A Model for Forecasting the Number of Cases and Distribution Pattern of Dengue Hemorrhagic Fever in Indonesia

Deni Mahdiana
Faculty of Mathematics and Natural Sciences
Universitas Gadjah Mada Yogyakarta, Indonesia and
Faculty of Information Technology Universitas Budi Luhur
Jakarta, Indonesia

Edi Winarko
Faculty of Mathematics and Natural Sciences
Universitas Gadjah Mada
Yogyakarta, Indonesia

Ahmad Ashari
Faculty of Mathematics and Natural Sciences
Universitas Gadjah Mada
Yogyakarta, Indonesia

Hari Kusnanto
Faculty of Medicine
Universitas Gadjah Mada
Yogyakarta, Indonesia

Abstract—**Dengue Hemorrhagic Fever (DHF) ourbreaks is one of the lethal health problems in Indonesia. Aedes aegypti type of insect prolefiration as the main vector of DHF has affected climate factors, such as temperature, humidity, rainfall, and irradiation time. Therefore, to project the number of DHF cases is a very important assignment for the Ministry of Health to initiate contingencies planning as a prevention step in confronting the increasing number of DHF cases in nearby future. This study aims in developing a forecasting model in anticipating the number of cases and distribution pattern of DHF with multivariate time series using Vector Autoregressive Spatial Autocorrelation (VARSA). VARSA model uses multivariate time series, such as a number of DHF case, minimum temperature, maximum temperature, rainfall, average humidity, irradiation time and population density. This modeling is done in two steps: Vector Autoregressive modeling to predict the number of DHF cases and Local Indicators of Spatial Association (LISA) method to visualize distribution pattern of DHF based on the spatial connectivity of the number of DHF cases among the neighboring districts. This study covers 17 districts in Sleman Yogyakarta, resulting in low errors with Root Means Square Error (RMSE) of 2.10 and Mean Absolute Error (MAE) of 1.51. This model produces smaller errors than using univariate time series methods, such as Linear regression and Autoregressive Integrated Moving Average (ARIMA).**

Keywords—Dengue Hemorrhagic Fever (DHF); Vector Autoregressive Spatial Autocorrelation (VARSA); forecasting; multivariate time series; Local Indicators of Spatial Association (LISA)

I. INTRODUCTION

Dengue Hemorrhagic Fever (DHF) is an acute and endemic disease which periodically causes outbreaks and even epidemics. DHF is widely found in tropical and sub-tropical areas [1]. Data from around the world shows Asia ranks first in the number of DHF patients each year. Meanwhile, from 1968 to 2009, the World Health Organization (WHO) listed Indonesia as the country with the highest dengue fever case in Southeast Asia [2].

DHF is a disease caused by dengue virus that is transmitted from person to person through the bite of aedes aegypti mosquitoes [3]. This disease is caused by the dengue virus of the genus flavivirus, flaviviridae family. Dengue transmitting mosquitoes are present in almost all corners of Indonesia, except in places with an altitude of more than 1000 meters above sea level. Several factors that influence the occurrence of DHF include low immune status and a population density of infectious mosquitoes due to mosquito breeding places that usually occur in the rainy season [4].

The number of DHF cases in 34 provinces in Indonesia based on data from the Ministry of Health in 2012 recorded as many as 90,425 cases and 816 people died. In 2013 recorded as many as 112,511 cases and 871 people died, In the year 2014 recorded as many as 71,668 case, and 641 people died , In the year 2015 recorded as many as 129,650 case, and 1,071 people died [5].

Preventive program namely the eradication of mosquito nest has been widely conducted nationally and regionally. This preventive activity is carried out by draining, closing, burying water reservoir and applying larvicide, keeping larvae fish and using a mosquito net, examination and eradication of larvae periodically no more than three months, and fumigation.

Nevertheless, those preventive actions have not been able to reduce the number of DHF patients nationally. The low ability in preventing of dengue fever is due to some factors. The first is the unpredictable time, place and number of events. The second is the unavailability of index and vulnerability maps of the region based on the time of the incident. The third is the unavailability of reliable model for forecasting DHF cases. Forecasting the number of DHF cases is very important for public health service to anticipate

affirmative planning for the prevention of the increasing number of DHF cases in the nearby future. Anticipated planning is needed to avoid problems such as unnecessaries or delayed precautions, or even with the increased casualties and lack of spaces and personnel for handling DHF patients.

In previous related studies the process of forecasting numbers of DHF cases were done by using ARIMA method [6] and SARIMA method [7]. Unfortunately, those studies focused on predicting the number of DHF cases without considering the spatial connectivity between an area and its neighborhoods. Moreover, this study aims to predict the number of case and distribution pattern of DHF using a combination of vector autoregressive and spatial autocorrelation (VARSA) model. This study is expected to help the public health service, especially for Sleman regency, Yogyakarta in the preventing the increasing numbers of Dengue Hemorrhagic Fever cases in the future.

The remaining of this paper is organized as follows. Section two deals with previous studies related to this study. The proposed method is presented in section three. The experiments study are discussed in section four, and the last section contains conclusion and recommendation for the future work.

II. RELATED WORKS

The purpose of applying time series, in general, is to understand future behavior through measurement of attribute data in a past time series by using trend, cyclic and seasonal indicators. [8]-[12]. The study of forecasting the number of dengue fever cases using the ARIMA method has been done by [13]-[16] and the SARIMA Method [17]-[21].

The study using the VAR method to predict the number of DHF cases has been done by [22]. The main advantage of VAR is that multivariate variables are both explained and explanatory variables. Hence, this model performs more accurate predictions using the relations between multiple variables [23].

This model is very popular in economics and explains the underlying causal mechanisms by using granger causality test [24], [25]. This test determines whether past variables can provide predictive information.

Forecasting time series analysis has been a major tool in many applications in meteorological phenomena, such as rainfall, humidity, and temperature [26], [27]. The development of the forecasting model by estimating the weight parameters by the least squares method on the hybridizing exponential smoothing model with neural network does not give optimal results [28]. This is because square errors in the emerging model will shift the fit curve to another point to reduce the accuracy of the forecasting result.

From the forecasting field review, the number of dengue cases can be analyzed that the application of the VAR method for forecasting the number of DHF case short-term and long-term (one or twelve periods) has good accuracy with low errors and meets the parsimony principle (simplicity) to be applied.

III. PROPOSED METHOD

Fig. 1 shows the proposed forecasting the number of case and distribution pattern of DHF using the VARSA model. The model starts from the input data in the form of climate, a number of DHF case and population density data. The climate variable data consists of such as minimum temperature, maximum temperature, average temperature, rainfall, average humidity and irradiation time. All data then do the preprocessing to eliminate the missing value and equalize the data period. The preprocessed data is then processed to forecast the number of DHF cases using the VAR method. The model is then tested using the RMSE and MAE methods, to see prediction errors compared to actual data. Forecasting results are then visualized using spatial autocorrelation by obtaining a digital district map data in Sleman. Visualization in this study using Local Indicator Spatial Association (LISA) method to see patterns of the distribution pattern of DHF case based on spatial connectivity among neighboring districts.
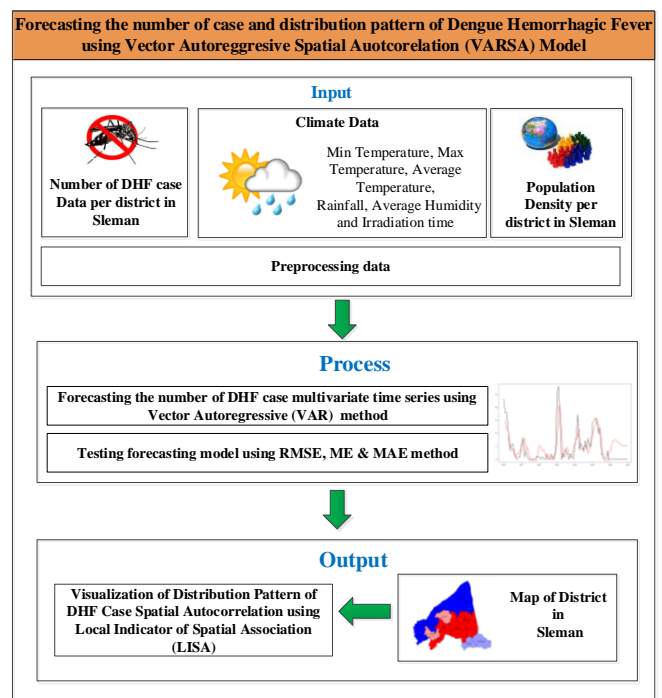

Fig. 1. Forecasting the number of case and distribution pattern of DHF using VARSA model.

Fig. 2 shows the research steps of forecasting the number of cases and distribution pattern of DHF using VARSA model. This model consists of four steps namely prepreprocessing data, create forecasting model, evaluation forecasting model and vizualitation distribution pattern of DHF wtih spatial autocorrelation.

In the first step, the preprocessing data, there are three kinds of data, namely, climate data, the number of DHF cases and population density data. Those data were collected from the district in Sleman, Yogyakarta-Indonesia, from 2010 to 2015. The climate data consist of minimum temperature,

maximum temperature, average temperature, humidity, rainfall and irradiation time variable. Firstly, the activities of climate data preprocessing are replacing the missing value using interpolation smoothing spline method, the correlation test to find out the influential climate variables using Keyser Meyer Olkin (KMO) method, and aggregating daily data of those influential climate variables into monthly. Secondly, the number of DHF cases is preprocessed by aggregation of data per incident into monthly data per district. Thirdly, the preprocessing of population density data is done by dividing the yearly data into monthly data per district.

After all data are clean and have the same period, then the second step is to build multivariate time series forecasting model using VAR method. The third step is to evaluate the forecasting model using MAE and RMSE method compared with actual data. The fourth step is to visualize the distribution pattern of DHF cases using Local Indicator of Spatial Association (LISA) method. This method is used to visualize the distribution patterns of DHF cases based on spatial connectivity between one district to its neighbors.
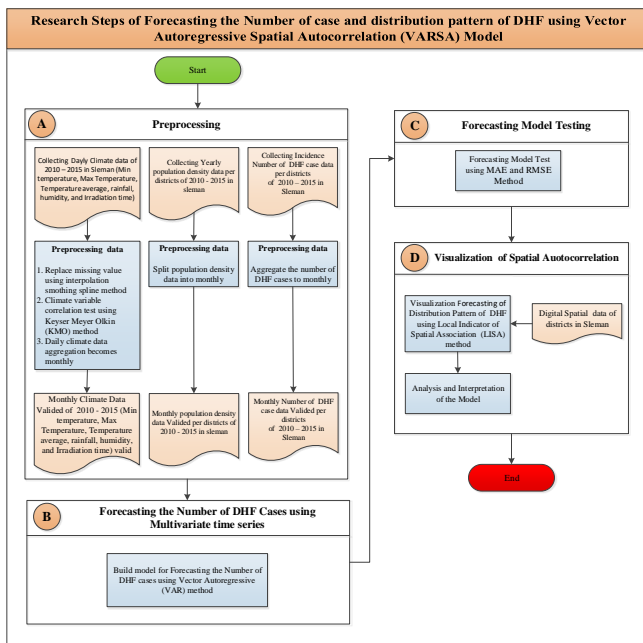

Fig. 2. Research steps of forecasting the number of case and distribution pattern of DHF using VARSA model.

A. Preprocessing

The preprocessing the climate data in Sleman regency Yogyakarta, the data used in this research is sourced from http://dataonline.bmkg.go.id the result of post measurement of geophysical station Yogyakarta. The research uses the district climate data (with a monthly period) from 2010 to 2015 consisting of minimum temperature, maximum temperature, average temperature, humidity, rainfall and irradiation time. The use of monthly data for climate forecasting is based on literature studies [10], [11], [13]. The more periods of the month used the data patterns will be more easily analyzed to determine the correct forecasting method. The climate data consisting of 17 districts in the Sleman included Moyudan,

Minggir, Seyegan, Godean, Gamping, Mlati, Depok, Berbah, Prambanan, Kalasan, Ngemplak, Ngaglik, Sleman, Tempel, Turi, Pakem, and Cangkringan.

The preprocessing data experiments were performed for all missing values of the total of 2,037 climate data. The missing value consists of unmeasured data reaching 5.1% for minimum temperature, 8.20% for maximum temperature, 17.43% for average temperature, 17.48% for average humidity, 3.29% for rainfall and 3.58% for time irradiation. From the results of climate data analysis is required preprocessing stage to obtain data quality, with data input quality (noise free) will be obtained an output that has the validity and high accuracy. The process that will be done is cleaning data replacing missing value to clean climate data from noise and distortion value. Starting with checking the input data whether it has an empty / tuple value, if there is an empty value for the attribute, the empty value can be filled with the new value using linear interpolation method [29]. Eliminating the missing value of the initial climate data done in linear regression and smoothing spline method. Interpolation using smoothing spline is given by can be seen in (1):

$$S(x) := \begin{cases} S_0(x) & x \in [x_0, x_1] \\ S_1(x) & x \in [x_1, x_2] \\ \quad \vdots & \quad \vdots \\ S_{n-1}(x) & x \in [x_{n-1}, x_n] \end{cases} \quad (1)$$

which each $S_i(x)$ is a polynomial of degree k.

The next process after the net climate data was correlation test to see the correlated variables using Keiser Mayer Olkin Factor Adequacy (KMO) method. Correlation test using KMO method is given by can be seen in (2).

$$KMO_j = \frac{\sum_{i \neq j} r^2_{ij}}{\sum_{i \neq j} r^2_{ij} + \sum_{i \neq j} u^2_{ij}} \quad (2)$$

Where,

$R = [r_{ij}]$ is the correlation matrix and

$U = [u_{ij}]$ is the partial covariance matrix.

B. Creating of Forecasting Model of DHF Multivariate Time Series using Vector Autoregressive.

The vector autoregression (VAR) model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series. The VAR model has proven to be especially useful for describing the dynamic behavior of economic and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model [22]. Vector Autoregressive in (3):

$$Y_t = c + \sum_{i=1}^{p} \Phi_i y_{t-i} + \varepsilon_t \qquad (3)$$

Where, y_t is the vector of *response* time series variables at time *t*. y_t has *n* elements. *c* is a constant vector of offsets, with *n* elements. Φ_i are *n*-by-*n* matrices for each *i*. The Φ_i are autoregressive matrices. There are *p* autoregressive matrices, and some can be entirely composed of zeros. ε_t is a vector of serially uncorrelated innovations, vectors of length *n*. The ε_t are multivariate normal random vectors with a covariance matrix Σ.

C. Evaluation Forecasting Model

The evaluation forecasting model, is done to find the best forecasting method to be used for forecasting the number of DHF cases. Evaluation model using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) to find the value of forecasting results, the smaller the error value than the more accurate forecasting results.

The Mean Absolute Error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. MAE is given by can be seen in (4). RMSE is a quadratic scoring rule that also measures the average magnitude of the error. It's the square root of the average of squared differences between prediction and actual observation. RMSE is given by can be seen in (5). Assuming that the difference or error is obtained using the Root Mean Square Error [28] :

$$MAE = \frac{1}{n} \sum_{j=1}^{n} \left| y_j - \hat{y}_j \right| \qquad (4)$$

Where, y_i is actual data, $\hat{y}_i$ is predicted data and n is the total number of data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n}} \qquad (5)$$

Where, P_i and O_i are the predicted and observed values, respectively; O is the average of O_i, and n is the total number of data.

The accuracy of the method is a measure to show how appropriate the method used to reproduce the data in the coming period, the smaller the error value, the better and the right method used.

Data forecasting results of the VAR method will be used to determine the pattern of DHF deployment. The model can be used for determination of dispersion pattern. Data forecasting results of the Vector autoregressive method will be used to determine the pattern of DHF deployment.

The model can be used for the determination of DHF dispersion pattern because it can perform Short and long-term Dengue forecasting (1-12 months) with a minimum error value.

D. Visualization of DHF Distribution Pattern using Spatial Autocorrelation

Spatial Autocorrelation (SA) is defined as the correlation between observation areas in the form of spatial patterns. The SA outline can be positive if the value of the observed region is similar to its neighbor's value and negative if the value of the observed region has a low degree of similarity with its neighbor.

The concept of spatial Autocorrelation indicator is divided into two methods of analysis namely the Global Indicator of Spatial Association (GISA) method and the Local Indicator of Spatial Association (LISA) method. GISA is used to determine the correlation within the entire set of observed data whereas LISA is used to determine the extent of spatially dependent variable data on the area around the observed location. In this research visualization pattern of dengue disease spread using LISA method.

The Local Indicator of Spatial Association (LISA) is a tool for the determination of spatial associations in each research area. The LISA method can indicate the region of centralization or the introduction of spatial phenomena in a region. LISA is as it is denoted in (6).

$$I = z_i x \sum_{j=1}^{n} w_{ij} z_j \qquad (6)$$

Which, z_i and z_j are deviations determined by (7).

$$Z_i = \frac{x_i - \bar{x}}{\delta} \qquad (7)$$

Value of W_{ij} is the weight matrix element determined by (8).

$$\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} = n \qquad (8)$$

The observed value of 'I' can be compared with the data distribution with the null hypothesis, it has no autocorrelation spatial, for example when the value of zi is independent of the value of z_J ($i \neq j$) in its neighbor region.

IV. EXPERIMENTAL RESULTS

Experimental results and model evaluation Forecast dengue hemorrhagic fever will be analyzed in Section IV.

A. Prapreprocessing Experiment Climate Data

In this experiment interpolation of climate data is performed with two methods of linear regression and smoothing spline. Graph of rainfall data and humidity of Sleman regency in 2010 until 2015 after preprocessing using linear regression method is presented in Fig. 3 and 4. In linear regression noise method is not seen on the chart with missing value but line connected does not follow curve pattern so it looks does not connect.
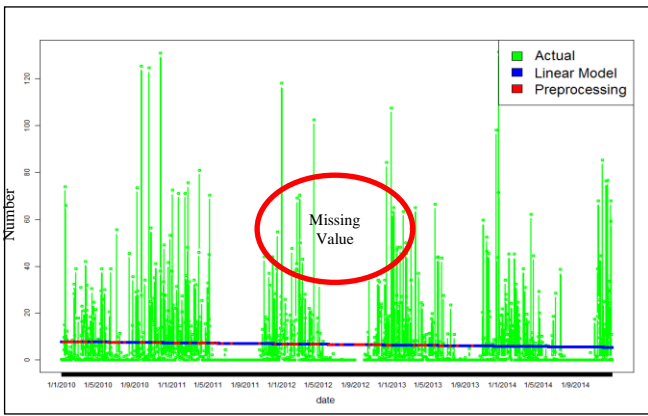
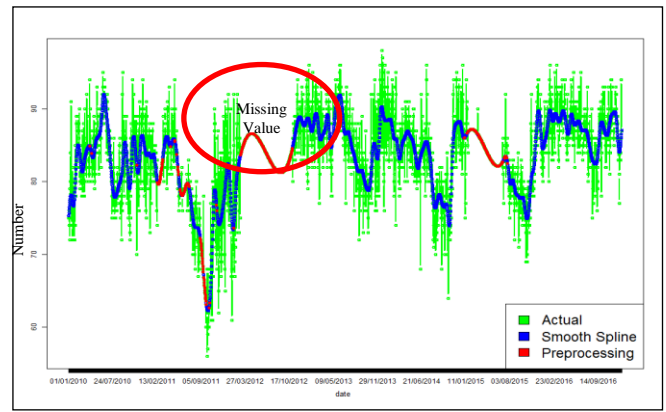Fig. 3.   Rainfall graph of 2010 until 2015 using linear regression after preprocessing.
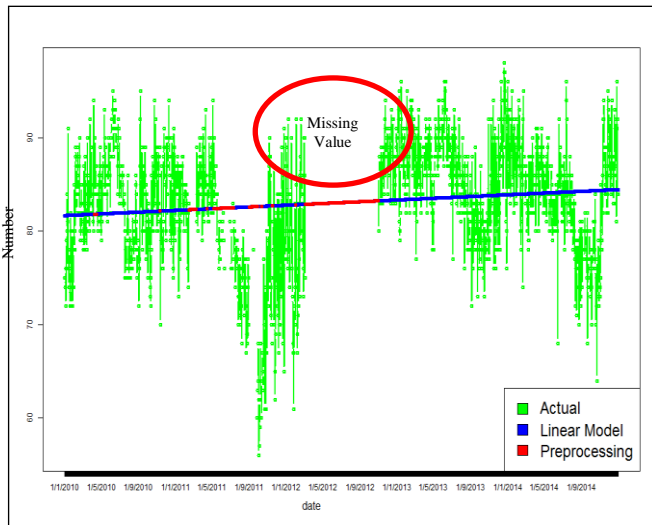

Fig. 4.   Graph of humidity in 2010 until 2015 using linear regression method after preprocessing.

The second method uses smoothing spline method. In this method noise is not visible on the graph with has missing value with the connecting line connecting rainfall data and average humidity while maintaining the smoothness of the curve as shown in Fig. 5 and 6.
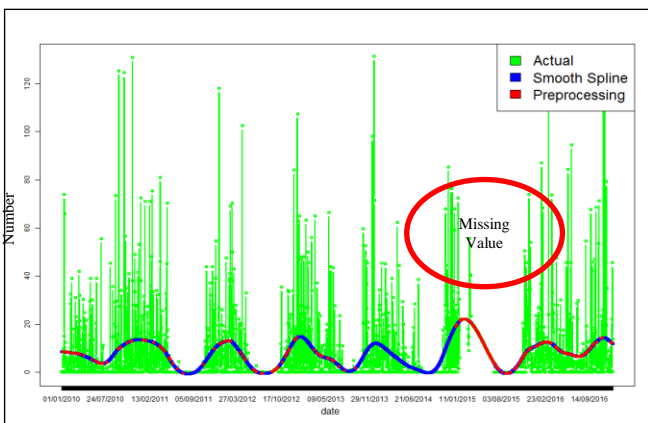

Fig. 5.   Rainfall graph of 2010 until 2015 using smoothing spline method after preprocessing.

Fig. 6.   The average humidity graph of 2010 to 2015 using smoothing spline method after preprocessing.
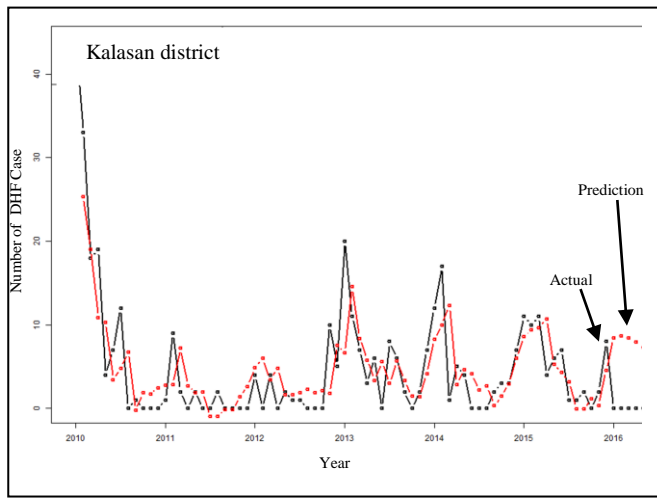
B. Experimental Forecasting Model the Number of DHF Cases using Vector Autoregressive Method

Fig. 7 shows predicted number of DHF cases in 2016 for Depok district (Fig. 7(a)), Kalasan District (Fig. 7(b)) and Mlati district (Fig. 7(c)). The Graph shows variation pattern of predicted data cycle of DHF cases from 2010-2015 following the previous period and according to the actual data cycle variation pattern. The graph shows a prediction point (red lines) almost equal to the actual point (black line) across almost the entire observation line. An important analysis of predictive experiments with the vector autoregressive method is that the 2016 prediction line shows a trend of seasonal patterns, the pattern indicates conformity with historical data patterns (actual data of previous years). Prediction pattern analysis shows that the corresponding vector autoregressive method is used to predict data on the number of DHF cases for the pattern of disease spread because the prediction pattern is seasonal. The analysis also shows the tendency of the vector autoregressive algorithm to be appropriate for short-term and long-term predictions.
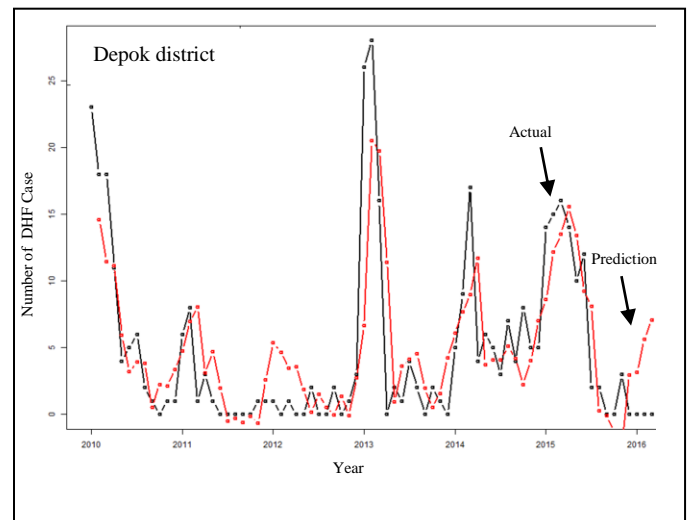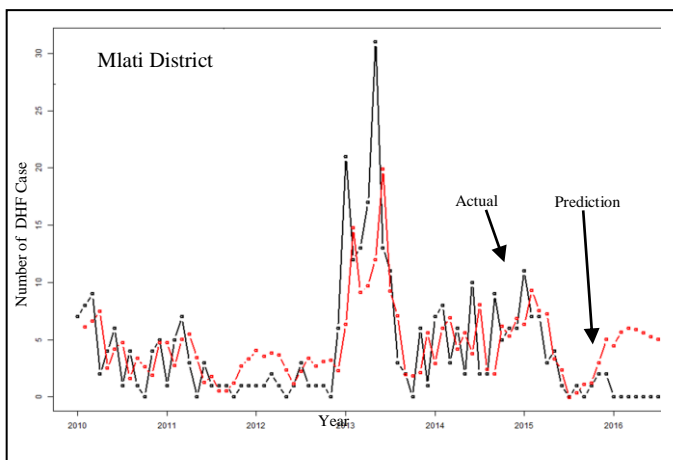

(a)

(b)

(c)

Fig. 7. Forecasting the Number of DHF case in 2016 at Depok (a), Kalasan (b), and Mlati District, (c) using Vector Autoregressive method.

The next step are to experiment the correlation test to determine the effect of each climate variable by using Keiser Mayer Olkin factor adequacy method as Table 1.

TABLE I.　RESULT CORRELATION TEST KMO METHOD

No.	Variable	KMO Value
1	Min Temperature	0.48
2	Max Temperature	0.62
3	Average Temperature	0.61
4	Average Humidity	0.55
5	Rainfall	0.69
6	Irradiation Time	0.70

From the results, KMO Test selected Variables that have a Value > 0.6 so that the variable of minimum temperature and average humidity removed [30].

C. Evaluation Forecasting Model

The results of experimental forecasting the number of DHF cases using VAR model for 17 Districts in Sleman, Yogyakarta-Indonesia showed a good suitability. This shows the conformity of the predicted data pattern of the number of dengue cases with the actual data. Table 2 shows the relatively small RMSE and MAE values.

TABLE II.　AVERAGE COMPARISON OF ERROR FITTING VALUES BETWEEN VECTOR AUTOREGRESSIVE ALGORITHMS WITH LINEAR REGRESSION, AND ARIMA ALGORITHMS FOR 17 DISTRICTS IN SLEMAN

No	Size test	Linear regression	ARIMA	VAR
1	RMSE	2.79	2.81	2.08
2	MAE	2.04	2.06	1.50

Evaluation fitting results the number of DHF cases in 17 districts using the VAR method has a low error, for the RMSE value of 2.08 and the MAE value of 1.50.

Comparison of error measurement between time series VAR algorithm with univariate time series algorithm using linear regression and ARIMA method for a number of DHF case in 2010 until 2015 for predicted data of 2016 could be seen in Table 3.

TABLE III.　AVERAGE COMPARISON OF ERROR FORECASTING VALUES BETWEEN VECTOR AUTOREGRESSIVE ALGORITHMS WITH LINEAR REGRESSION, AND ARIMA ALGORITHMS FOR 17 DISTRICTS IN SLEMAN

No	Test Size	Linier Regression	ARIMA	VAR
1	RMSE	3.49	3.55	3.16
2	MAE	2.71	2.68	2.47

The results of the predictive data test of the number of DHF cases using linear regression algorithm for the RMSE value of 3,49, and MAE value of 2,71. ARIMA method for the RMSE value of 3.55 and MAE value of 2.68.

The comparison of test size values between time series univariate algorithms with the proposed time series multivariate model for 17 districts in Sleman Regency in Table 3 shows that the proposed model has lower Error value compared with linear regression and ARIMA methods.

In this study forecasting the number of cases will be used to determine the distribution pattern of DHF so that required model that can make a prediction of the short-term or long-term a number of DHF case (minimum 1-12 months) with a minimum error value. Recommendation based on analysis of experimental results, the model for forecasting the number of DHF cases appropriate for the planning needs prevention of the spread of DHF disease is vector autoregressive spatial autocorrelation.

D. Visualization of Forecasting the Distribution Pattern of DHF Case using Spatial Autocorrelation

Fig. 8 shows districts that have hotspots are in Gamping, Depok and Mlati districts. The hotspots area indicates that the districts have the potential to spread DHF diseases so that the control of the spread needs to be done. While the districts that have coldspots are Moyudan, Minggir, Seyegan, Sleman, Ngemplak and Prambanan. Coldspots areas have the potential to become vulnerable to the spread of dengue fever which is transmitted by the surrounding area is high.
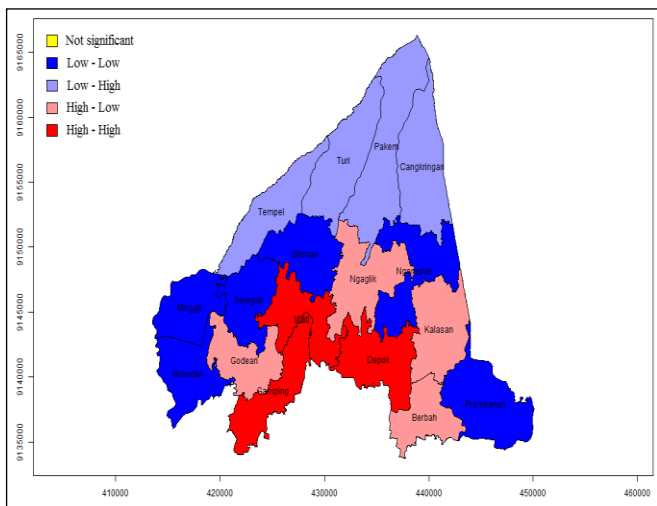
Fig. 8. LISA plot of DHF case forecasting.

Spatial autocorrelation analysis showed three districts in Sleman regency significant to High - High (HH) in 2010 until 2015; the area is Depok, Mlati, and Gamping District. This indicates that controlling the number of DHF case need to be done by the public health service in the three districts.

The hotspots (HH) area is a connectivity of areas with high dengue fever cases in the study sites.

V. CONCLUSION

Forecasting DHF using the vector autoregressive spatial autocorrelation (VARSA) model is proven to perform both short and long-term forecasting with relatively lower error values than linear regression and ARIMA method. This model can also see the pattern of the spread of dengue cases, based on the spatial connectivity of the number of dengue fever cases among the neighboring districts.

Spatial autocorrelation analysis showed three districts in Sleman regency significant to High - High (HH). the area is Depok, Mlati and Gamping district. This indicates that controlling the number of DHF patients needs to be done by the public health service in the three districts.

Future research is recommended and suggested to expand the scope of the research study area (scope of DI Yogyakarta or all provinces within Indonesia). Subsequent research includes the addition of variables forecasting the number of cases of dengue fever with socioeconomic data and population density as well as conducting a comparative study of forecasting methods in are needed this research using artificial neural network method.

REFERENCES

[1] Depkes, Regulation of the Ministry of Health *No 949/MENKES/SK/VIII/2004* About "the Guidelines for Implementing Early Warning System of Outbreaks". 2004.

[2] Kemenkes RI, Dengue Hemorrhagic Fever Bulletin, vol. 2. 2010

[3] Gibbons RV and Vaughn DW. "Dengue: an escalating problem". BMJ. 2002;324: 1563–1566. Pmid:12089096, 2002

[4] Raharja, A., Wiwik, A., and Retno, A.V., "Application of Exponential Smoothing Method For Forecasting Use of Telephone Time at PT. Telkom Divre 2 Surabaya", Information Systems, Faculty of Information Technology, Institut Teknologi Sepuluh November, SISFO-Information System Journal, 2009.

[5] MOH, "Indonesia health profile of 2015", Ministry of Health. 2016.

[6] Earnest, A., Tan, S. B., Wilder-Smith, A. and Machin, D. Comparing statistical models to predict dengue fever notifications. C*omputational and mathematical methods in medicine,* 2012.

[7] E. Z. Martinez, E. A. S. da Silva, and A. L. D. Fabbro, "A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of São Paulo, Brazil.," *Rev. Soc. Bras. Med. Trop.*, vol. 44, no. 4, pp. 436–40, 2011

[8] Ghil M., Allen M.R., Dettinger M.D., Ide K., Kondrashov D., Mann M.E., Robertson A.W., Saunders A., Tian Y., Varadi F., andYiou P., "Advanced Spectral Methods For Climatic Time Series", Reviews of Geophysics, 40, pp. 1-1 – 1-41, 2002.

[9] Ai, T.J., "Optimization of Exponential Smoothing Prediction One Variable Using Non-Linear Programming Algorithm", Journal of Industrial Technology), pp. 139 – 148, 1999.

[10] Unkel, S. and Farrington, C.P., "Statistical methods for the prospective detection of infectious disease outbreaks: A review", Department of Mathematics and Statistics, Faculty of Mathematics, Computing and Technology, The Open University, Walton Hall, Milton Keynes, UK, 2010.

[11] Lu, H.M., Zeng, D., and Chen, H., "Prospective Infectious Disease Outbreak Detection Using Markov Switching Models," Management Information Systems Department, University of Arizona, Tucson, 2008.

[12] Luz, P. M., Mendes, B. V., Codeco, C. T., Struchiner, C. J. and Galvani, A. P. Time series analysis of dengue incidence in Rio de Janeiro,Brazil. *Am. J. Trop. Med. Hyg.,* 933–939, 2008

[13] Rua-Uribe, G. L., Suarez-Acosta, C., Chauca, J., Ventosilla, P. and Almanza, R. Modelling the effect of local climatic variability on dengue transmission in Medellin (Colombia) by means of time series analysis. *Biomedica: revista del Instituto Nacional de Salud* Suppl 1, 142–152, 2013

[14] Eastin, M. D., Delmelle, E., Casas, I., Wexler, J. and Self, C. Intra-and interseasonal autoregressive prediction of dengue outbreaks using local climate and regional climate for a tropical environment in Colombia. *The American journal of tropical medicine and hygiene***,** 598–610. 2014.

[15] Wilder-Smith, A., Earnest, A., Tan, S., Ooi, E. and Gubler, D. Lack of association of dengue activity with haze. *Epidemiol. Infect.* 962–967, 2010.

[16] Anggraeni, W., Pramana, P. A. D., Samopa, F., Riksakomara, E., Wibowo, R., P., Condro, L., and Pujiadi, "Forecasting The Number Of Dengue Fever Cases In Malang Regency Indonesia," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 1, 2017

[17] A. C. Cruz, J. A. Lubrica, B. V. D. C. Punzalan, and M. C. Martin, "Forecasting dengue incidence in the National Capital Region, Philippines : using time series analysis with climate variables as predictors," vol. 60, pp. 19–26, 2012

[18] Hu, W., Clements, A., Williams, G. and Tong, S. Dengue fever and El Nino/Southern Oscillation in Queensland, Australia: a time series predictive model. *Occup. Environ. Med.***,** 307–311, doi: 10.1136/oem.2008.044966, 2010.

[19] Gharbi, M., Quenel, P., Gustave, J., Cassadou, S., La Ruche G., Girdary L., and Marrama L., Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. *BMC Infect Dis***,** 166, doi: 0.1186/1471- 334-11-166, 2011

[20] Bhatnagar, S., Lal, V., Gupta, S. D. and Gupta, O. P. Forecasting incidence of dengue in Rajasthan, using time series analyses. *Indian J. Public Health***,** 281–285, doi: 10.4103/0019-557X.106415, 2012.

[21] Halide, H. and Ridd, P. A. predictive model for Dengue Hemorrhagic Fever epidemics. *Int. J. Environ. Health Res.* 253–265, doi: 10.1080/09603120801966043, 2008.

[22] Goto, K., Kumarendran, B., Mettananda, S., Gunasekara, D., Fujii, Y., and Kaneko, S., "Analysis of effects of meteorological factors on dengue incidence in Sri Lanka using time series data". *PLoS One*, e63717, doi: 10.1371/ journal.pone.0063717, 2013.

[23] Kumar S, Managi S, Matsuda A, "Stock prices of clean energy firms, oil and carbon markets: A vector autoregressive analysis". Energy Economics 34: 215–226. 2012

[24] Opgen-Rhein R, Strimmer K Learning causal networks from systems biology time course data: an effective model selection procedure for the vector autoregressive process. BMC Bioinformatics 8 Suppl 2: S3. 2007

[25] Akinboade OA, Braimoh LA, "International tourism and economic development in South Africa: a Granger causality test". International Journal of Tourism Research 12: 149–163. *in medicine.* 2012.

[26] Unkel, S. and Farrington, C.P., "Statistical methods for the prospective detection of infectious disease outbreaks: A review", Department of Mathematics and Statistics, Faculty of Mathematics, Computing and Technology, The Open University, Walton Hall, Milton Keynes, UK, 2010.

[27] Sofla, M.J., Silahi, B., and Masomi, M. T., "Germi County Seasonal Precipitation Routing and Analysis, Using Holt-Winter Method for Times with Non-seasonal changes", Technical Journal of Engineering and Applied Sciences, Vol 3(11): 950-953, 2013.

[28] Mara, M.N., Satyahadewi, N., and Yundari, "Theoretical Review Hybridizing Exponential Smoothing and Neural Network For Time-series Forecasting", Scientific Bimaster Mat. Stat. And Applied (Bimaster), Vol. 2 No. 3, pp. 205-210, Pontianak, 2013.

[29] Hartomo, K., D., Winarko, E., and Subanar, "ESSPI : Exponential smoothing seasonal Platting Index, A New Algorithm for prediction Rainfall in Indonesia" International journal of computer science and information security, vol 14 no. 6, June edition, 2016

[30] Cerny, C.A., and Kaiser, H.F.. A study of a measure of sampling adequacy for factor-analytic correlation matrices. Multivariate Behavioral Research, 12(1), 43-47, 1977