

## Relative Risk Estimation for Dengue Disease Mapping in Malaysia based on Besag, York and Mollié Model

Samat, N. A.\* and Pei Zhen, W.

*Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak, Malaysia*

### ABSTRACT

In the study of disease mapping, relative risk estimation is the focus of analysis. Many methods have been introduced to estimate relative risk. In this paper, one of the common spatial models known as Besag, York and Mollié (BYM) model is discussed, and its application to dengue data for epidemiology weeks 1 to 52 of the year 2013 for 16 states in Malaysia is studied. Findings show that Selangor has the highest relative risk of dengue in comparison with other states. Data on the estimated relative risks are presented in the form of risk maps which can be used as a tool for the prevention and control of dengue.

*Keywords:* Relative risk estimation, disease mapping, dengue disease, BYM model

### INTRODUCTION

Disease mapping is the visual representation of the spatial and geographic distribution of specified diseases and the corresponding risk. In recent years, disease mapping has become an important tool and technique of epidemiologists, biostatisticians, medical demographics and academics working on disease epidemiology and health services. A disease map displays the geographic areas with high and low incidence, prevalence or mortality rates of a specific diseases. In this paper, dengue is the disease of interest.

Dengue is a major global health problem in which the viruses are transmitted by the infective bite of female mosquito of species genus *Aedes*. This paper investigates the dengue disease mapping in Malaysia. In this research, Besag, York and Mollié (BYM) model is used as the alternative method to estimate the relative risk of dengue occurrence in human populations in Malaysia.

First, a brief introduction to the BYM method will be discussed. Data on dengue cases in Malaysia are used to estimate relative risk using WinBUGS software to carry out Markov chain Monte Carlo (MCMC)

---

*Article history:*

Received: 27 May 2016

Accepted: 14 November 2016

---

*E-mail addresses:*

[norazah@fsm.upsi.edu.my](mailto:norazah@fsm.upsi.edu.my) (Samat, N. A.),

[Peizhen123@yahoo.com](mailto:Peizhen123@yahoo.com) (Pei Zhen, W.)

\*Corresponding Author

computations of Bayesian models. Expected disease risk is presented and analysed using graph and risk map. The disease risk map will display the high and low disease risk incidences of the dengue diseases in 16 states of Malaysia.

### BESAG, YORK AND MOLLIE (BYM) MODEL

In this paper, the Besag, York and Mollié model is applied to the observed dengue disease data to estimate the relative risks for the 16 states in Malaysia. This model was first introduced by Clayton & Kaldor (1987) as an empirical Bayesian and later developed to a fully Bayesian setting by Besag, York and Mollié (Besag, York, & Mollié, 1991).

There are three stages of the BYM model which are described in Lawson, Browne & Rodeiro (2003). In the first stage, for  $i = 1, 2, \dots, M$  study regions, the observed numbers of dengue cases  $y_i$  are assumed to follow a Poisson distribution. The mean and variance for this distribution are the same which is  $e_i \theta_i$ . Here,  $e_i$  represents the expected number of new infective, while  $\theta_i$  refers to the relative risk. Hence, the model is formulated as follows:

$$y_i \sim \text{Poisson}(e_i \theta_i) \tag{1}$$

$$\log(\theta_i) = \alpha + u_i + v_i \tag{2}$$

From Equation (2), the term  $\alpha$  represents the intercept of the model which is an overall level of the relative risk, while the terms  $u_i$  and  $v_i$  represent the correlated and uncorrelated heterogeneities, respectively.

In the second stage, a Bayesian model needs specification of prior distribution for the random effects and intercept term. Here, the distribution model for  $v_i$  does not depend on geographic location and is assumed to follow a normal distribution which can be written as follows.

$$v_i \sim N(0, \tau_v^2) \tag{3}$$

While for the spatial correlation structure,  $u_i$  is used. Here, the risk estimation in any area depends on the spatial correlation between neighbouring areas. This distribution model is called the conditional autoregressive (CAR) model which is written in Equation (4).

$$[u_i | u_{p, i \neq j}, \tau_u^2] \sim N(\bar{u}_i, \tau_i^2) \tag{4}$$

where

$$\bar{u}_i = \frac{\sum_j u_j \omega_{ij}}{\sum_j \omega_{ij}}, \quad \tau_i^2 = \frac{\tau_u^2}{\sum_j \omega_{ij}}$$

For each  $u_i$ , the prior mean is in the form of weighted average of the other  $u_j$  in which  $i \neq j$ . The weightage,  $w_{ij}$ , defines the relationship between the areas  $i$  and its adjacent areas  $j$  based on the following scheme.

$$w_{ij} = \begin{cases} 1 & \text{if } i, j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

The variability for the random effect is controlled by the precision parameters  $\tau_u^2$ .

Finally, in the final stage, as a fully Bayesian analysis, the model specification is completed by adding hyper-prior distributions to the parameters  $\tau_v^2$  and  $\tau_u^2$ . As suggested by Johnson (2004), the Gamma hyper prior distribution was chosen.

## APPLICATION OF RELATIVE RISK ESTIMATION TO DENGUE DISEASE MAPPING

This section displays the result of relative risk estimation based on the application of Besag, York and Mollié (BYM) model to observed dengue data in Malaysia.

### Data Set

The data used in this analysis were provided by the Ministry of Health Malaysia, consisting of the observed count dengue data for every state in Malaysia from epidemiology week 1 to epidemiology week 52 of the year 2013. 16 states were involved in this study, which were Perlis, Kedah, Pulau Pinang, Perak, Kelantan, Terengganu, Pahang, Selangor, Kuala Lumpur, Putrajaya, Negeri Sembilan, Melaka, Johor, Sarawak, Labuan and Sabah.

### Analysis and Results: Relative risk estimation based on BYM Model

This section presents the results of the analysis based on the BYM model. Figure 1 depicts the time series plot of cases of dengue disease for all 16 states in Malaysia from epidemiology week 1 to epidemiology week 52. It can be seen for every epidemiology week, Selangor had the highest number of cases. The number of cases shows an increasing trend and reached the peak in epidemiology week 52 with 1408 cases. While for other states, most states had less than 200 cases for most of the epidemiology weeks in 2013. The lowest number of cases is in Labuan, which reported only 14 cases of dengue. It is free from dengue for the first 20 weeks to increase to by week 49.

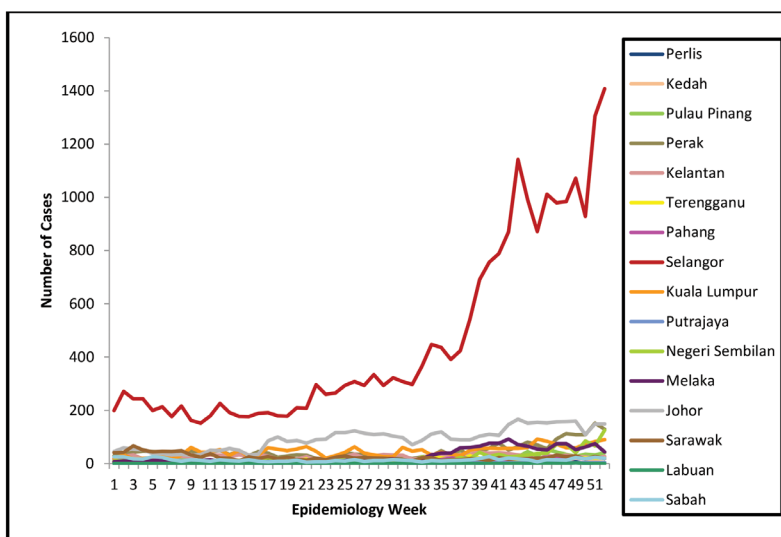


Figure 1. Time series plot for the number of dengue cases for all states in Malaysia

Figure 2 represents the time series graph of the estimated relative risk based on Besag, York and Mollié (BYM) model for dengue disease in Malaysia in 2013. The graph shows that Selangor had the highest number of relative risk of contracting dengue having relative risks above 2 for most epidemiology weeks. Based on the definition of relative risk found in Samat & Percy (2012) and Samat & Ma'arof (2013), a relative risk greater than 2 means that people in these states have a higher likelihood of contracting dengue compared to the rest of the country.

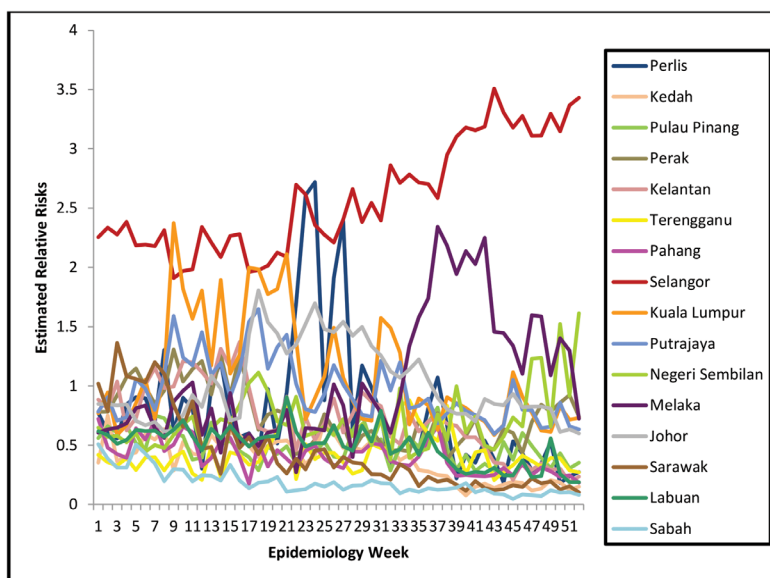


Figure 2. Time series plots for the estimated relative risks based on the BYM Model for all states in Malaysia

Table 1 shows the posterior means for the relative risk in 16 states of Malaysia obtained using the BYM model. The results give specific estimated relative risks for epidemiology week 52 and it can be seen that the state of Selangor has the highest risk.

Table 1  
*Estimated Relative Risks Based On BYM Model for Epidemiology Week 52*

States	Relative Risk	Standard Deviation
Perlis	0.2738	0.1154
Kedah	0.1507	0.0313
Pulau Pinang	0.3535	0.0539
Perak	0.7407	0.0648
Kelantan	0.2354	0.0436
Terengganu	0.2723	0.0571
Pahang	0.1830	0.0398
Selangor	3.4310	0.0924
Kuala Lumpur	0.7309	0.0771
Putrajaya	0.6341	0.2851
Negeri Sembilan	1.6130	0.1439
Melaka	0.7242	0.1081
Johor	0.6006	0.0496
Sarawak	0.1018	0.0230
Labuan	0.1879	0.1242
Sabah	0.0812	0.0180

### Dengue Risk Maps

In this section, the estimated relative risks are displayed in a form of risk maps. It provides a quick overview of the collected information, so that further attention and action can be given to these priority states. The interpretation of the maps is based on the tone of colours used. Darkest colour represents very high risk while the lightest colour represents very low risk.

Figures 3 and 4 show thematic risk maps for dengue disease based on the number of observed cases and the expected relative risk estimated by BYM model, respectively. Both maps demonstrate the results for epidemiology week 52 in the year 2013 for all 16 states in Malaysia. The relative risks are categorized into five different levels of risk, which are very low, low, medium, high and very high with respective intervals.

Figure 3 represents the risk map for dengue disease based on the number of cases for epidemiology week 52. The five category of risk are determined based on the intervals of  $[0,300)$ ,  $[300,600)$ ,  $[600,900)$ ,  $[900,1200)$ , and  $[1200,\infty)$ , respectively. This figure depicts that state with very high risk for dengue occurrences is Selangor. The other 15 states are classified as very low risk areas.

Furthermore, Figure 4 represents the risk map for dengue disease based on BYM model for epidemiology week 52. There are five categories of risk area which are determined based on the intervals of  $[0.0, 0.5)$ ,  $[0.5, 1.0)$ ,  $[1.0, 1.5)$ ,  $[1.5, 2.0)$ , and  $[2.0, \infty)$ , respectively. From this figure, Selangor showed has a very high risk area of dengue occurrence, followed by Negeri Sembilan with high risk. States with low risk include Perak, Kuala Lumpur, Putrajaya, Melaka and Johor. The other states with very low risk are Perlis, Pulau Pinang, Kelantan, Terengganu, Sarawak, and Labuan in Sabah.

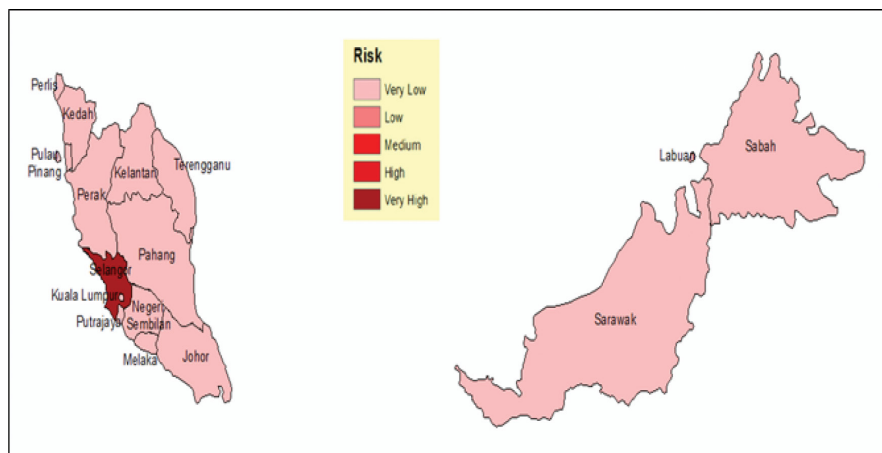


Figure 3. Risk map for dengue disease based on number of cases for Epidemiology Week 52

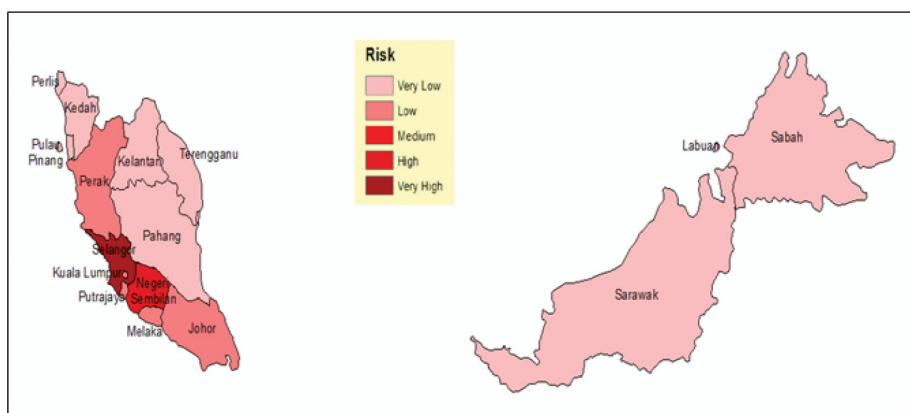


Figure 4. Risk map for dengue disease based on BYM model for Epidemiology Week 52

Comparisons among the map of Figures 3 and 4, specifically for epidemiology week 52, demonstrate no obvious differences in terms of the estimated risks for dengue disease in 16 states. However, further action and precautionary steps need to be implemented in every state, particularly in Selangor in which the people has very high risk to contract with dengue disease compared to people in overall population.

## DISCUSSION AND CONCLUSION

Dengue is a major public health issue in Malaysia and is likely to remain endemic for a long time. The production of an accurate disease map relies on statistical modelling used to estimate the relative risk. The purpose of exploratory disease mapping is to provide insight into possible causes, effects and trends, and impression of the geographical or spatial distribution of disease.

In this paper, Besag, York and Mollié (BYM) model was used to estimate the relative risk of dengue. The results of the analysis showed that Selangor has very high risk of dengue, arising from higher rainfall, humidity, temperature, and urbanization. In order to avoid the spread of dengue surrounding area should be clean and free of water reservoir, requiring public and community commitment to hygiene. Disease risk maps can be useful for prevention and control strategies of dengue.

## ACKNOWLEDGEMENT

The authors acknowledge Universiti Pendidikan Sultan Idris and the Ministry of Higher Education in Malaysia for their financial support in respect of this study.

## REFERENCES

- Besag, J., York, J., & Mollié, A. (1991). Bayesian Image Restoration, With Two Applications in Spatial Statistics. *Annals Of The Institute Of Statistical Mathematics*, 43(1), 1-20.
- Clayton, D., & Kaldor, J. (1987). Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics*, 43(3), 671-681.
- Johnson, G. D. (2004). Small Area Mapping of Prostate Cancer Incidence in New York State (USA) using Fully Bayesian Hierarchical Modelling. *International Journal of Health Geographics*, 3(1), 1-12.
- Lawson, A. B., Browne, W. J., & Rodeiro, C. L. V. (2003). *Disease mapping with WinBUGS and MLwiN* (Vol. 11). England: John Wiley & Sons.
- Samat, N. A., & Ma'arof, S. M. I. (2013). Dengue Disease Mapping with Standardized Morbidity Ratio and Poisson-gamma Model: An Analysis of Dengue Disease in Perak, Malaysia. *International Journal of Mathematical Science and Engineering*, 7(8), 1299-1303.
- Samat, N. A., & Percy, D. F. (2012). Dengue Disease Mapping in Malaysia Based on Stochastic SIR Models in Human Populations. *International Conference on Statistics in Science, Business, and Engineering (ICSSBE 2012)* (pp. 1-5). IEEE.

