International Journal of Research in Advent Technology, Vol. 4, No. 9, September 2016
E-WASSN: 2321-9637
Available online at www.ijrat.org

Forecasting Dengue Hemorrhagic Fever Cases Using Time Series in East Java Province, Indonesia

Hasirun[1], Windhu P[2], Chatarina U. W[3], Avie Sri Harivianti[4]
Department of Epidemiologi[1,3,] Department of Biostatistics[2], Public Health Faculty Airlangga University, Surabaya,[4]East Java Province Health Office, Indonesia
Email: run_fetp@yahoo.co.id[1]

Abstract-Dengue incidence continues to increase dramatically worldwide in the past decade. Indonesia is one of countries with highest dengue cases and one of the provinces with the highest dengue cases namely East Java. The objectives of this study were to forecasting model of dengue fever cases in East Java Province for 2015 and determine the major forecasting dengue fever cases in East Java province for a period of one year in 2015.
This used time series approach. The data used was the number of dengue cases in the province of East Java in 2005 to 2014. The data was collected by study of documentation. The unit of analysis was the months amounted to 120 months. The analysis used was time series with statistical application that aimed to forecast dengue cases in the province of East Java.
Forecasting results obtained ARIMA (0,1,1) $(1,1,1)^{12}$. The model was the best models of several possible models that existed because the only model that had a significant parameter, white noise and met the assumption of independence. The results could be interpreted that the large number of dengue fever patients from January till December 2015 was decreased. This model also had mean absolute error about 403 cases. Based on the results of forecasting, preferably East Java Provincial Health Office can take action in anticipation of an increase in dengue cases in 2015.

Index Terms- Dengue, forecast, model

1. INTRODUCTION

Dengue fever (DF) and dengue hemorrhagic fever (DHF) are caused by the dengue virus.[9] It is expanding geographic distribution of both the viruses and the mosquito vectors and increasing frequency of epidemics, development of hyperendemicity and the emergence of DHF in new areas.[10] Dengue Hemorrhagic Fever (DHF) is one of the problems in public health in tropical countries and sub-tropical countries in Southeast Asia, Central America, Latin and western pacific. It was estimated that about 500,000 cases of dengue require hospitalization each year with most sufferers are children. Figures of case fatality rate (CFR) could reach 20% if there was no proper treatment, while the numbers will decline less than 1% if cases are treated with intensive therapy.[1]

Dengue incidence continues to increase dramatically worldwide in the past decade. The actual number of dengue cases are often underreported and also many cases with wrong classification. In addition, the disease often leads to outbreak and causes a great burden for society.[2]

Indonesia is one of high suitability countries for transmission of dengue. From trend of dengue cases in Indonesia from 2002-2014 year showed that dengue fever cases occurred most commonly with the incidence rate of 71.78 per 100,000 in 2007. Dengue cases decreased very significantly to 27.67 per 100,000 people in next 4 years. This figure then rose again in 2012-2014.[4]

Since 2005, CFR dengue fever has decreased until 2010 (0.87%). However, the rate increased again in 2011 to 2014 reached 0.9%. This disease often caused outbreak in some areas and were included in the five diseases with the highest frequency to be outbreaks. The province with the highest dengue cases was namely West Java, while East Java comes out on second place with the number of cases about 8177 cases.[4] Since 2010, dengue cases were on the decline, but increased again in 2013 (14,534 cases) and decreased again in the next year.[3]

The purpose of this study was to acquire forecasting model of dengue fever cases in East Java Province for 2015. In addition, the study also aims to determine the major forecasting dengue fever cases in East Java province for a period of one year in 2015.

2. METHOD

This research used time series approach. Variable was the number of dengue cases in the province of East Java. The data used was the number of dengue cases in the province of East Java in 2005 to 2014. The unit of analysis was month that amounted to 120 months. The data obtained by documentation study. The analysis was time series using Minitab 16 application that aimed to forecast dengue cases in the province of East Java.

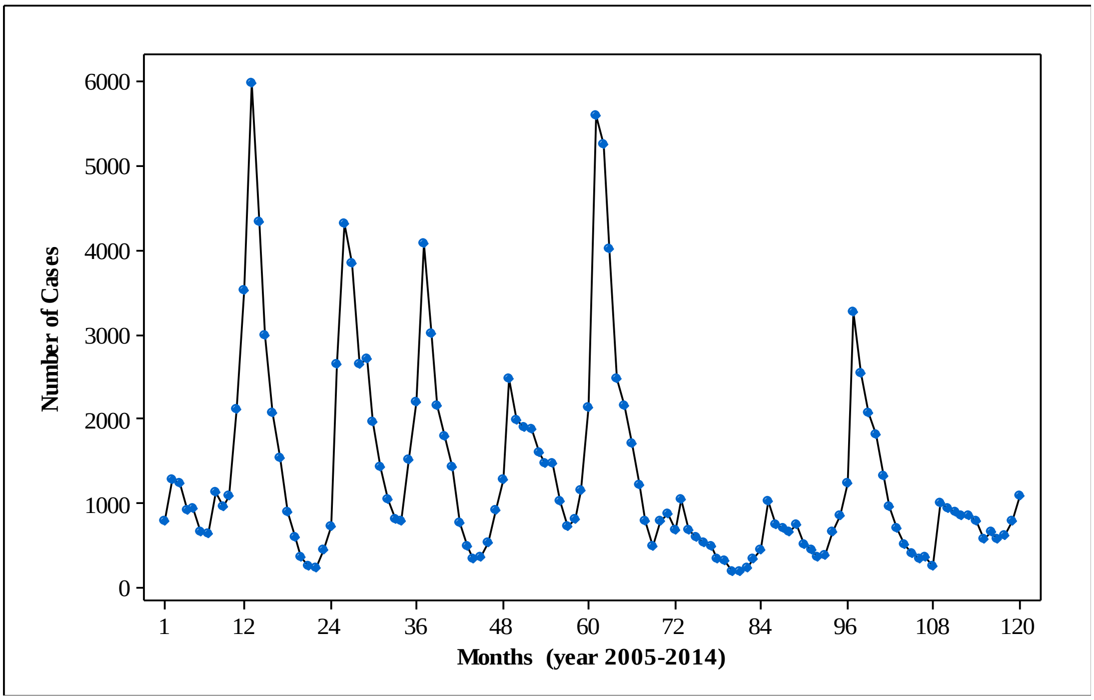

Figure 1. Plot of dengue cases from 2005 to 2014 in East Java Province, Indonesia

3. RESULT AND DISCUSSION

Trend data analysis of dengue cases (see Figure 1) showed that dengue cases had increased in multiples of 12 months and then decreased in the middle of each year. This trend includes time series are not stationary.

The data also needed to be seen its stationarity toward variance. Therefore it was necessary for the Box-Cox transformation test on the original case data. In this study it was conducted a few times of transformation. The first transformation values obtained Rounded Value of 0.00, it could be said that the data was not stationary against the variance. To make the variance stationery it was performed log transformation. After transformation was done then data was tested using Box-Cox. However, the data was still not stationary on the variance, so it was necessary to transform the data again. Based on latest data transformation was obtained Rounded Value about 1.00, it could be concluded that the data was already stationary against the variance.

Autocorrelation function (ACF) and the partial autocorrelation function (PACF) needed to be done at model identification step to determine the stationarity of time series data. From result of ACF and PACF, it looked that the data had not been stationary in mean whether seasonal and non seasonal. Because the data was still not stationary, the data must be differenced. For non seasonal variation, differencing was conducted using data after transformation and differenced with Lag 1. For seasonal variation, differencing was done using the differenced-non seasonal data then differenced again with Lag 12.

Result of ACF and PACF seemed that the data had been stationary in mean of seasonal and non seasonal. It was assumed that data was stationary at non seasonal variation because after Lag disconnected then there was no a lag further out of the boundary line.

It could be stated that the data was already well stationary on both non seasonal and seasonal. It was stationary on non seasonal variation because after Lag disconnected, no Lag else out of the boundary line in both the ACF and PACF except Lag 12, because Lag multiples of 12 showed seasonal variation. It was stationary at seasonal variation due to a multiple of 12 no existing Lag that was out of bounds line. Only in lag multiples of 12 that could be out of bounds then it was said stationary line (see Figure 2 and 3).

Because the data had been stationary in both the variance and the mean on seasonal and non seasonal, then from ACF and PACF could be determined initial model for this time series data. For non seasonal variation, it appeared that on its ACF disconnected after Lag 1 and on its PACF disconnected after Lag 1. So, for non seasonal variation, potential model was ARIMA (1,1,1), ARI (1,1) or IMA (1,1). For seasonal variation, it appeared that on its ACF disconnected after Lag 12 as well as on PACF disconnected after Lag 12. So for seasonal variation, likely model might be ARIMA $(1,1,1)^{12}$. Furthermore, models are combined into nine possible models i.e: ARIMA (1,1,1) $(1,1,1)^{12}$, ARIMA (1,1,1)

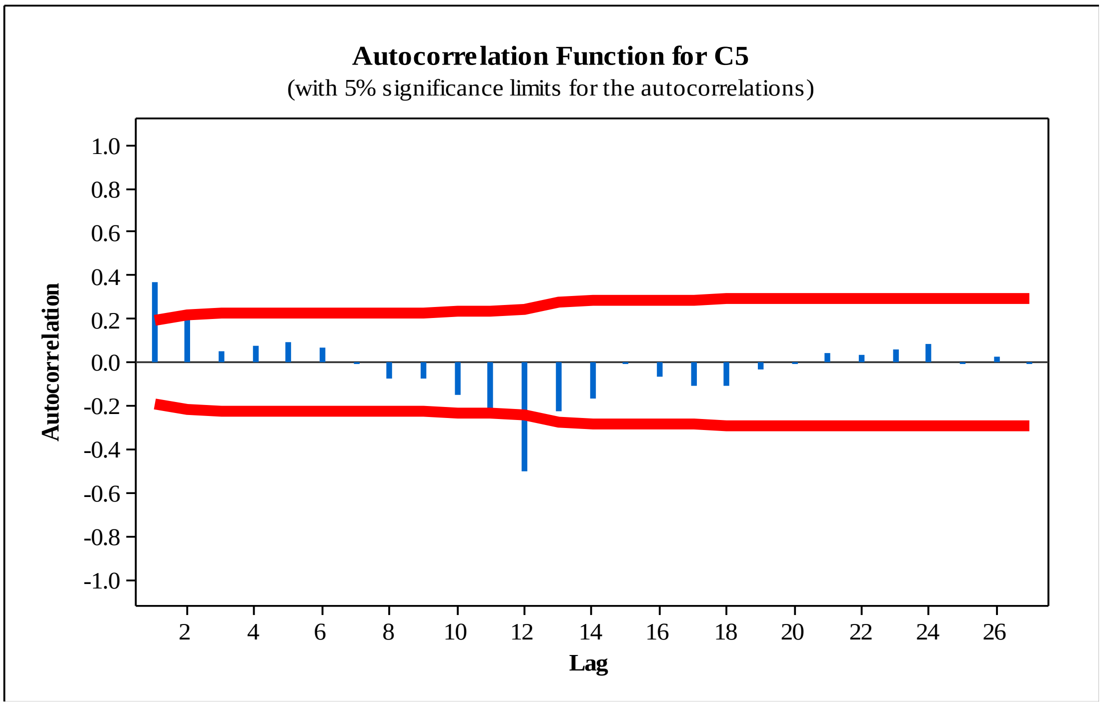

Figure 3. Forecasting Cases of dengue in 2015 in East Java Province

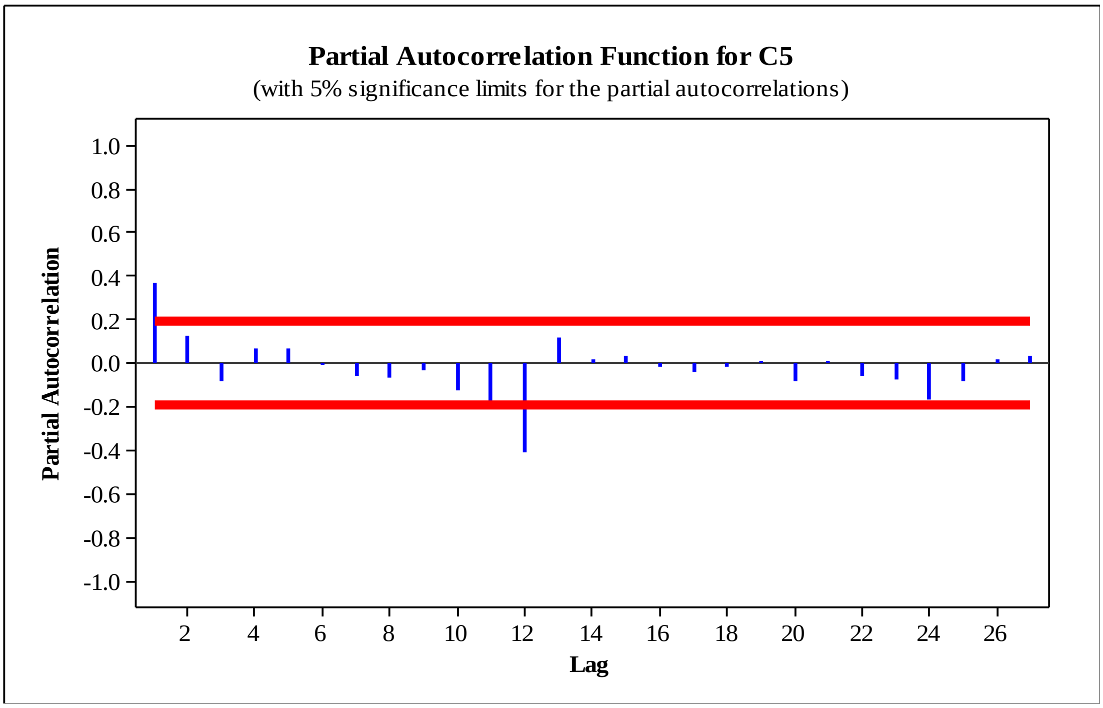

Figure 4. Forecasting Cases of dengue in 2015 in East Java Province

$(0,1,1)^{12}$, ARIMA (1, 1.1) $(1,1,0)^{12}$, ARIMA (1,1,0) $(1,1,1)^{12}$, ARIMA (1,1,0) $(0,1,1)^{12}$, ARIMA (1,1.0) $(1,1,0)^{12}$ ARIMA (0,1,1) $(1,1,1)^{12}$, ARIMA (0,1,1) $(0,1,1)^{12}$, ARIMA (0,1.1) $(1,1,0)^{12}$.

Furthermore, from 9 models are possible, looked for significant models. Next, see its Ljung-Box whether met the assumptions of white noise or not. Test of significance of the parameters and assumptions of white noise was with critical area to reject H0 if the p-value < alpha (0.05). Of the nine models which might have the best model was a model that had significant parameters could only be continued for further analysis to forecast.

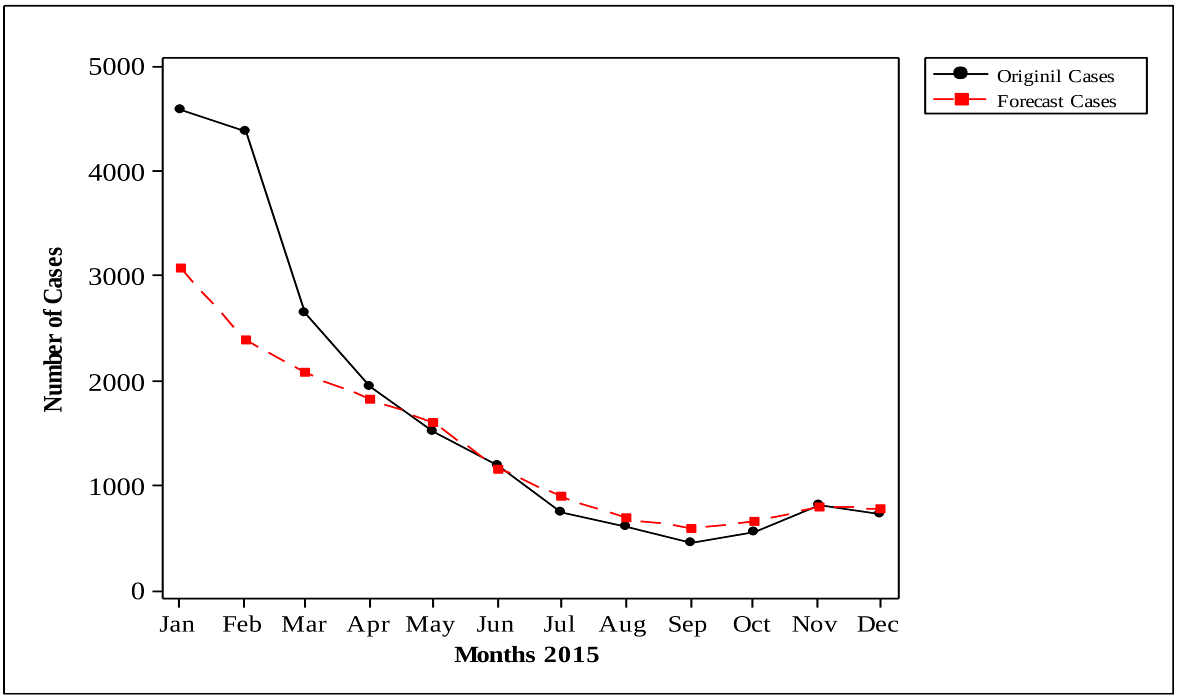

Figure 5. Forecasting Cases of dengue in 2015 in East Java Province

Model, that obtained from 9 models tested and met the best model assumptions, was ARIMA (0,1,1) $(1,1,1)^{12}$. This model was found that had a significant parameter. From Ljung-Box, it was showed there was no correlation between its Lags. Thus, this model was said to be continued for further analysis to forecasting. Of the nine models that may turn out to elect the best model that was ARIMA (0,1,1) $(1,1,1)^{12}$.

Then it was conducted test of the second assumption namely the assumption of normality residual of ARIMA $(0,1,1)(1,1,1)^{12}$. For ARIMA (0,1,1) $(1,1,1)^{12}$, it was obtained residual normality plot. It was obtained p-value 0.150, it can be inferred that residual was distributed normally. This means that the model meets the independence assumption. Because of the assumption of residual normality met then it can be continued forecasting dengue cases in East Java using the ARIMA (0,1,1) $(1,1,1)^{12}$.

Forecasting result using ARIMA $(0,1,1)(1,1,1)^{12}$ was undergone, but it must be transformed into original data because at the time of the determination of the model the initial data was performed twice transformation. Result of forecasting using ARIMA can be seen on figure 5.

Based on the result of forecasting Minitab software, it was known that the majority incidence of dengue fever in East Java occurred in January with 3073 cases, while the least incidence occurred in September with prediction about 590 cases. Also, the accuracy of forecast using mean absolute error was 403,7 cases.

4. CONCLUSION

It can be concluded that seasonal data forecasting ARIMA (0,1,1) $(1,1,1)^{12}$. The model was the best model of several possible models. The model has been chosen because has a significant parameter, white noise and met the assumption of independence. Based on the results of forecasting, preferably East Java Provincial Health Office can take action in anticipation of an increase in dengue cases in 2016.

Acknowledgments

Thanks to East Java Provincial Health Office for providing dengue cases data on this research.

REFERENCES

[1] WHO. (2006): Dengue Haemorrhagic Fever: early recognition, diagnosis and hospital management. Geneva: Department of Epidemi and Pandemic Alert and Response.
[2] WHO. (2012): Global Strategy for Dengue Prevention and Control. Geneva, Switzerland; World Health Organization.
[3] East Java Provincial Health Office. Dengue cases report in 2005-2015.
[4] Ministry of health. (2014): Health Profile of Indonesia 2014.

[5] Baroroh, Ali. (2013): Analisis Multivariate dan Time Series dengan SPSS 21. Kompas Gramedia; Jakarta.
[6] Choudhury, Z. M., Banu, S., & Islam, A. M. (2008): Forecasting dengue incidence in Dhaka, Bangladesh: A time series analysis.
[7] Wei, W. W. S. (1994): *Time series analysis*. Reading: Addison-Wesley publ.
[8] Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
[9] Srichaikul, T., & Nimmannitya, S. (2000): Haematology in dengue and dengue haemorrhagic fever. *Best Practice & Research Clinical Haematology*, *13*(2), 261-276.
[10] Gubler, D. J. (1998): The global pandemic of dengue/dengue haemorrhagic fever: current status and prospects for the future. *Annals of the Academy of Medicine, Singapore*, *27*(2), 227-234.