



Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Disease

journal homepage: www.elsevier.com/locate/apjtd

Document heading

doi:10.1016/S2222-1808(14)60608-3

© 2014 by the Asian Pacific Journal of Tropical Disease. All rights reserved.

Time series model on hand, foot and mouth disease in Sarawak, Malaysia

Noraishah Mohammad Sham^{1*}, Isthinayagy Krishnarajah^{1,2}, Mahendran Shitan^{1,2}, Munn–Sann Lye³¹Laboratory of Computational Statistics and Operations Research, Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia²Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia³Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia

PEER REVIEW

ABSTRACT

Peer reviewer

Shamarina Shohaimi, Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

Tel: 03–89466637

E-mail: shamarina@upm.edu.my

Comments

The present study is a good study because of the lack of time series models in predicting the outbreak of HFMD in Sarawak. This is important because poor control of this disease could lead to fatalities. The ARMA model for HFMD has a forecast ability of approximately 94%, thus is able to make good predictions of future HFMD outbreaks in Sarawak.

Details on Page 472

Objective: To model and forecast hand, foot and mouth disease (HFMD) cases since there has not been any studies on time series modelling of HFMD in Sarawak, Malaysia.

Methods: The longitudinal data of HFMD collected over the past 7 years (2006–2012) showed a cyclical pattern of outbreak every 2 to 3 years. The data consisted of weekly number of HFMD reported cases from 2006 to 2012. An auto regressive moving average (ARMA) model was fitted to the data. Batting average was used to measure the performance of our forecast for 2013.

Results: The results indicate that ARMA model fit the trends of HFMD in Sarawak very close to the actual data during our study period. Our prediction for the number of HFMD reported cases lies within the 90% interval range for the first 16 weeks in 2013.

Conclusions: An ARMA(1,4) model fits the data well and has good prediction ability of over 90% accuracy.

KEYWORDS

HFMD, Time series, ARMA model, Sarawak

1. Introduction

Sarawak, the largest state in Malaysia located on the island of Borneo (Figure 1), comprises of 11 districts. Hand, foot and mouth disease (HFMD) is endemic in this state. HFMD is a tropical disease and generally manifests itself in the form of blister-like rash on hands, feet and mouth, usually in young children^[1], with previous outbreaks of HFMD having been reported in Sarawak^[2]. Although this disease mainly affects children, non-immune adults may also be affected. Adults are less susceptible to the

disease due to immunity from previous exposures to this disease. HFMD is usually diagnosed by a complete physical examination of the patient. A throat and/or blister swab is used to determine which enterovirus caused the illness^[2]. So far, there is no specific effective antiviral drugs for the treatment of, or vaccine available for prevention from HFMD^[1]. Symptomatic treatment is usually given to provide relief from fever, aches, or pain from the mouth ulcers. Children with HFMD are advised to take enough liquids to maintain hydration^[1]. The disease can be fatal in some instances.

*Corresponding author: Noraishah Mohammad Sham, Laboratory of Computational Statistics and Operations Research, Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.

Tel: +60122870461

E-mail: aishahsham@yahoo.com

Foundation Project: Supported by the Ministry of Higher Education under the Fundamental Research Grant Scheme.

Article history:

Received 5 May 2014

Received in revised form 16 May, 2nd revised form 25 May, 3rd revised form 30 May 2014

Accepted 17 Jun 2014

Available online 10 Jul 2014

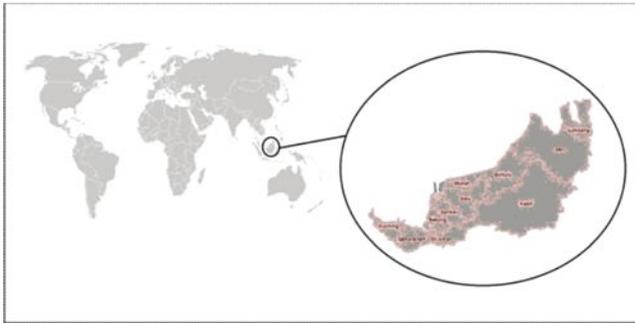


Figure 1. Location of Sarawak, Malaysia.

Numerous studies have been conducted using time series analysis for various infectious diseases such as dengue fever, hemorrhagic fever with renal syndrome, viral respiratory illness, malaria, cholera, etc. For example, a study done in Brazil found that autoregressive integrated moving average (ARIMA) models are useful in monitoring dengue incidence[3]. For modelling hemorrhagic fever with renal syndrome in China, ARIMA models have also been used[4]. Another study done in United States used time series modelling techniques in forecasting the burden of severe viral respiratory illness in paediatric intensive care unit[5]. In Malaysia, ARIMA model with climate variables (rainfall, relative humidity and temperature) was fitted to describe the trends of dengue cases[6].

However, time series modelling of HFMD in Sarawak has not been undertaken as yet. As such, the objective of this study is to fit a time series model for HFMD data in Sarawak and use it for forecasting purposes.

2. Materials and methods

The weekly HFMD data from 2006 to 2012 was obtained from Sarawak State Health Department. Figure 2 shows the time series plot of HFMD dataset for 7 years period that is used in this study. In our study, we fit an auto regressive moving average (ARMA) model to the data. The following equation depicts the ARMA time series sequence:

$$(1 - \sum_{i=1}^p \phi_i L^i) X_t = (1 + \sum_{i=1}^q \theta_i L^i) Z_t$$

Where X_t is the number of HFMD cases, ϕ_i is the autoregressive parameter to be estimated at order i , L^i is the lag operator, θ_i is the moving average parameter to be estimated at order i , Z_t is white noise error term at time t .

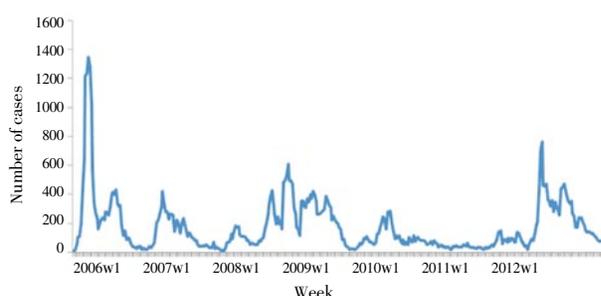


Figure 2. Time series of weekly HFMD cases in 2006–2012.

The fitting of ARMA model is based on Box–Jenkins methodology that is based on the assumption that concerns the characteristics of the initial data series[7]. In order to test the stationarity of the data, we performed augmented Dickey–Fuller procedure using an appropriate command in R version 3.0.2 where the hypothesis test is as follows[8]:

H_0 : Unit root exists

H_a : Unit root does not exist

We will reject the null hypothesis if P -value is less than $\alpha=0.05$.

In order to select the best model for this study, we look at the minimum value of Akaike's information criterion (AICc) which assesses the goodness-of-fit of the model[9]. AICc is AIC with a correction for finite sample sizes that measures the fitness of ARIMA model. It measures the penalty on the likelihood for each additional term included in the model. All values are produced by R version 3.0.2.

Basically, when the best model has been selected we are capable of generating forecast values and intervals. In this study we use batting average to measure the performance of our forecast. This method is chosen because it is straightforward and easy to understand. We present the results as the percentage of number of HFMD reported cases which lies within the 90% interval range for the first 16 weeks in 2013. The batting average count is as given below:

$$\text{Batting average} = \frac{\text{Number of HFMD reported cases}}{\text{Number of week}} \times 100$$

3. Results

The augmented Dickey–Fuller test was conducted and the results showed that there was evidence to reject H_0 at 5% significance level with $P=0.01$. Therefore, we can conclude that the time series was stationary and it would be appropriate to fit an ARMA model.

The autocorrelation (ACF) and partial autocorrelation (PACF) functions were produced to indicate the temporal dependence structure in the stationary time series. It was important in the identification of possible model order. A few models were constructed based on ACF and PACF correlograms (Table 1). The ACF and PACF are graphically presented in Figure 3. All the ACF values exceed the significance level while for the PACF plots partial autocorrelations at lag 1, 2, 3 and 6 exceed the significance level. Based on both plots, we found that the AR(4), MA(6), ARMA(1,4), ARMA(1,5) and ARMA(2,2) models could be the possible models that suited the HFMD data.

ARMA(1,4) gave the minimum value of AICc among all the possible models (Table 1). Therefore, ARMA(1,4) appeared to be the best fit for the data. The fitted ARMA(1,4) model for the mean-corrected data is given as:

$$(1 - 0.773L)X_t = Z_t + 0.3905Z_{t-1} + 0.3919Z_{t-2} + 0.2614Z_{t-3} + 0.2594Z_{t-4}$$

where $Z_t \sim WN(0, 3686.7)$.

Table 1

Models based on ACF and PACF correlograms.

| Model form | AICc |
|------------|---------|
| AR(4) | 4046.12 |
| MA(6) | 4062.35 |
| ARMA(1,4) | 4038.25 |
| ARMA(1,5) | 4039.66 |
| ARMA(2,2) | 4045.21 |

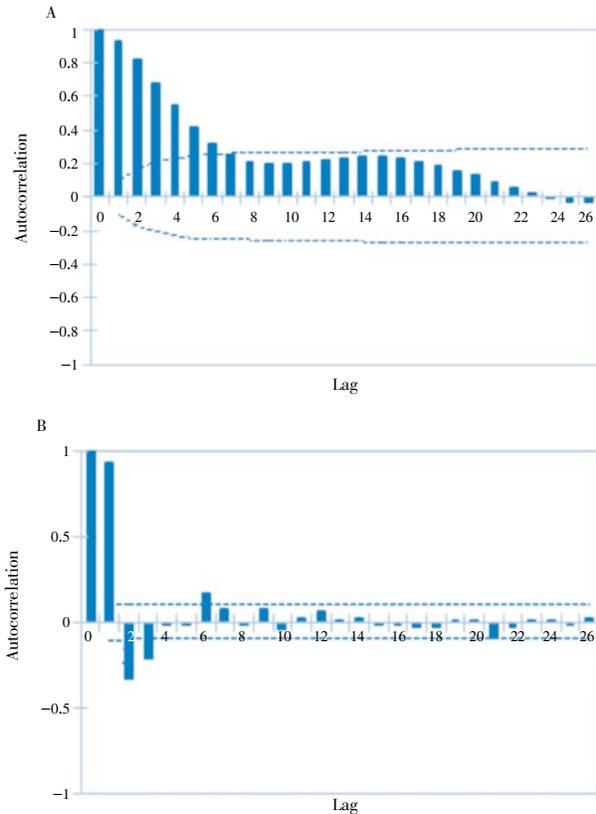


Figure 3. Autocorrelation (a) and partial autocorrelation (b) functions of HFMD incidence for the period 2006–2010.

We plotted the fitted and reported HFMD cases as shown in Figure 4. The plot showed that the predicted cases using ARMA(1,4) match the actual cases very closely. However, the diagnostic for residuals is still needed for assessing the model adequacy. Based on the randomness tests (Table 2), none of the tests rejects H_0 at 5% significance level. Hence, the fitted model appears to be good for this data.

We use the ARMA(1,4) model to obtain the 90% forecast interval for the number of HFMD cases. We then compared it with the observed HFMD cases. According to Table 3, 15 out of 16 (94%) HFMD reported cases in 2013 fall within the 90% forecast interval.

Table 2

Randomness tests for residuals.

| Statistic | DF | Value | P-value |
|------------|----|--------|---------|
| Box-Pierce | 25 | 33.793 | 0.112 |
| Ljung-Box | 25 | 34.707 | 0.094 |
| McLeod-Li | 25 | 31.891 | 0.161 |

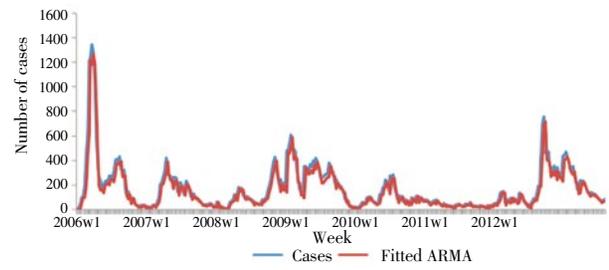


Figure 4. Time series of actual number of HFMD cases and fitted cases from ARMA(1,4).

Table 3

Forecast of HFMD reported cases.

| Week | 90% forecast interval | HFMD reported cases | Correctly forecasted |
|------|-----------------------|---------------------|----------------------|
| 1 | (0, 195) | 89 | √ |
| 2 | (0, 262) | 173 | √ |
| 3 | (0, 321) | 196 | √ |
| 4 | (0, 367) | 195 | √ |
| 5 | (0, 405) | 296 | √ |
| 6 | (0, 429) | 295 | √ |
| 7 | (0, 443) | 232 | √ |
| 8 | (0, 453) | 426 | √ |
| 9 | (0, 459) | 512 | X |
| 10 | (0, 463) | 422 | √ |
| 11 | (0, 466) | 450 | √ |
| 12 | (0, 469) | 402 | √ |
| 13 | (0, 470) | 231 | √ |
| 14 | (0, 471) | 367 | √ |
| 15 | (0, 472) | 278 | √ |
| 16 | (0, 473) | 449 | √ |

4. Discussion

The objective of this study was to fit a simple time series model that could be used for forecasting HFMD cases. This objective is met where we were able to fit an ARMA(1,4) model to the HFMD data in Sarawak, Malaysia and it is seen from the results obtained the model suits the data well. The best model selected is purely based on historical data from 2006 to 2012. Results obtained show that the ARMA(1,4) model has a forecast ability of approximately 94%. Hence, we can say that the model is capable of giving good prediction. This model could also be used to predict future outbreaks. An advantage of this model is that it is purely based on historical data and as such would not require other variables such as rainfall and temperature. Furthermore, since the model is simple, it is easy to incorporate bigger datasets and there is also an added advantage whereby it is easy to update the model in future as more data become available. The results of this study would be useful to health policy makers by predicting the size of the outbreak based on this model in order to help them to prepare adequate facilities to control the disease. However, this model is not able to forecast too far into the future as it is a short memory time series model.

In this study, factors such as rainfall and temperature are not considered because it is not realistic to forecast these

variables. Nevertheless, the simple ARMA(1,4) model gives a good fit. Thus, we can say that a time series model based purely on historical data is sufficient for this HFMD dataset.

In future study, we plan to develop models to predict impending outbreaks and investigate factors that may contribute to outbreaks of HFMD in Sarawak.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgements

We thank Crisis Preparedness and Response Centre, Communicable Disease Control Section, Sarawak Health Department for the invaluable assistance on compiling the seven years dataset; many thanks to Institute for Medical Research and Ministry of Science, Technology and Innovation for the financial support. This project was funded by the Ministry of Higher Education under the Fundamental Research Grant Scheme.

Comments

Background

HFMD which usually affects young children is endemic in the state of Sarawak, Malaysia. Persons affected with the disease develop blister-like rash on hands, feet and mouth and if untreated can lead to death. However, there is no specific effective antiviral drug for the treatment of HFMD and those affected are usually treated by providing relief from fever, aches, or pain from the mouth ulcers.

Research frontiers

Currently there is no reported research on time series modelling of HFMD in Sarawak. The present study addresses this by fitting a time series model which is the ARMA model for HFMD data in Sarawak for forecasting purposes.

Related reports

There have been studies on time series analysis for various infectious diseases using ARMA models such as dengue fever in Brazil, hemorrhagic fever with renal syndrome in China, viral respiratory illness in United States and in Malaysia. The ARMA model with climate variables (rainfall, relative humidity and temperature) was fitted to describe the trends of dengue cases.

Innovations & breakthroughs

With a forecast ability of approximately 94%, the model

is able to accurately predict future outbreaks of HFMD and because it is based on historical data the model does not require the use of other variables. It is also easy to use bigger datasets and easy to update the model in future because the model is simple.

Applications

The model would be useful to health policy makers in terms of making prediction of future outbreaks and providing adequate resources such as facilities and personnel in order to control this disease.

Peer review

The present study is a good study because of the lack of time series models in predicting the outbreak of HFMD in Sarawak. This is important because poor control of this disease could lead to fatalities. The ARMA model for HFMD has a forecast ability of approximately 94%, thus is able to make good predictions of future HFMD outbreaks in Sarawak.

References

- [1] Centers for Disease Control and Prevention. Hand, foot, and mouth disease. Atlanta: Centers for Disease Control and Prevention; 2013. [Online] Available from: <http://www.cdc.gov/features/handfootmouthdisease/> [Accessed on 20th May, 2014]
- [2] Sarawak State Health Department. What are the clinical signs and symptoms. Sarawak: Sarawak State Health Department; 2013. [Online] Available from: http://jksarawak.moh.gov.my/en/modules/mastop_publish/?tac=91 [Accessed on 20th May, 2014]
- [3] Luz PM, Mendes BV, Codeço CT, Struchiner CJ, Galvani AP. Time series analysis of dengue incidence in Rio de Janeiro, Brazil. *Am J Trop Med Hyg* 2008; **79**(6): 933–939.
- [4] Li Q, Guo NN, Han ZY, Zhang YB, Qi SX, Xu YG, et al. Application of an autoregressive integrated moving average model for predicting the incidence of hemorrhagic fever with renal syndrome. *Am J Trop Med Hyg* 2012; **87**(2): 364–370.
- [5] Spaeder MC, Fackler JC. Time series model to predict burden of viral respiratory illness on a pediatric intensive care unit. *Med Decis Making* 2011; **31**(3): 494–499.
- [6] Dom NC, Hassan AA, Latif ZA, Ismail R. Generating temporal model using climate variables for the prediction of dengue cases in Subang Jaya, Malaysia. *Asian Pac J Trop Dis* 2013; **3**(5): 352–361.
- [7] Box GE, Jenkins GM, Reinsel GC. *Time series analysis: forecasting and control*. New York: John Wiley & Sons, Inc.; 2008.
- [8] Said SE, Dickey DA. Testing for unit roots in autoregressive moving average model of unknown order. *Biometrika* 1984; **71**(3): 599–607.
- [9] Brockwell PJ, Davis RA. *Time series: theory and methods*. Berlin: Springer; 1991.