Penerbit Akademia Baru

Journal of Advanced Research Design

Journal homepage: www.akademiabaru.com/ard.html
ISSN: 2289-7984

Mathematical modelling of dengue pattern in Penang, Malaysia

Open Access

Norzaida Abas[1,*], Rafidah Mohd Shamsuddin[1], Syafrina Abd. Halim[2], Noor Shazwani Osman[1]

[1] UTM Razak School of Engineering and Advanced Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia
[2] Department of Science in Engineering, Kuliyyah of Engineering, International Islamic University of Malaysia,53100 Gombak, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 5 June 2017
Received in revised form 14 August 2017
Accepted 16 August 2017
Available online 22 August 2017

Keywords:
Dengue, urban area, forecasting, double exponential smoothing method

ABSTRACT

Dengue fever is an endemic disease in many tropical and subtropical regions. In Malaysia, it is the leading public health challenge despite the extensive intervention programs by the related authorities. Distribution of dengue cases in Malaysia varies according to states and districts where cases are more distinct in urban and suburban areas. Preparedness strategies of dengue cases could be more successful with some comprehensive and technical analysis on disease incidences. Hence, the present study analyses dengue cases using mathematical modelling in the state of Penang, one of the more urbanised state. In particular, two time series models are fitted to the dengue data from the region in order to identify the mathematical model that best describe the data. Results show that both proposed models are able to represent the cases rather well; however numerical inspection revealed that Double Exponential Smoothing method is the better choice. Subsequently, the identified model is used to make forecasting on the number of expected cases. Results show that dengue cases in Penang are expected to increase gradually.

1. Introduction

Dengue is an endemic disease not only in Malaysia but in many tropical regions. Dengue viruses are spread to humans by means of the bites of infected female Aedes mosquitoes. According to WHO [1], worldwide, the occurrence of dengue cases is escalating, from 2.2 million in 2010 to 3.2 million cases in 2015. Similarly, Malaysia experiences an alarming rise of dengue cases, despite numerous efforts of prevention and control measures. Data from Malaysia i-Dengue Data Portal (see Figure1) indicated that the number of reported dengue cases for the period 1995-2015 [2] showed an 18-fold increase. Distributions of dengue cases in Malaysia vary considerably between states and districts where dengue cases are more pronounced in urban and suburban areas. This disparity is not a surprise since the distribution of dengue cases is shaped by factors such as urbanization, weather

[*] *Corresponding author.*
E-mail address: Norzaida Abas (zaida.kl@utm.my)

condition and poor environmental hygiene. Furthermore, Malaysia's tropical weather makes it conducive to dengue breeding.

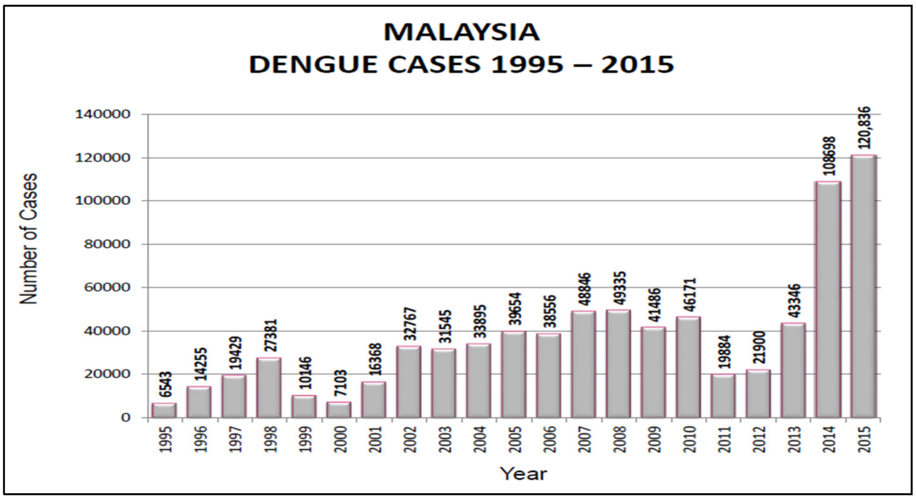

Fig. 1. Dengue Cases in Malaysia for 1995-2015

Monitoring, curbing and prevention strategies of dengue cases could be more effective with the inclusion of some comprehensive analysis on the disease incidences. In general, analysis using mathematical modelling of infectious diseases such as the dengue fever helps in monitoring the dynamics of the disease [3-4]. Modelling not only examines the present scenario of the infectious diseases but anticipates the trend of future spread, which in turn, could be used to evaluate strategies for combating the diseases. Forecasting information may aid in managing the distribution of resources and forestalls outbreaks. Literature shows there are various approaches of mathematical modelling being carried out in Malaysia such as time series analysis [5], neural networks and nonlinear regression models [6].

The present study aims to use time series approach to analyse dengue cases in Penang, one of the most urbanised states. Although Penang is the second smallest state in Malaysia, the rate of dengue cases is quite disturbing (see Figure 2). Coincidently, the first dengue being reported in 1902 was in Penang [7]. Time series models are fitted to the state's dengue data to determine the mathematical model that best represents the dengue incidences. Subsequently, the identified model is used to make forecasting on the number of expected cases.

2. Model Development

2.1 Data

The region under study is the state of Penang, located on the northwest coast of Peninsular Malaysia. The state which comprises of the Penang Island and Seberang Perai on the mainland, is densely populated, sharing 2nd place with Putrajaya regarding population. Economically, Penang relies on manufacturing, with numerous multinational companies engaging in electronics, engineering and other technology-related activities are situated. Penang is also a popular tourist destination for both local and foreign travellers. Rapid urbanization in Penang seems to contribute to the rising rate of dengue cases (see Figure 3), with the cases for period 2010-2015 showed a 2-fold increase.

This study examined dengue cases using time series approach. Time series is a set of historical data arranged sequentially according to time. Past studies showed that time series forecasting has been used to analyse dengue cases in many affected regions [8-10]. In this study, two time series model namely Double Exponential Smoothing and Holt Winters' method are fitted to the historical data. The performance of these two models are evaluated and compared, and subsequently, the best model is used to make forecasting. Seven years (2010 – 2016) of dengue data in the form of weekly data are obtained from the official Malaysia i-Dengue Data Portal, Ministry of Health Malaysia. Specifically, weekly data from week 8, 2010 to week 46, 2016 (351 weeks) are analysed.

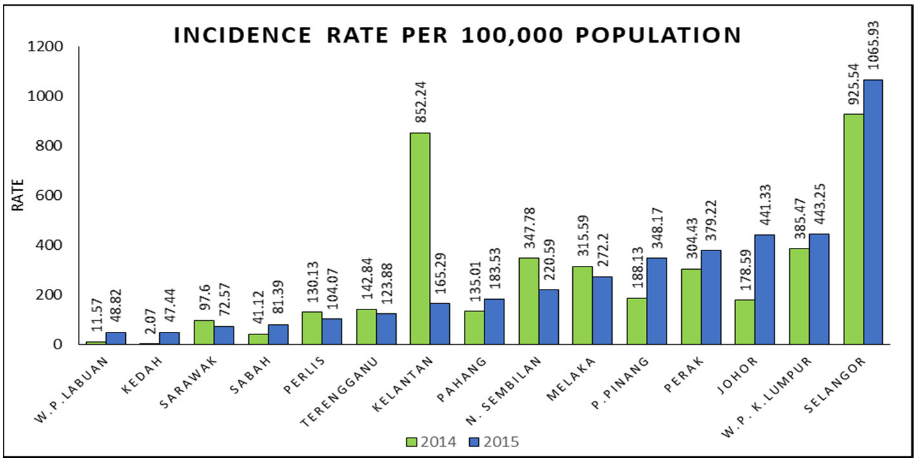

Fig. 2. Breakdown of dengue cases according to state for 2014 and 2015

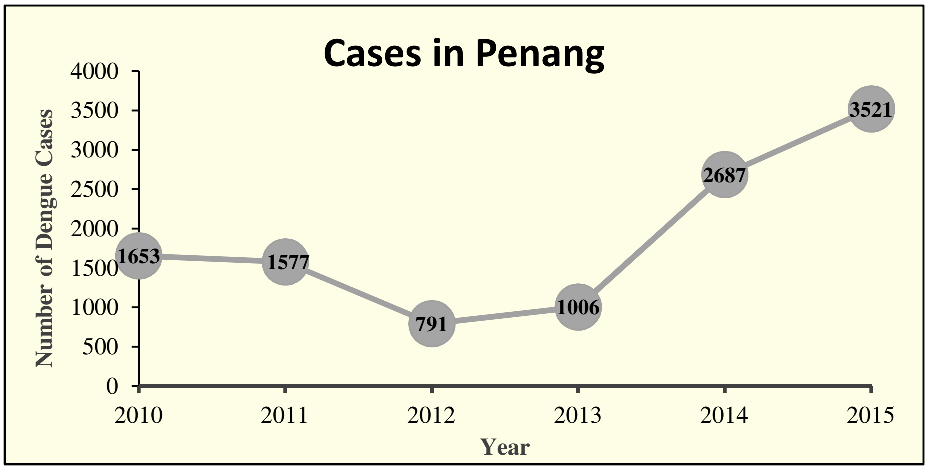

Fig. 3. Dengue Cases in Penang for 2010-2015

2.2 Exponential Smoothing

In Exponential smoothing [9] technique, an exponentially weighted average of past observations is used as a basis to make forecasting. The size of the weight assigned to each observation depends on the arrangement of data with regards to time. The biggest weight is allocated to the present observation while less weight is assigned to older data.

Penerbit Akademia Baru

2.2.1 Double exponential smoothing

Double exponential smoothing technique smooths out the data when a trend is present. Exponential smoothing with a pattern works much like simple smoothing apart from that two components must be updated every period - level and trend. The level is a smoothed estimate of the value of the data at the end of every period. The trend is a smoothed gauge of the estimation of average growth at the end of every period. The formula for simple exponential smoothing S_t is:

$$S_t = \alpha * y_t + (1 - \alpha) * (S_{t-1} + b_{t-1}),\ 0 < \alpha < 1 \quad (1)$$

where $b_t = \gamma * (S_t - S_{t-1}) + (1 - \gamma) * b_{t-1}\ ,\ 0 < \gamma < 1$ (2)

Note that the present value of the series is used to calculate its smoothed value replacement in double exponential smoothing. There are numerous techniques to choose the initial values for S_t and b_t. S_1 is in general set to y_1.
Three recommendations for b_1

$b_1 = y_2 - y_1$
$b_1 = [(y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3)]/3$
$b_1 = (y_n - y_1)/(n - 1)$

2.2.2 Holt-Winters Forecasting

Holt's method is used to manage time series when there are trend and seasonal variations. There are two forms in Holt-Winters method, multiplicative and additive, the utilization of which relies on the characteristics of the specific time series [8].

The general forecast function for the multiplicative Holt-Winters method is:

$$\hat{y}_{n+l|n} = (m_n + l b_n) c_{n-s+l} \qquad l = 1, 2, \ldots \quad (3)$$

where m_n is the component of level, b_n is the component of the slope, and c_{n-s+l} is the relevant seasonal component, with s signifying the seasonal period (e.g. 12 for monthly data and 4 for quarterly data)

Therefore if a monthly time series is considered, the one step ahead forecast is given by:

$$\hat{y}_{n+l|n} = (m_n + b_n) c_{n-11} \quad (4)$$

The updating formulae for the three components will each require a smoothing constant. If once again α_0 is used as the parameter for the level and α_1 for the slope, and a third constant α_2, is added as the smoothing constant for the seasonal factor, the updating equations will be:

$$m_t = \alpha_0 \frac{y_t}{c_{t-s}} + (1 - \alpha_0)(m_{t-1} + b_{t-1}) \quad (5)$$

$$b_t = \alpha_1 (m_t - m_{t-1}) + (1 - \alpha_1) b_{t-1} \quad (6)$$

$$c_t = \alpha_2 \frac{y_t}{m_t} + (1 - \alpha_2)c_{t-s} \tag{7}$$

Once again, α_0, α_1, and α_2 all lie between zero and one. If the aforementioned additive version of Holt-Winters was used, the seasonal factor is simply added as opposed to multiplied into the one step ahead forecast function, thus:

$$\hat{y}_{n+l|n} = m_n + b_n + c_{n-11} \tag{8}$$

and the level and seasonal updating equations involve differences as opposed to ratios:

$$m_t = \alpha_0(y_t - c_{t-s}) + (1 - \alpha_0)(m_{t-1} + b_{t-1}) \tag{9}$$

$$c_t = \alpha_2(y_t - m_t) + (1 - \alpha_2)c_{t-s} \tag{10}$$

The slope component, b_t, remains unchanged. Normally, the smoothing parameters are chosen between the range of 0.02 and 0.2.

3. Results and Discussion

The comparison between observed data of dengue cases and fitted data for both Double Exponential Smoothing and Winters' Method are individually constructed in Figure 4 and 5 respectively. The performance of both methods is very similar where both techniques are suitable for representing the dengue cases. Further comparison is conducted to determine the better model by using mean absolute percentage error (MAPE), mean absolute deviation (MAD) and mean squared deviation (MSD) (see Table 1). The lowest value (in bold) indicates the better model. The comparison results are consistent, suggesting that Double Exponential Smoothing is more suitable in representing dengue cases in Penang. Subsequently, Double Exponential Smoothing model is used to forecast for 13 weeks of cases with the results shown in Table 2. The forecast results anticipate that dengue cases for Penang would slowly escalate for the next 13 weeks.

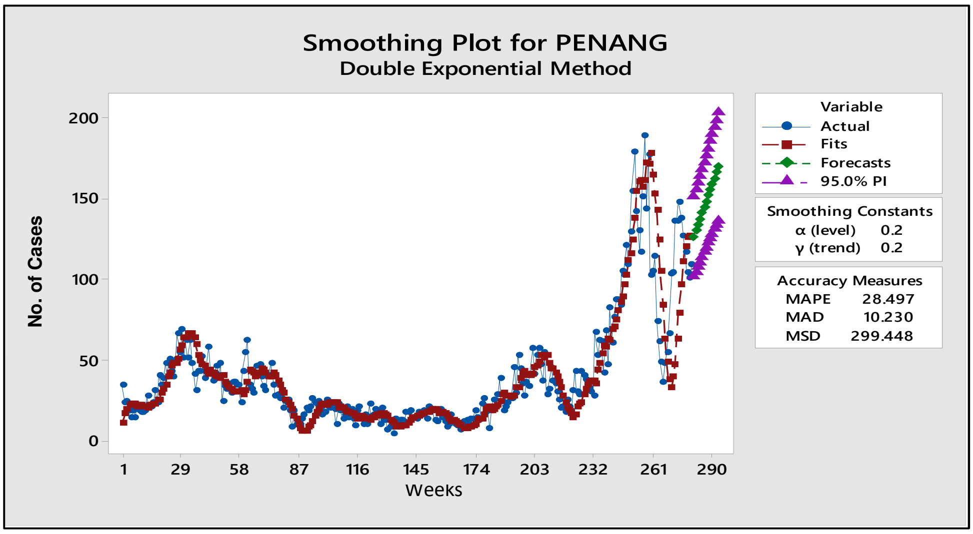

Fig. 4. Double Exponential Smoothing Plot for Penang

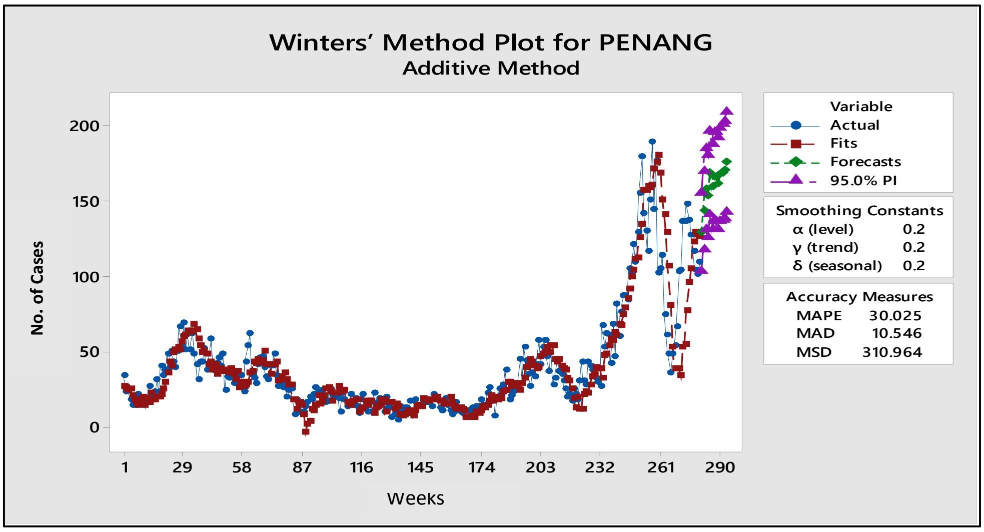

Fig. 5. Winters' Method Plot for Penang

Table 1
Numerical comparisons of both models

Accuracy Measures	Double Exponential Smoothing Method α (level) 0.2, γ (trend) 0.2	Winters' Method (Additive) α (level) 0.2, γ (trend) 0.2 δ (seasonal) 0.2
MAPE	28.50	30.03
MAD	10.23	10.55
MSD	299.45	310.96

Table 2
Forecasting result for week 47, 2016 to week 7, 2017

Week	**Expected Cases**
1	126
2	130
3	134
4	137
5	141
6	144
7	148
8	152
9	155
10	159
11	163
12	166
13	170

4. Conclusion

This study focuses on modelling dengue cases in Penang and make predictions on future cases base on the time series approach. Identifying the correct mathematical model is crucial and useful in understanding, predicting and guiding dengue control. Results show that both proposed model is able to capture the pattern of the dengue cases. Numerical comparison results also show that the values of all three criteria for both models are almost similar, nevertheless the better model is Double Exponential Smoothing. Forecasting of future dengue cases using the identified model reveal that for the next 13 weeks, dengue cases are expected to slowly increase (see Table 2).

It should be noted that the work in this study is limited to the available data posted at the Ministry of Health website. Normally, data or information on dengue is limited to press statements and press releases to the general public issued by the Ministry of Health and also from the Minister of Health. Nevertheless, this study's methodology and results could be used as a basis for providing public and clinical health authorities with the information needed to implement appropriate actions and efficient use of limited resources.

Acknowledgement
We like to thank the Ministry of Health for the dengue data used in this study. We also thank University Teknologi Malaysia for the support and funding of the study with UTM GUP Vote: Q.K130000.2540.11H63

References

[1] Dengue and severe dengue: Fact Sheet No 117, Geneva: World Health Organization (WHO); 2017. [extracted May 3, 2017]. www.who.int/mediacentre/factsheets/fs117

[2] Department of Statistics Malaysia, Official Portal [webpage on the Internet]. [extracted May 3, 2017]. Available from: https://www.dosm.gov.my/v1/index.php

[3] Anderson, Roy M., Robert M. May, and B. Anderson. *Infectious diseases of humans: dynamics and control*. Vol. 28. Oxford: Oxford university press, 1992.

[4] de Castro Medeiros, Líliam César, Cesar Augusto Rodrigues Castilho, Cynthia Braga, Wayner Vieira de Souza, Leda Regis, and Antonio Miguel Vieira Monteiro. "Modeling the dynamic transmission of dengue fever: investigating disease persistence." *PLOS neglected tropical diseases* 5, no. 1 (2011): e942.

[5] Cheong, Yoon Ling, Katrin Burkart, Pedro J. Leitão, and Tobia Lakes. "Assessing weather effects on dengue disease in Malaysia." *International journal of environmental research and public health* 10, no. 12 (2013): 6319-6334.

[6] Husin, Nor Azura, and Naomie Salim. "A comparative study for back propagation neural network and nonlinear regression models for predicting dengue outbreak." *Jurnal Teknologi Maklumat Bil* 4 (2008): 97-112.

[7] Skae, F. M. T. "Dengue fever in Penang." *British Medical Journal* 2, no. 2185 (1902): 1581.

[8] Promprou, S., M. Jaroensutasinee, and K. Jaroensutasinee. "Forecasting Dengue Haemorrhagic Fever Cases in Southern Thailand using ARIMA Models." (2006).

[9] Choudhury, Zamil MAH, Shahera Banu, and Amirul M. Islam. "Forecasting dengue incidence in Dhaka, Bangladesh: A time series analysis." (2008).

[10] Ho, Chiung Ching, and Choo-Yee Ting. "Time Series Analysis and Forecasting of Dengue Using Open Data." In *International Visual Informatics Conference*, pp. 51-63. Springer, Cham, 2015.