18 and widened to 11.6 to 59.1 cases by 2026-W25. Therefore, the point forecast remained below the monitoring threshold while the upper plausible range increased. The immediate result was a moderate point forecast with substantial longer-horizon uncertainty, not evidence that future outbreak risk was absent.

5. Public Health Implications and Discussion

5.1. Public Health Meaning of the Forecast

The findings are to be viewed as the evidence of surveillance and not an exhaustive description of dengue transmission. The chosen ARIMA(4,1,0)+Fourier(2) model is based on only reported weekly case counts, thus it estimates the anticipated short-term trend of reported burden but does not determine biological causes. Its practical significance is that official case notifications are routinely available and can be revised without waiting until climate, vector, serotype or mobility data are available.

This case-only design is also friendly to public health communication. The inputs, transformation, differencing, candidate comparison, residual checks and rolling validation can be seen, which makes the users know why the model was chosen, what it uses and what it cannot tell them.

5.2. Early Warning and Situational Awareness

Short-term situational awareness is the primary use of the model in public health. The forecast one week ahead provides a reference value to be used in the following surveillance week. In case the number of observed cases increases beyond the projected trajectory, or when the predicted range shifts towards the study monitoring limit, this can be interpreted by the staff as an indicator of closer monitoring instead of being an automatic outbreak call.

This is depicted in the eight-week forecast. Between 2026-W18 and 2026-W25, all point forecasts were below the 27.0-case monitoring level with a high of 26.5 cases in 2026-W25. The 90% interval, however, expanded to 15.7-30.2 to 11.6-59.1 cases, therefore, the outcome must be interpreted as moderate burden with further uncertainty, rather than no risk of outbreak.

5.3. Resource Planning and Preparedness

Short-term preparedness can also be facilitated by forecasts. The control of Dengue needs field inspection, planning of vector-control, laboratory preparedness, awareness of healthcare and communication to the people. Previous Singapore models tended to use weather, vector or serotype data to justify longer-horizon warning [1, 2, 11]. The current model has a more limited but practical benefit: it can be updated using only weekly case counts, thus providing a low-cost baseline signal in cases where external covariates are delayed, unavailable, or hard to interpret.

The high-alert weeks are not the only practical value. Moderate forecast can justify routine readiness, and an increasing forecast or widening upper interval can trigger review of inspection schedules, laboratory capacity, risk communication, and field intelligence before a significant increase is clearly visible.

5.4. Interpreting Uncertainty and Reliability

The uncertainty interval should be taken into account by forecast users, in addition to the point forecast. MAE was 11.6 cases and RMSE was 16.0 cases in the final 90/10 evaluation that demonstrates practical short-term point accuracy. The empirical coverage of the 90% prediction interval however was 80%. This implies that the range was smaller than desired during the test period. This is a significant limitation when it comes to public health use. A forecast range must be considered as a planning range rather than a guaranteed range.

Model adequacy was supported by residual diagnostics, there was no significant residual autocorrelation in the first 12 weekly lags and a Ljung-Box p-value of 0.9999 at lag 12. Nonetheless, dengue transmission may abruptly vary due to weather, mosquito activity, behavior, reporting, serotype circulation or interventions; forecasts must be integrated with field data and expert opinion.

5.5. Limitations and Future Use

The main limitation is the deliberate case-only scope. It can describe temporal dependence in reported cases, but it cannot explain why transmission changes or estimate the effect of climate, mosquito density, serotype patterns, spatial clusters, population movement, or interventions.

The second limitation is that the model has been fitted to Singapore surveillance data. The framework can be reproduced, but the parameters and performance chosen should not be applied to a different country or timeframe without being revalidated. The model should be prospectively tested in future work, use residual bootstrap techniques to better calibrate 90% intervals without altering the ARIMA-Fourier structure, and compare fixed thresholds with operational thresholds selected by public health users. Climate, vector, spatial, serotype, and intervention variables may be incorporated as external predictors when reliable weekly data are available. The current study is thus an open standard of short-term dengue surveillance, rather than the ultimate shape of an operational warning system.

6. Conclusion

This paper evaluated the ability of routinely reported weekly dengue cases to aid in transparent short-term surveillance forecasting in Singapore. It tracked a reproducible case-only pipeline using 748 epidemiological weeks between 2012-W01 and 2026-W17, which included validation, transformation, assessment of stationarity, model comparison, diagnosis of residuals, and time-ordered forecast evaluation. This was to test the extent to which official case notifications can provide information on short-term warning by themselves.

The findings indicate that the latest weekly cases have valuable one-week-ahead forecasting information. The chosen ARIMA(4,1,0)+Fourier(2) model had an MAE of 11.6 cases and RMSE of 16.0 cases in the last 90/10 test. Adequacy was also supported by residual checks where there was no material autocorrelation at the first 12 weekly lags and a Ljung-Box p-value of 0.9999 at lag 12.

The last model kept two Fourier harmonics since this specification gave the optimal trade-off of forecast error, information criteria, adequacy of residuals and parsimony. Increasing Fourier orders increased in-sample fit but failed to increase rolling forecast performance indicating a risk of overfitting. The last model was thus chosen based on predefined evidence as opposed to a favorable measure.

In practice, the framework can support one-week-ahead monitoring and can be updated when new weekly case counts become available. For 2026-W18 to 2026-W25, point forecasts remained below the study monitoring threshold, but the interval widened to about 11.6 to 59.1 cases by 2026-W25 and empirical 90% coverage was only 80%. The forecast should therefore be interpreted as decision-support evidence, not as a guarantee that outbreak risk is absent.

The primary weakness is the case-only nature. Climate, vector, serotype, spatial and intervention predictors should be added in future work when these are available on a weekly basis, probabilistic calibration should be improved and operational thresholds for alerts tested prospectively. Within this context, an interpretable Fourier-ARIMA model is shown to deliver valuable short-term support for Singapore's dengue surveillance system.

References

[1] Hii Y. L., Zhu H., Ng N., et al. Forecast of dengue incidence using temperature and rainfall. PLOS Neglected Tropical Diseases, 2012, 6(11), e1908. https://doi.org/10.1371/journal.pntd.0001908

[2] Shi Y., Liu X., Kok S. Y., et al. Three-month real-time dengue forecast models: An early warning system for outbreak alerts and policy decision support in Singapore. Environmental Health Perspectives, 2016, 124(9), 1369-1375. https://doi.org/10.1289/ehp.1509981

[3] Earnest A., Tan S. B., Wilder-Smith A., et al. Comparing statistical models to predict dengue fever notifications. Computational and Mathematical Methods in Medicine, 2012, Article 758674. https://doi.org/10.1155/2012/758674

[4] Leung X. Y., Islam R. M., Adhami, M., et al. A systematic review of dengue outbreak prediction models: Current scenario and future directions. PLOS Neglected Tropical Diseases, 2023, 17(2), e0010631. https://doi.org/10.1371/journal.pntd.0010631

[5] Dayama P., Kameshwaran S. Predicting the dengue incidence in Singapore using univariate time series models. AMIA Annual Symposium Proceedings, 2013, 285-292. https://pubmed.ncbi.nlm.nih.gov/24551338/

[6] Fang L., Hu W., Pan G. Meteorological factors cannot be ignored in machine learning-based methods for predicting dengue, a systematic review. International Journal of Biometeorology, 2024, 68, 401-410. https://doi.org/10.1007/s00484-023-02605-1

[7] Lima C. L., da Silva A. C. G., Moreno G. M. M., et al. Temporal and spatiotemporal arboviruses forecasting by machine learning: A systematic review. Frontiers in Public Health, 2022, 10, 900077. https://doi.org/10.3389/fpubh.2022.900077

[8] Sylvestre E., Joachim C., Cécilia-Joseph, E., et al. Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review. PLOS Neglected Tropical Diseases, 2022, 16(1), e0010056. https://doi.org/10.1371/journal.pntd.0010056

[9] Tian N., Zheng J. X., Li L. H., et al. Precision prediction for dengue fever in Singapore: A machine learning approach incorporating meteorological data. Tropical Medicine and Infectious Disease, 2024, 9(4), 72. https://doi.org/10.3390/tropicalmed9040072

[10] Zhao X., Li K., Ang C. K. E., et al. A deep learning based hybrid architecture for weekly dengue incidences forecasting. Chaos, Solitons & Fractals, 2023, 168, 113170. https://doi.org/10.1016/j.chaos.2023.113170

[11] Finch E., Chang C. C., Kucharski A., et al. Climate variation and serotype competition drive dengue outbreak dynamics in Singapore. Nature Communications, 2025, 16, Article 11364. https://doi.org/10.1038/s41467-025-66411-6

[12] Frasca M., La Torre D., Pravettoni G., Cutica I. Explainable and interpretable artificial intelligence in medicine: A systematic bibliometric review. Discover Artificial Intelligence, 2024, 4, 15. https://doi.org/10.1007/s44163-024-00114-7

[13] Ministry of Health Singapore. Weekly Infectious Disease Bulletin Cases (2024) Dataset, data.gov.sg, 2016. Retrieved July 7, 2026, from https://data.gov.sg/datasets/ d_ca168b2cb763640d72c 4600a68f 9909e/ view

[14] Communicable Diseases Agency Singapore. Weekly infectious diseases bulletin 2023. Retrieved July 7, 2026, https://www.cda.gov.sg/resources/weekly-infectious-diseases-bulletin-2023/

[15] Communicable Diseases Agency Singapore. Weekly infectious diseases bulletin 2024. Retrieved July 7, 2026, https://www.cda.gov.sg/resources/weekly-infectious-diseases-bulletin-2024/

[16] Communicable Diseases Agency Singapore. Weekly infectious diseases bulletin 2025. Retrieved July 7, 2026, https://www.cda.gov.sg/resources/weekly-infectious-diseases-bulletin-2025/

[17] Communicable Diseases Agency Singapore. Weekly infectious diseases bulletin 2026. Retrieved July 7, 2026,https://www.cda.gov.sg/resources/weekly-infectious-diseases-bulletin-2026/

[18] Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice (3rd ed.), sections on Fourier series, dynamic harmonic regression, and time-series cross-validation. OTexts, 2021. https://otexts.com/fpp3/