Prediction Dengue Fever Cases Semarang City Indonesia

Hanif Pandu Suhito[1]*, Mahalul Azam[1], Dina Nur Anggraini Ningrum[1], Anan Nugroho[1], Sholikun[2]

[1]Universitas Negeri Semarang, Central Java, Indonesia

[2]Dinas Kesehatan Kota Semarang, Central Java, Indonesia

Corresponding author: mashanifps@gmail.com

Abstract: Dengue fever is often in the spotlight in Indonesia every year. Data from the Ministry of Health shows that until the 17th week of 2024, there were 88,593 cases of dengue fever with 621 deaths in Indonesia. Semarang City is still an endemic dengue case, so a comprehensive control of the incidence of cases in the community is needed through the calculation of case predictions. To predict the number of dengue fever cases per month per sub-district and compare with the real cases that occurred during January - August in 2024. We used the exponential smoothing method in this study. Data source: The data used is dengue fever case data taken from the dengue fever information system report (Tunggal Dara) for the period 2016-2023. The variables included include the number of cases per month, the population, and sub districts, while the data set obtained is 1616 data. The prediction of DHF cases per month for 2024 per sub-district area was produced, with an average MAE value of 1.55, indicating that the average difference between the predicted value and the observed number of DHF patients is relatively small. The average value of alpha in smoothing coefficients is 0.286, meaning that the influence of the latest data on the predicted value has a moderate weight. Tugu sub-district had an RMSE value of 1.101, meaning that the predicted values were close to the real cases and during the 8 months of observation, the results were correct. The prediction of cases per month per sub-district is used for prevention efforts against the occurrence of DHF cases, the hope is not to exceed the predictions produced. If it can be below the prediction results, it means that there is a success of promotive and preventive efforts made by all parties.

Keywords: Dengue Fever, Prediction, Exponential Smoothing Methods

INTRODUCTION

Dengue hemorrhagic fever (DHF) is an infectious disease that garners significant attention in Indonesia. The dengue virus, which can cause death, transmits the disease through the bite of Aedes aegypti and Aedes albopictus mosquitoes (Lutfiana Inda Hapsari et al., 2022). As of the 17th week of 2024, the Ministry of Health reported 88,593 cases and 621 fatalities in Indonesia. (Kementerian Kesehatan RI, 2024). Semarang City remains endemic for DHF cases, necessitating comprehensive control of the incidence of cases in the community through case prediction calculations.

Forecasting is defined as making future predictions based on past and present data and analyzing trends. Time-series forecasting uses a model to predict future values on the basis of previously observed values (GEP Box, GM Jenkins, GC Reinsel, 1994). Predicting the number of dengue cases is an attempt to overcome disease prevention and control. Forecasting is a method used to predict or estimate a value in the future or future period using data from the previous period (Alex & Nur Rahmawati, 2023). Forecasting is very useful to prepare us for conditions that may occur in the coming period and can also be used by companies for decision-making (Nurul Hudaningsih et al., n.d.). With accurate predictions, health authorities can take more appropriate measures, such as the distribution of medical resources and community counseling. Moreover, case prediction supports the development of data-informed health policy, thereby enhancing dengue prevention efforts.

This study aims to predict the number of dengue cases per month in each sub-district in Semarang City in 2024 and compare the prediction results with the real data of cases that occurred in the period January to August 2024, using the exponential smoothing method. One of the time series forecasting techniques, the exponential smoothing method, predicts data with trend and seasonal patterns by identifying regular patterns that persist into the future and assigning a weight to the latest data (Tableau, 2024). The exponential smoothing method is a development technique of the moving averages method. The basic idea of exponential smoothing method is that the closer to the predicted point in a time series, the greater the effect. The further away from the predicted point, the less effect it has. This method forecasts continuously using the most recent data, employing single, double, and triple calculations (Kurniagara, 2017). Evaluation of forecasting techniques is carried out by the level of accuracy of the utilization of Mean Forecast Error (MFE), Mean Absolute Deviation (MAD), Mean Square Error (MSE), and Root Mean Squared Error (RMSE)(Agomoh & Ukabuiro, 2023). The weight of different data is weighted according to the distance in time and the weight is decayed exponentially. The method has 3 parameters to control: Alpha, Gamma, and Delta, corresponding to the level at the current point in time, the slope of the trend part, and the seasonal part, respectively. The parameters Alpha, Gamma, and Delta all have values ranging from 0 to 1, and the closer they are to 1, the greater their weight in the prediction (Zaijin et al., 2022)

The data used is dengue fever case data taken from the Semarang City Dengue Information System (Tunggal Dara) for the period 2016 to 2023, which includes the number of dengue cases per month, population, and sub-district location.By reporting patient cases online and in real time, Tunggal Dara serves as an integrated system that bridges the barriers in efforts to control dengue incidence. (Dinas Kesehatan Kota Semarang, 2024). This analysis can use the prediction results as a tool to prevent and manage dengue cases.

METHOD

The exponential smoothing method predicts the incidence of dengue fever (DHF) cases in Semarang City from January to December 2024. The exponential smoothing method employs previous patterns to predict data, emphasizing the most recent information while diminishing the significance of earlier data. Exponential smoothing effectively captures trends and seasonal patterns in time series data, especially when predicting DHF cases influenced by seasonal variables such as weather, temperature, and humidity. This method enhances predictive accuracy, enabling local governments to implement preventive actions proactively based on estimated case variations in each sub-district. The Exponential Smoothing approach incorporates trend, error, and seasonal components through smoothing calculations (Lusiana & Yuliarty, 2020). The smoothing calculation equation, Yt = Tt + St + Et (1), utilizes an additive decomposition by summing the trend (T), seasonal (S), and error (E) components to get the observed series. The equation's general form, Yt = Tt - St - Et (2), combines the trend (T), seasonal (S), and error (E) components to get the observed series (Nurmalasari et al., 2021).

The exponential smoothing method is employed to predict the incidence of dengue cases, accounting for seasonal variations in each sub-district. Each subdistrict in Semarang has distinct seasonal patterns, characterized by an escalation of cases in some months (the peak seasonal effect) and a decline in other months (the minimal seasonal effect). Furthermore, we utilize smoothing coefficients to improve the model's accuracy in addressing monthly variations in data patterns. The evaluation of the model utilized RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and AIC (Akaike Information Criterion) metrics to assess the accuracy of prediction outcomes in the sub-district.

RESULTS

Figure 1 shows the distribution of predicted DHF cases from January to December 2024. Each sub-district has a different prediction trend, which can be used as a reference in evaluating the effectiveness of dengue prevention programs in each region.

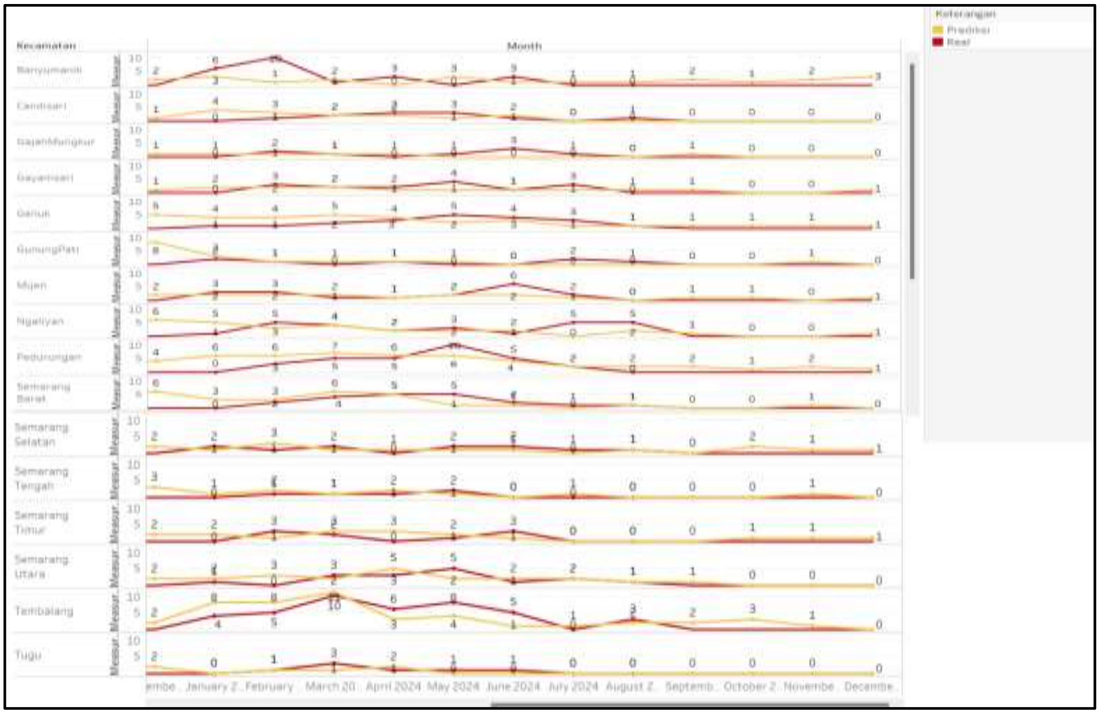

Figure 1. Prediction Chart of DHF from January to December 2024

The study of the prediction gap for dengue fever cases in Semarang City from January to December 2024. On January to August 2024 reveals a difference between the actual cases and what was predicted in each sub-district, as illustrated in Table 1.

Table 1. Gap Evaluation of Case Prediction January-August 2024

Kecamatan	Jan	Feb	Mar	Apr	Mey	Jun	Jul	Aug
Banyumanik	-3	-9	-1	-3	3	-2	1	1
GajahMungkur	1	-1	0	1	0	-3	0	0
Mijen	-1	-1	1	0	2	-4	1	0
Tugu	0	0	-2	1	0	-1	0	0
SemrangSelatan	-1	2	-1	1	1	-1	0	0
Gunung Pati	1	0	1	0	1	0	0	-1
Ngaliyan	4	-2	0	0	2	1	0	-3
Gayamsari	2	-1	0	-1	1	0	1	1
SemarangTimur	2	-2	1	3	2	-2	0	0
SemarangTengah	1	1	0	1	1	0	1	0
Tembalang	4	3	1	-3	4	-4	1	-1
Semarang Barat	3	1	2	0	1	-1	0	0
Candisari	4	2	0	0	1	1	0	-1
Semarang Utara	1	3	-1	2	2	1	2	0
Genuk	3	3	3	1	2	-1	1	0
Pedurungan	6	3	2	1	6	-1	2	2

Table 2. Evaluation of Predictive Accuracy by Month from January to August 2024

Kecamatan	Prediksi Tepat	Persentase Prediksi Tepat %	Under estimasi	Persentase Under estimasi %	Over estimasi	Persentase Over estimasi %
Banyumanik	0	0	5	63	3	38
GajahMungkur	4	50	2	25	2	25
Mijen	2	25	3	38	3	38
Tugu	5	63	2	25	1	13
SemrangSelatan	2	25	3	38	3	38
Gunung Pati	4	50	1	13	3	38
Ngaliyan	3	38	2	25	3	38
Gayamsari	2	25	2	25	4	50
SemarangTimur	2	25	4	50	2	25
SemarangTengah	3	38	0	0	5	63
Tembalang	0	0	3	38	5	63
Semarang Barat	3	38	1	13	4	50
Candisari	3	38	1	13	4	50
Semarang Utara	1	13	1	13	6	75
Genuk	1	13	1	13	6	75
Pedurungan	0	0	1	13	7	88

Table 2 presents data on the duration of months from January to August 2024 that exhibited discrepancies between predictions and actual cases, classified into accurate predictions, overestimations, and underestimations. Accurate prediction (Prediction = real) occurs when the predicted value matches the actual case, indicating the model's precision in forecasting cases in the region. Overestimation (Prediction > Real) occurs when the predicted value exceeds the actual case results, indicating that the model's predictions are too high compared to reality. Underestimation (Prediction < Real) occurs when the predicted value is lower than the actual results, indicating that what the model produces is lower than the real cases. It is anticipated that the predictive models developed within illness preventive activities would not be undervalued.have been produced.

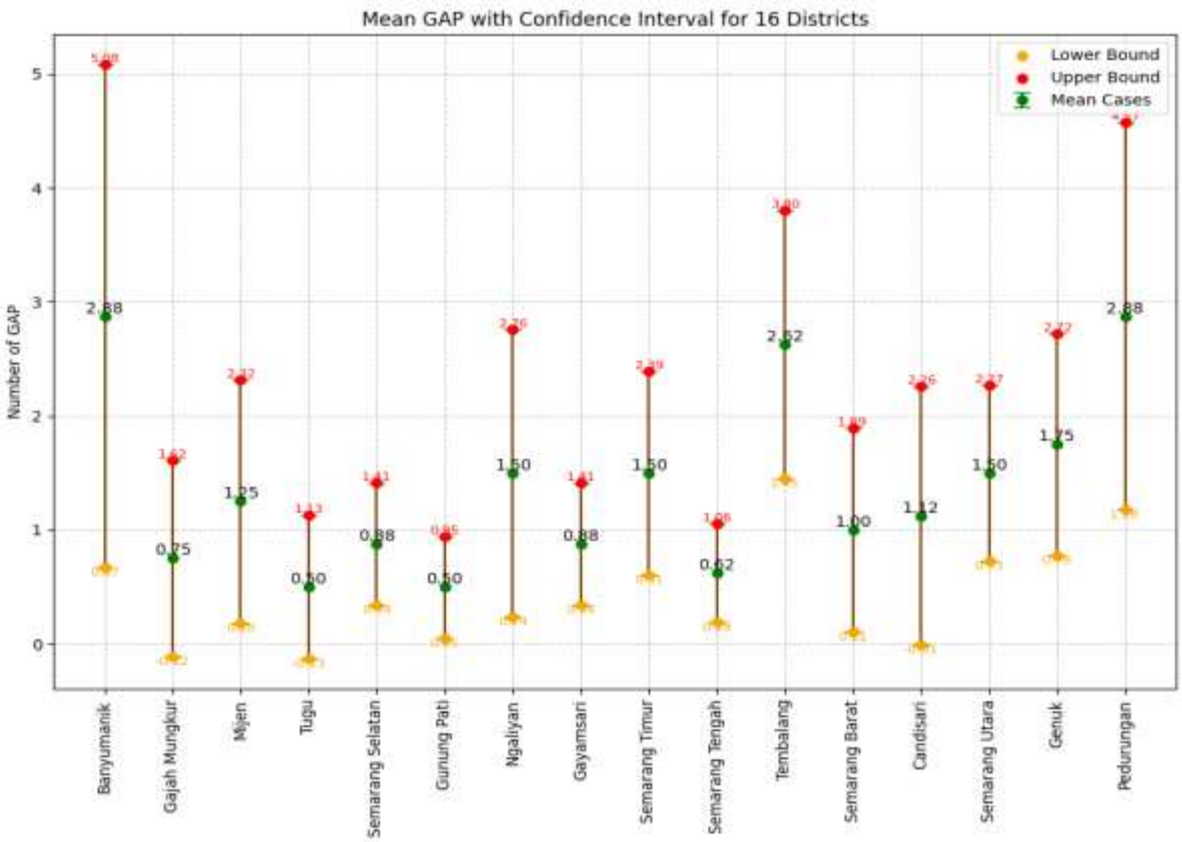

Figure 2. Mean Gap with Confidence interval for 16 Districs

Figure 2 illustrates the confidence interval for the average number of cases (GAP) across 16 districts in Semarang. In the graph depicted, each green dot signifies the mean number of cases in the district, while the orange and red lines and dots denote the lower and upper limits of the confidence interval at a 95% confidence level.

Table 3. Quality Matrics Smoothing Coefficients

Kecamatan	RMSE	MAE	MASE	MAPE	AIC
Tugu	1.101	0.682	0.81	28.0%	48
Tembalang	4.256	3.198	0.52	73.0%	308
Semarang Utara	2.308	1.619	0.63	42.8%	191
Semarang Timur	1.557	1.179	0.78	34.8%	115
Semarang Tengah	1.201	0.883	0.83	23.2%	65
Semarang Selatan	1.385	1.066	0.72	36.5%	93
Semarang Barat	2.182	1.671	0.64	47.8%	180
Pedurungan	2.867	2.179	0.64	56.2%	232
Ngaliyan	2.998	2.103	0.54	49.8%	241
Mijen	1.955	1.409	0.66	41.9%	159
Gunung Pati	1.487	1.069	0.67	38.0%	106
Genuk	2.279	1.745	0.76	63.2%	188
Gayamsari	1.269	0.908	0.79	26.5%	76
Gajah Mungkur	1.503	1.036	0.64	33.0%	108
Candisari	2.316	1.661	0.59	49.1%	191
Banyumanik	3.213	2.388	0.48	52.3%	254

The Quality Metrics Smoothing Coefficients Table indicates that the projected monthly dengue cases for each district in Semarang City in 2024 have an average Mean Absolute Error (MAE) of 1.55, signifying a minimal average discrepancy between the predicted and actual dengue case numbers. This indicates a somewhat acceptable overall forecast accuracy.

Particularly in the Tugu District, the predictive model demonstrates high accuracy with an RMSE value of 1.101, signifying that the forecasts for dengue cases closely align with the actual case numbers. Over an eight-month observation period in Tugu, the predictive outcomes continually shown excellent accuracy, positioning this area as the most successful in comparison to other sub-districts.

Table 3. Smoothing Coefficients

Kecamatan	Alpha	Beta	Gamma
Tugu	0.212	0.000	0.180
Tembalang	0.500	0.000	0.004
Semarang Utara	0.493	0.000	0.179
Semarang Timur	0.266	0.000	0.175
Semarang Tengah	0.000	0.000	0.247
Semarang Selatan	0.252	0.000	0.180
Semarang Barat	0.343	0.000	0.161
Pedurungan	0.278	0.000	0.090
Ngaliyan	0.500	0.000	0.000
Mijen	0.400	0.000	0.047
Gunung Pati	0.249	0.000	0.141
Genuk	0.230	0.000	0.171
Gayamsari	0.000	0.000	0.125
Gajah Mungkur	0.264	0.000	0.100
Candisari	0.337	0.000	0.048
Banyumanik	0.466	0.000	0.026

The mean alpha value of 0.286 for the smoothing factors signifies that the latest data exerts a moderate impact on the prediction. The alpha weight ensures that the most recent data substantially influences the predictions without entirely overshadowing the expected values. The prediction findings offer a clear depiction of dengue case trends by district, characterized by a minimal average error and a balanced impact of the most recent data in the predictive model.

DISCUSSION

Smoothing Exponential is a powerful forecasting method that may serve as an alternative to the well-known Box-Jenkins ARIMA family of methods. This parameter regulates the speed at which the impact of observations from previous time steps diminishes exponentially. (Bhattacharjee et al., 2023). The predicted cases of dengue hemorrhagic fever (DHF) from January to December 2024 show significant variation among sub-districts in Semarang City, as illustrated in Figure 1. This distribution of predictions reflects different case trends in each area and can serve as an important tool for evaluating the effectiveness of DHF prevention programs in each subdistrict. According to the evaluation results of the gap between predictions and actual cases from January to August 2024 (Table 1), several districts, have a small prediction gap. This suggests that the predictive model functions effectively in the area, enabling more effective adjustments to preventive interventions based on the prediction results. But also several sub-districts, such as Pedurungan and Genuk have larger prediction gaps.

From the results of the Quality Metrics Smoothing Coefficients (Table 3), the average mean absolute error (MAE) value of 1.55 indicates that the difference between predictions and actual dengue cases is relatively small in most sub-districts. Particularly in Tugu Sub-district, the low Root Mean Square Error (RMSE) value of 1.101 reinforces the conclusion that predictions in that area are very accurate. This can be used as a guide to plan more efficient preventive actions in the future. In addition, the analysis of the smoothing coefficients (Table 3) shows that the average alpha value of 0.286 indicates a fairly balanced influence of recent data on the predictions. This value guarantees that the prediction model incorporates current trends, avoiding undue influence from insignificant short-term data changes.

Overall, the prediction results provide information about the distribution patterns of dengue cases by district. The relatively high accuracy in several areas indicates the predictions' potential as a tool to improve dengue prevention programs, while areas with significant prediction gaps require further evaluation to improve the model's effectiveness.

CONCLUSION

Use the monthly prediction of dengue fever (DHF) cases per district as a reference to prevent the number of cases from exceeding the projected figures. If all parties involved successfully reduce the number of cases below the predicted results, it signifies the success of their promotional and preventive efforts. The Tugu and Gajah Mungkur districts, where the prediction gap is small, demonstrate the effectiveness of this model in approaching the actual number of cases. This suggests that those areas are already effectively implementing preventive measures. On the other hand, districts like Pedurungan and Genuk, with larger prediction gaps, require evaluation and improvement of prevention strategies. Overall, accurate predictions, with a low average RMSE in several districts, help guide prevention efforts more effectively. We hope that districts that have successfully reduced the number of cases below predictions can serve as examples for other regions, implementing more optimal promotional and preventive measures.

Conflict of Interest

The author(s) declare that they have no conflict of interest.

Acknowledgment

A thank you to the agencies and parties who supported this research.

REFERENCES

Agomoh, S. A., & Ukabuiro, I. K. (2023). *Intelligent Sales Forecasting Technique Application*. European Journal of Theoretical and Applied Sciences, *1*(6), 641–653. https://doi.org/10.59324/ejtas.2023.1(6).64.

Alex, M. A. H., & Nur Rahmawati. (2023). *Application of the Single Moving Average, Weighted Moving Average and Exponential Smoothing Methods For Forecasting Demand At Boy Delivery*. Tibuana, *6*(1), 32–37. https://doi.org/10.36456/tibuana.6.1.6442.32-37.

Baharom, M., Ahmad, N., Hod, R., & Abdul Manaf, M. R. (2022). *Dengue early warning system as outbreak prediction tool: A systematic review. Risk Management and Healthcare Policy*, 15, 871-886. https://doi.org/10.2147/RMHP.S361106.

Bhattacharjee, Atanu, Gajendra K. Vishwakarma, Namrata Gajare, Neha Singh.(2023).*Time Series Analysis Using Different Forecast Methods and Case Fatality Rate for Covid-19 Pandemic,Regional Science Policy & Practice*, Volume 15, Issue 3, Pg 506-520, ISSN 1757-7802, https://doi.org/10.1111/rsp3.12555.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Bukhsh, Z. A., Stipanovic, I., Saeed, A., & Doree, A. G. (2020). *Maintenance intervention predictions using entity-embedding neural networks*. Automation in Construction, 116, 103202. https://doi.org/10.1016/j.autcon.2020.103202.

Dinas Kesehatan Kota Semarang. (2024). Sistem Informasi Demam Berdarah Kota Semarang (Tunggal Dara). Retrieved from http://116.254.113.136:8080/tunggaldara/

Götz, T., Altmeier, N., Bock, W., Rockenfeller, R., Sutimin, & Wijaya, K. P. (2017). *Modeling dengue data from Semarang, Indonesia*. Ecological Complexity, 30, 57-62. https://doi.org/10.1016/j.ecocom.2017.02.002.

Lusiana, A., & Yuliarty, P. (2020). Penerapan Metode Peramalan (Forecasting) pada Permintaan Atap di PT.X.

Lutfiana Inda Hapsari, Nonik Eka Martyastuti, & Erwin Lestyaningsih. (2022). PENA MEDIKA *Implementasi Kompres Hangat Pada Pasien Anak DHF (Dengue Haemorrhagic Fever) Dengan Hipertermi*. In *MEDIKA* (Vol. 12, Issue 2). http://jurnal.unikal.ac.id/index.php/medika.

Nurmalasari, M., Zuama Qomarania, W., Temesvari, N. A., & Pertiwi, T. S. (2021). *Bimbingan Teknis Peramalan Jumlah Kunjungan Pasien Dengan Tableau*. In Indonesian Journal of Health Information Management Service *(IJHIMS* (Vol. 1, Issue 1). https://public.tableau.com/

Nurul Hudaningsih, Silvia Firda Utami, & Wari Ammar Abdul Jabbar. (n.d.). *Perbandingan Peramalan Penjualan Produk AKNIL PT. Sunthi Sepuri Menggunakan Metode Single Moving Average dan Single Exponential Smoothing*. https://doi.org/https://doi.org/10.51401/jinteks.v2i1.554.

Tableau. (2024). *Understanding exponential smoothing in time series analysis*. https://www.tableau.com/learn/articles/exponential-smoothing.

Zaijin G, Hao G, Luojing Z. SARIMA model and Holt-Winters index smoothing method in predicting the number of tuberculosis cases in Jiangsu Province. Dis Surveill. 2022;37(8):1042–7.