Original Research

EPIDEMIOLOGY FORECASTING ANALYSIS OF DENGUE HAEMORRHAGIC FEVER WITH SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE IN TROPICAL AREA

Siswanto*, Risva, Nana Marliana

Epidemiology Department, Public Health Faculty, Mulawarman University, Indonesia

Received: 11 April 2019 | Revised: 9 May 2019 | Accepted: 2 June 2019

Correspondence:
Siswanto
Epidemiology Department, Public Health Faculty, Mulawarman University
South Sempaja, North Samarinda, Samarinda City, East Kalimantan 75242, Indonesia
Email: siswanto2013@gmail.com

Copyright: © the author(s), YCAB publisher and Public Health of Indonesia. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: Health problems that often occur in tropical countries are infectious diseases, one of which often causes outbreaks was Dengue Hemorrhagic Fever (DHF). This disease often causes problems especially in endemic areas and even outbreaks that occur with death from sufferers.

Objectives: To forecasting of the Dengue Hemorrhagic Fever in the working area of the Puskesmas Temindung.

Methods: This was analytical descriptive research with forecasting design using secondary data and primary from informant who understand the problem. Forecasting using SARIMA method (Seasonal Autoregressive Integrated Moving Average).

Results: The results showed that the total of DHF cases in Temindung Health Center could be predicted by the SARIMA (1,1,1) (1,0,0) model with means square error (MSE) of 0.001394688 forecasting results obtained from October 2018 to September 2019 cases, which tend to fluctuate but illustrates an increase in cases of DHF compared to the previous year's data.

Conclusion: Forecast of the DHF is for the next 12 months starting from October 2018 as many as 7 cases, in November 4 cases, in December 4 cases; then starting in January 2019 as many as 3 cases, February 2 cases, March 3 cases, April 3 cases, May 3 cases, June 4 cases, July 3 cases, August 3 cases and September 3 cases with a total number of 42. Forecasting results show dengue cases tend to fluctuate every month but have increased cases from the previous year.

Keywords: SARIMA, dengue hemorrhagic fever, tropical area

BACKGROUND

Health problems that often occur in tropical countries are infectious diseases that are closely related to environmental conditions (Gubler, 2002; Tosepu, 2017). One infectious disease that often causes outbreaks was Dengue Hemorrhagic Fever (Ishak & Kasman, 2018; Rahman et al., 2002). This disease often causes problems especially in endemic areas and even outbreaks that occur with death from sufferers. The losses incurred are not only material but also material.

Incidence rate of DHF in East Kalimantan endemic areas in 2017 has 918 cases with 25.68%. And this case was highest in Samarinda City with 2,814 people with Case Fatality Rate of 18 people. Based on the data, the Puskesmas work area with the highest cases of DHF was in Puskesmas of Temindung with 41 cases (Office, 2016).

Based on the problem, an analysis was needed using the forecasting method. That was to...
predict problems and what should be prepared and carried out for prevention, handling and control. Mathematical and statistical models provide a substantial contribution in understanding the dynamics of dengue fever transmission and the growing trend of cases of disease. There are several types of forecasting methods used, one of which was the Time Series method that was used to predict the future based on past values and variables in order to find patterns in the data and explore the future (Busenberg & Cooke, 2012; Vynnycky & White, 2010).

Time series characteristics, which are stationary, seasonal, etc. require a systematic approach to obtain an overview of basic models, one of which was the SARIMA method (Seasonal Autoregressive Integrated Moving Average). This seasonal model was useful in situations where time series data shows repeated periodic fluctuations with the same intensity each year (Hota, 2014; Khashei & Bijari, 2011).

The characteristics of the SARIMA model are adequate or appropriate for studies of monthly dengue fever data, so it can be seen the results of forecasting the number of cases of dengue fever in a population tend to be the subject for seasonal variations, looking at maximum data in the rainy season and minimum during the dry season (Martinez et al., 2011).

Epidemiological forecasting analysis was a method of analyzing health problems using module 1 problem 100 solutions in a holistic and comprehensive manner with the principles of Low Cost, High Impact and Continuous. The problem of dengue in endemic areas from the past until now needs forecasting analysis that looks at all aspects.

METHODS

Study design
This research was analytical descriptive research with forecasting design using secondary data from monthly reports on dengue and primary cases from interviews of informants who understand the problem. Descriptive analytic research is research that serves to describe or give an overview of the object under study through data or samples that have been collected, take a problem or focus on problems as they are when the research is carried out, the results of the research are then processed and analyzed to draw conclusions for see and explore how health phenomena occur.

Sample
The sample was secondary data from Puskesmas of Temindung in Samarinda City. We have used totally monthly dengue hemorrhagic fever cases at Puskesmas of Temindung from 2014 to 2018.

Data analysis
Method of forecasting data analysis using software Minitab Pro 16.1.0.0, then from the data of the last 5 years, data was input into one of the forecasting software. Furthermore, stage determines the ARIMA model by using the Box-Jenkins procedure. Forecasting results were analyzed by epidemiological forecasting analysis using 1 problem 100 solutions module. This are the results of the analysis using the Epidemiological Strategy and are aligned with the strategies of the scientific Health Policy Administration and Health Promotion. Module 1 Problem 100 Solution is a module that contains methods in making planning based on evidence using the principles of Low Cost, High Impact and Continuous.

RESULTS

Number of DHF Cases in Puskesmas of Temindung Samarinda

<table>
<thead>
<tr>
<th>Months</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>10</td>
<td>5</td>
<td>13</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>February</td>
<td>3</td>
<td>21</td>
<td>6</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>March</td>
<td>1</td>
<td>9</td>
<td>12</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>April</td>
<td>1</td>
<td>0</td>
<td>14</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>May</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>June</td>
<td>2</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Data Calculation Using SARIMA Method

The formation of data models for dengue cases in Puskesmas of Temindung is by looking at the results of the ACF and PACF charts. Where obtained 10 temporary models namely ARIMA (1.1.0), ARIMA (1.1.1), ARIMA (2.1.1), ARIMA (0.1.1), ARIMA (2.1.0), ARIMA (1.2.1), ARIMA (1.2.0), ARIMA (0.2.1), ARIMA (2.2.0), and ARIMA (2.2.1).

Based on the test results using parameter significance, white noise and residual normality, it was known that 10 temporary models do not fulfill the three tests. This was caused by the presence of seasonal data in cases of DHF, so it was concluded that the use of the ARIMA model was not appropriate in this data. In accordance with statistic procedures, a re-modeling was carried out for data on the incidence of DHF using the SARIMA model.

There was two identification of data stationarity in the SARIMA model with non-seasonal and seasonal stationarity. Formation of time series charts at the stage of identification of non-seasonal stationarity data on cases of dengue in Puskesmas of Temindung in averages and variances can be seen in Figure 1.

Based on the data in Figure 1 it was said that the data was not stationary both in average and in non-seasonal variances. This was because there was a change in the average time and variance (distribution of data). So the data must be stationary first. Furthermore, to stationary data that was not stationary in the mean and variance (distribution of data), the parameter value estimation $\lambda$ was carried out by carrying out Box-Cox transformations and differencing methods. Data was re-identified by forming a time series graph (Figure 2).
Based on Figure 2 it can be seen that the pattern of data spreads around the average, so it was concluded that twice the data differencing have been stationary in non-seasonal averages.

![Figure 2](image)

**Figure 2** Graph of ACF data on dengue incidence of cases

Based on Figure 3 it can be seen that the ACF value was cut off in lag 1 and after lag 1 drops rapidly. So it was concluded that the data has been stationary in averages and non-seasonal variances. Seasonal Stationary Identification

Based on Figure 3, it can be seen that the ACF chart have a pattern of cut-off in lag 4 and falls rapidly after lag 4, so the data was stationary in the mean and seasonal variance.

**Identification of the SARIMA Model**

Based on Figure 3 it can be seen that the data has a seasonal period of \( s = 4 \), and based on Figure 4.6 it was known that the order for non-seasonal MA was 0 and 1 (according to the ACF cut off value in lag 1). It was known that the order for Non-seasonal AR was 0.1 and 2 (according to PACF cut-off values in lags 1 and 2), with differencing \( (d) \) 2 times.

![Figure 3](image)

**Figure 3** Graph of ACF data on dengue incidence of cases

![Figure 4](image)

**Figure 4** PACF graph of data on seasonal dengue cases is 1, and seasonal differencing order \((D)\) is 0, \( D = 0 \) because there was no differencing process for seasonal.

Public Health of Indonesia, Volume 5, Issue 2, April - June 2019

41
Table 2 The combination of SARIMA Model

<table>
<thead>
<tr>
<th>Models</th>
<th>Models</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA (1,1,0)(1,0,1)^4</td>
<td>SARIMA (1,1,0)(1,0,0)^4</td>
<td>SARIMA (1,1,0)(0,0,1)^4</td>
</tr>
<tr>
<td>SARIMA (1,1,0)(1,0,1)^4</td>
<td>SARIMA (1,1,0)(1,0,0)^4</td>
<td>SARIMA (1,1,0)(0,0,1)^4</td>
</tr>
<tr>
<td>SARIMA (1,1,0)(1,0,1)^4</td>
<td>SARIMA (1,1,0)(1,0,0)^4</td>
<td>SARIMA (1,1,0)(0,0,1)^4</td>
</tr>
<tr>
<td>SARIMA (1,1,0)(1,0,1)^4</td>
<td>SARIMA (1,1,0)(1,0,0)^4</td>
<td>SARIMA (1,1,0)(0,0,1)^4</td>
</tr>
</tbody>
</table>

Based on table 2 it can be concluded that there was thirty temporary models obtained. Where from the thirty models will be tested the significance of parameters, White Noise and Residual Normality.

Selection of the best model

Determination of the best model between the two SARIMA models that have fulfilled the requirements in testing parameter significance, white noise and residual normality can be done by looking at the MSE (Means Square Error) value of the two models. The best model was a model that has the smallest MSE value. The MSE value of the SARIMA (1,1,0) (1,0,0) 4 model is 0.001394688 (0.0014) and the SARIMA (1,2,0) (1,0,0) 4 model was 0.001974108 (0.0019), so that it can be concluded that the SARIMA (1,1,0) (1,0,0) 4 model was a more appropriate model for predicting the number of cases of DHF in Puskesmas of Temindung.

Forecasting

The results of forecasting cases of dengue in Puskesmas of Temindung for October 2018 to September 2019 was shown in table 3. Based on Figure 4, it was known that the ACF chart was cut off on lag 4, so that the order for the seasonal MA was 1. Based on Figure 4, it can be seen that the cut-off chart was in lag 4, so the order for AR

Table 3 Results of Forecasting Cases of DHF

<table>
<thead>
<tr>
<th>Months</th>
<th>Results of Forecasting</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 2018</td>
<td>7</td>
</tr>
<tr>
<td>November 2018</td>
<td>4</td>
</tr>
<tr>
<td>December 2018</td>
<td>4</td>
</tr>
<tr>
<td>January 2019</td>
<td>3</td>
</tr>
<tr>
<td>February 2019</td>
<td>2</td>
</tr>
<tr>
<td>March 2019</td>
<td>3</td>
</tr>
<tr>
<td>April 2019</td>
<td>3</td>
</tr>
<tr>
<td>May 2019</td>
<td>3</td>
</tr>
<tr>
<td>June 2019</td>
<td>4</td>
</tr>
<tr>
<td>July 2019</td>
<td>3</td>
</tr>
<tr>
<td>August 2019</td>
<td>3</td>
</tr>
<tr>
<td>September 2019</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
</tr>
</tbody>
</table>

From table 3 it can be concluded that during October 2018 - September 2019 there was an illustration of the number of cases of 42 cases. Based on the results of forecasting that has been carried out there are 3 months of case data that are used as differences with the results in the field, namely in October, November and December as follows.

Table 4 Results of Forecasting and Number of DHF Cases in 2018

<table>
<thead>
<tr>
<th>Months</th>
<th>Results of forecasting year 2018</th>
<th>Number of cases DHF year 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>November</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>December</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Based on table 4, it can be concluded that the results of forecasting that have been carried out there was a match between the number of cases with forecasting in November and December and there are differences in the number of cases with the results of forecasting in October which caused the value of MSE. The following are the results of the analysis and discussion of the gap at the end of the month of 2018, beginning of the month and mid-month of 2019.
DISCUSSION

The table above shows that forecasting results that use forecasting software are fluctuating for the next 12 months. From table it can be seen the comparison of forecasting data with the previous year's data shows an increase in the number of sufferers of DHF cases in Puskesmas of Temindung. The description of forecasting results was an increase in cases from the previous year where in October 2017 - September 2018 there were 32 cases that occurred while based on the results of forecasting in October 2018 - September 2019 amounted to 42 cases.

Based on statistic theory the prediction was basically an estimate or estimate of the occurrence of an event at a time when it will come (Supranto, 1983). In this quantitative forecasting study using the SARIMA method because it was seen based on the results of the graph of the data of the last year case of DHF showing seasonal movements. This was in line with Milasari's (2010) study that the branch of statistic science using forecasting methods can be applied to life. In DHF cases where in the rainy season the development of Aedes Aegypty mosquitoes increases and results in an increase in the number of people with Dengue Fever. So that the right forecasting model was the Seasonal ARIMA models.

The SARIMA method was used to measure seasonal data using seasonal indices to calculate time series data forecasts (Tseng et al., 2001). In ARIMA forecasting with seasonal aspects there are general notations used, namely (p, d, q), (P, D, Q) s. In this forecast, the MSE value was 0.001394688 (0.0014) from the forecast that have been done.

In a forecasting there was several disadvantages, one of which was the existence of a gap (difference in data). There are several factors that can influence the emergence of gaps, some of which are due to external variables such as environmental aspects (natural disasters, weather), population aspects (demographic conditions), behavioral aspects (migration), service aspects (prevention efforts) and so on so that there are differences that affect the results of forecasting with the results in the field.

Based on data at the end of 2017 there was 1 case from the puskesmas. Forecasting results at the same time as many as 7 cases. Then after being seen from cases that have passed as many as 4 cases. From the results of forecasting there was a decrease in the number of cases with the fact that this was influenced by the presence of self-awareness of the incidence of DHF by increasing prevention programs that have been carried out from the previous month by the DHF program holders. Efforts made in the form of Epidemiology Investigation, Indonesia call was Epidemiologi Investigation and fogging and Indonesia call was Eradication of Mosquito Nests directly after there are reports of dengue cases from Indonesia call was household or local residents.

From the results of the interviews to validate the results of the research analysts with the holders of the DHF program at the Puskesmas of Temindung, it was found that there were several risk factors that were still present which caused an increase in cases based on the results of forecasting with the previous year's data. This was because of the population aspect with a high level of mobilization in endemic in adjacent puskesmas, service aspects (prevention efforts) with limited number of program holders and population ratio.

Some of the reasons stated were that they were busy working and collaborating for environmental hygiene. In addition, the activities of the community was more busy in matters of work and do not have time to implement to drain, close, bury and Eradications of Mosquito Nests programs. This was in accordance with analysis from the aspects of people's behavior that do not maintain environmental cleanliness.

This was supported by research with the results of the study there was a significant
relationship between 3M Plus PSN behavior towards the incidence of DHF (Priesley, 2018). In addition, based on observations in several homes, there were still many houses that have water reservoirs both inside and outside the house of unused items that have the potential to be a breeding place. This was in accordance with the larva-free number Indonesia call was free numbers of flick data which was carried out every month which is reached 74-79% where the data is still far from the target of Puskesmas of Temindung which was 95%.

Based on the results of the forecast and reality data there are differences in the results of the forecast of 7 cases with the results of the reality of 4 cases. This difference was influenced by the level of error of the Means Square Error (MSE) calculation of 0.001394688 (0.0014) with the MAPE value of 29.283%. This was in accordance with the principle theory of forecasting according to Sofyan (2015) that forecasting involves errors, forecasting only reduces uncertainty but does not eliminate it, almost never found that the forecasting results are like reality in the field and forecasting must include the size of the error because forecasting always contains error, the user needs to know the error amount in the form of range in units or percentages (Sofyan, 2015).

At the beginning of 2018 there were 0 dengue cases. Based on the results of forecasting there were 3 cases. From these results it appears that there are differences, namely there was an increase in the number of forecast cases that occur. This happens because of several risk factors. Among them are from the environmental aspects, namely climate conditions (rainy season) that occur at the beginning of the month so that the potential for an increase in the number of breeding places both inside and outside the house is unknown (Fathi, 2005). The rainy and transition season that happened at that time was due to the west monsoon which had an impact on the increase in rainfall so that the breeding of the Aedes Aegypti mosquito also increased. This was supported by research by Wirayoga (2013) in 2006-2011 where the results of the study explain that air temperature, rainfall and humidity have a significant relationship to the incidence of dengue hemorrhagic fever (Wirayoga, 2013).

Based on case report data in the Puskesmas of Temindung area in the middle of 2018 there were 5 cases of DHF. Based on the results of forecasting there are as many as 3 cases. From the results of the data and predictions, there was a picture of the decline in dengue cases that occur. From the results of interviews with DHF program holders to validate the results, this was influenced by several factors, one of which was the aspect of service (prevention, handling and control) carried out by the Puskesmas better quality and quantity than before. Efforts made to improve programs such as fogging directly to the source are indicative of the presence of the Aedes mosquito.

In addition, by promoting Epidemiological Investigation directly, there are case reports and follow up actions to reduce the spread of dengue cases. Such as counseling on prevention of dengue and distribution of abatitation in adjacent areas with a radius of 200 meters from the homes of patients with DHF reported from residents or head of the household.

With the increase in the prevention, handling and control of dengue events that have been carried out by health workers, especially in the Puskesmas of Temindung work area so that early self-awareness Indonesia call is SKD carried out by Puskesmas of Temindung was better which causes a decrease in the number of dengue events from the forecast results, this obtained from the results of the quota conducted by researchers with the DHF program holders. In addition to increasing prevention, handling and control efforts by the puskesmas, the support of technological advancements, Indonesia namely SIKDA (Regional Health Information System) was very helpful, namely an application for collecting data on predictions or effectiveness.
of SKD (early alertness self-control) which was very useful for controlling DHF.

From the results of forecasting there was mid-2018 there is an illustration of the increase in the number of cases from the results of the forecast with the previous year. Where in 2018 was 1 case. Based on the results of forecasting in mid-2019 there are 4 cases. Counseling about DHF and information, and if there are reports from local residents or from the head of the household Puskesmas of Temindung officers.

From the results of the forecasting, it can be seen that the case picture have increased even though prevention efforts have been carried out by health personnel, but there are still cases that occur because of several risk factors including population aspects where the level of community mobilization was high. the interaction was still within the DHF endemic area so that it have the potential to be in the Aedes mosquito feeding place.

Other aspects of population are population density which was one of the risk factors for the incidence of dengue. Population density was related to the distance of flying mosquitoes and transmission of DHF, where the more densely populated, the easier the transmission of dengue was due to the flight distance of mosquitoes was estimated to be less than 200 meters between sufferers one with other sufferers.

And from environmental factors where Samarinda city, including the humid tropical rainforest region with relatively high rainfall, was the endemic of DHF. So that breeding places were scattered everywhere. Several observations from a number of Aedes larvae were found in gutters, zinc roofs, bottles, cans, used pots, tarps or former waterlogged banners with a surface height of more than 1 cm.

In addition, the coverage area of Puskesmas of Temindung was prone and endemic areas where there were always dengue cases every year in the area because the area was a collection of used goods so that there are many containers that can hold rainwater into Aedes Aegypti mosquito breeding grounds, This was supported by Gamma's research (2010) where the presence of containers > 3 was the risk factor for dengue fever, the risk was greater than those with containers <3 (Gama, 2010).

Based on the results of interviews with DHF program holders, this was supported by aspects of community behavior where most residents in the protected health center work area possessed rainwater shelters where such shelters were rarely drained so they could be at risk of becoming mosquito breeding grounds for DHF. also with the research conducted by Carundeng (2015) which explains that the results of the research carried out there was a relationship between the habit of draining water reservoirs with the incidence of dengue hemorrhagic fever / p = 0.000 (p <0.05) and OR = 5.9 (CI = 2.137- 16.342) (Carundeng, 2015).

Another opinion that supports the Susmaneli Research (2011) states that the results of multivariate analysis showed a significant related variable was the water reservoir OR = 3.849 (95% CI: 2.399-6175), availability of water reservoir cover OR = 2.248 (95% CI: 1.403- 3.603), frequency of drainage of water reservoir OR = 2.238 (95% CI: 1.399-3.579), house density OR = 4.049 (95% CI: 2.486-6.596), age OR = 2.845 (95% CI: 1.768-4.577), brainwashing sex OR = 0.613 (95% CI: 0.379-0.992) (Susmaneli, 2011).

Based on the previous discussion, it was found that the forecasting results that had been carried out using forecasting software showed that dengue cases were predicted to have an increase in cases of dengue when compared with the previous year's data, factors that influence the incidence of dengue in the region are due to some aspects, namely aspects of population, aspects of community behavior, aspects of health services and environmental aspects in the endemwas region where the occurrence of dengue cases in the region occurs every year. And there is still the attitude of the people who are less concerned about the cleanliness of their environment so they do not have time to implement the 3M +
and PSN program because they are busy working. So that these factors support the increase or continuous presence of dengue cases and breeding sites for Aedes mosquitoes. Coupled with the existence of collectors (collecting used goods) so that they become mosquito breeding grounds.

The results of forecasting for the next 12 months starting from October 2018 to September 2019 are 7, 4, 4, 3, 2, 3, 3, 4, 3, 3, 3, and 3 with amounts of 42 with MSE values of 0.001394688 (0.0014) and MAPE of 29.283%. The forecasting results show that DHF cases tend to fluctuate but have an increase in the number of dengue cases that will occur in 2019 compared to the data in 2018.

CONCLUSION

The data plot illustrates that the data have been stationary in averages and variances after data transformation have been carried out. And differencing was done twice from data that have been stationary in non-seasonal averages and was not carried out differencing in seasonal data. From the ACF and PACF charts, the order for non-seasonal MA was 0 and 1 (ACF cut off value in lag 1), for non-seasonal AR was 0.1 and 2, with differencing (d) 2 times. Seasonal AR was 1 and the seasonal differencing order (D) was 0 so that the differencing process was not seasonal. Determination of the best model was done by using parameter significance testing, white noise and residual normality based on Box-Jenkins procedures. Determination of the best model for forecasting DHF cases in puskesmas of Temindung in the next 12 months was to use the SARIMA method (according to the Box-Jenkins methodological flow chart). This best SARIMA model has the smallest MSE value in the SARIMA (1,1,1) (1,0,0) model 4. Forecast results for the next 12 months starting from October 2018 as many as 7 cases, in November 4 cases, in December 4 cases. Then starting in January 2019 as many as 3 cases, February 2 cases, March 3 cases, April 3 cases, May 3 cases, June 4 cases, July 3 cases, August 3 cases and September 3 cases with a total number of 42. Forecasting results show dengue cases tend to fluctuate every month but have increased cases from the previous year.

REFERENCES


