Sains Malaysiana 40(8)(2011): 937–943

Evaluation of the Spatial Risk Factors for High Incidence of Dengue Fever and Dengue Hemorrhagic Fever Using GIS Application

(Penilaian Faktor-faktor Risiko Ruangan Bagi Kejadian Penyakit Demam Denggi dan Demam Denggi Berdarah Menggunakan Aplikasi GIS)

AZIZ SHAFIE*

ABSTRACT

In Malaysia, the incidence of Dengue Fever (DF) and Dengue Hemorrhagic Fever (DHF) have risen dramatically in the last twenty years. With the use of Geographical Information System an explanation for the spread and control of these diseases can be obtained. This study aims to develop a spatial modeling that can predict the risks for DF and DHF based on environmental factors such as physical surroundings, land use, rainfall, temperature and GIS application using logistic regression. A total of 16 variables were used in the process of spatial modeling development. At the significant level of 0.05, the results of logistic regression showed that only 10 out of 16 significant variables in the modeling process. The accuracy of the resulting model is 70.3%. A crucial feature of this study is a risk area map for incidence of DF and DHF in the study area. This study also highlights the application of spatial analysis in planning and implementing the process for the prevention and control activities of DF and DHF in Malaysia.

Keywords: Dengue fever; dengue hemorrhagic fever; geographic information system (GIS); risk mapping; spatial modeling

ABSTRAK

Di Malaysia, kejadian kes Demam Denggi (DD) dan Demam Denggi Berdarah (DBD) telah meningkat secara dramatik dalam dua puluh tahun terakhir. Dengan penggunaan Sistem Maklumat Geografi (GIS), penjelasan tentang penyebaran dan pengawalan penyakit ini boleh dilakukan. Kajian ini bertujuan untuk membangunkan model ruangan yang dapat meramalkan risiko kejadian penyakit DF dan DBD berdasarkan faktor-faktor persekitaran seperti persekitaran fizikal, guna tanah, hujan, suhu dan aplikasi GIS menggunakan kaedah regreasi logistik. Sebanyak 16 pemboleh telah digunakan dalam proses pembentukan model ruangan. Pada aras signifikan 0.05, keputusan regrasi logistik mendapati hanya 10 daripada 16 pembolehubah signifikan dalam proses pembentukan model. Ketepatan model yang dihasilkan ialah sebanyak 70.3%. Hasil penting kajian ini adalah peta risiko kejadian DF dan DBD di kawasan kajian. Kajian ini juga menyoroti penggunaan analisis ruangan dalam proses perancangan dan pelaksanaan aktiviti untuk pencegahan dan kawalan penyakit DD dan DDB di Malyasia.

Kata kunci: Demam denggi; demam denggi berdarah; pemetaan risiko; permodelan ruangan: sistem maklumat geografi (GIS)

INTRODUCTION

Over the years, the incidences of dengue fever (DF) and dengue hemorrhagic fever (DHF) reported in Malaysia have been on the rise. In 2008 alone, Malaysia reported 49,335 cases with 112 deaths. Until 5 September 2009, there were 30,110 cases with 69 deaths reported for the whole country (MOH 2009). The incidence of DF and DHF in Malaysia are caused by viruses transmitted through the bites of infective female *Aedes albopictus* and *Aedes aegypti* mosquitoes. These mosquitoes only breed in areas of clear water storage. According to Wegbreit and Reisen (2000), Pedigo (1999), and Thomson and Connor (2000), the breeding of the mosquito vector species is strongly associated with total rainfall, altitude, tidal force, land use and other factors influencing the breeding behavior of mosquito. The Ministry of Health, Malaysia has identified areas such as parks, cemeteries, vacant land, public infrastructure areas, construction sites and rubbish dumping sites as favorable areas for mosquito breeding in Malaysia.

The mosquito population as well as the density of mosquitoes are the two main indicators that determine the risk of dengue in an area. The technique of measuring mosquito population requires enormous effort and time and this has to be done continuously. At the same time, the decision to implement vector control activities must be carried out immediately so that the incidence of dengue can be prevented. With the advancement in information technology, computer, communication systems, especially GIS, management and prevention activities for these diseases can be done immediately, without delay. The use of GIS allows us to integrate fully the environmental and time elements related to mosquito breeding and disease spreading without depending on indicators like mosquito population and mosquito density. GIS is a computer-based system that can integrate various spatial and non-spatial data to study the mosquitoes' habitats. There have been a growing number of studies utilizing GIS to analyze *Aedes* breeding habitats and dengue fever risk (Aziz 2008; Bohra & Andrianasolo 2001; Mohammad et al. 2003; Pei et al. 2009 and Tzai et al. 2006).

This study utilised a GIS application to integrate the locational data for the incidence of DF and DHF, along with data on land use and demography to identify areas that are prone to these diseases. It however excludes utilising data on population density of mosquitoes, which is the current practice in Malaysia. The ability to identify the probable areas for dengue incidence allows for a prompt and effective implementation of disease management and prevention activities.

MATERIALS AND METHODS

STUDY AREA

This study is divided into two sections namely the development of spatial model for areas with high dengue risk and the prediction of risk for DF and DHF in each area. By utilising the data on dengue incidences in Georgetown area from 2000 till 2004, a spatial model for dengue risk prediction is developed using backward logistic regression equation. Based on this spatial model, dengue incidence for the whole area of study, which is Penang island - one of the states of Malaysia comprising of an area of about 240.5 km^2, has been calculated (Figure 1).

DATA REQUIREMENT

Data on 882 cases of DF and DHF in Georgetown area from 2000 to 2004, provided by Ministry of Health, Malaysia have been used as the basis in developing the spatial distribution model. Based on patients' residential addresses, the location of dengue incidences were determined using GPS (Garmin GPSMAP 60CSx) equipments. ArcGIS 9.3 (ESRI, Inc., Redlands, CA, USA) software was then used to map the location for the incidence of these diseases. Other data used in this study are digital data for land use, acquired from Penang Geographic Information System (PEGIS), and data for rainfall and temperature, obtained from Malaysian Meteorological Department. Administrative maps at states, districts and territories levels from the Department of Survey and Mapping, Malaysia and population data, provided by the Department of Statistics, Malaysia are other equally essential materials for this study. In developing the spatial distribution model, the SPSS 17.0 (SPSS Inc. Chicago, IL, USA) software was used.

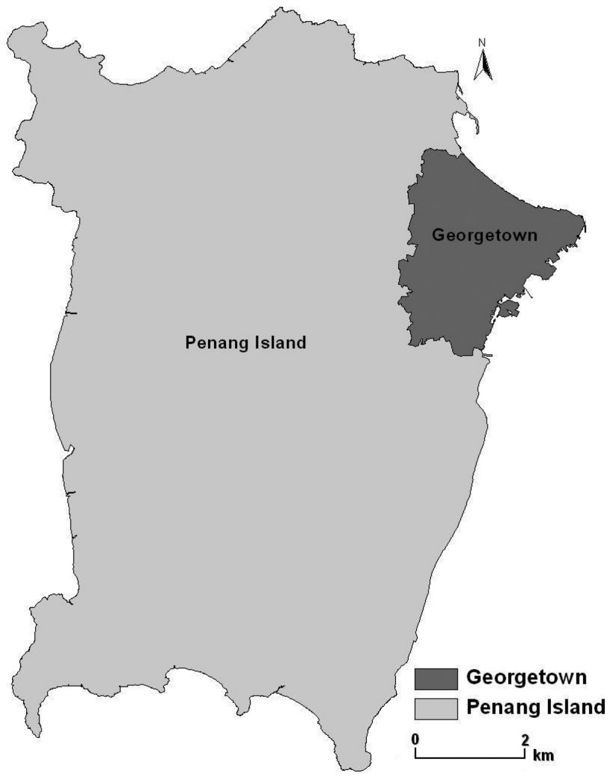

FIGURE 1. Study area

DEVELOPMENT OF SPATIAL DISTRIBUTION MODEL

In developing the spatial distribution model, the stepwise logistic regression analysis was used. According to Field (2005), the stepwise logistic regression equation to estimate the risk of disease is represented by:

$$Y = \frac{1}{1 + e^{-(Bo+B1x1+B2x2+Bnxn)}}$$

where Y refers to the risk of disease incidence, b_0 refers to intercept and $x_1, x_2, \ldots, x_n$ are independent variables. The location risk for dengue in each site was calculated based on the distance of each of 882 real cases to the location of the independent variables. These independent variables represent locations which include areas that influence incidence of diseases, areas for mosquito breeding, areas related to demography factor and others. In theory, the set of control cases are needed in order to perform logistic regression. Control case is a false case created solely to offset the modeling process. Typically, the number of control cases is equal to the number of actual cases used in the modeling process. Location of control cases must be balanced and cover the whole area of research. In this study, 882 cases were produced randomly using ArcGIS software. The location of these control cases was spread across whole study area.

Previous studies (Aziz 2008 & Richards 2010; Barbazan et al. 2003; Beaty & Marquard 1996; Bora & Andriasolo 2001; Chan 1985; Cheong 1986; Gubler & Kuno 1997; Haliza 2000; Mohammad et al. 2003; Mondini & Chiaravalloti 2008; Rotela et al. 2007; Tan & Teo 1998; Wegbreit & Reisen 2000; Wellmer 1983) have identified 16 independent variables essential in developing the spatial distribution model and they include the following:

1. Proximity of the real cases and control cases to cemeteries [CEMETERY]
2. Proximity of the real cases and control cases to swamp and forest areas [FOREST]
3. Proximity of the real cases and control cases to river bank areas [RIVER]
4. Proximity of the real cases and control cases to places of religious areas [RELIGIOUS]
5. Proximity of the real cases and control cases to public infrastructure areas [INFRA]
6. Proximity of the real cases and control cases to public utility areas [UTILITY]
7. Proximity of the real cases and control cases to parks and vacant land [VACANT]
8. Proximity of the real cases and control cases to school areas [SCHOOL]
9. Proximity of the real cases and control cases to industrial areas [INDUSTRY]
10. Proximity of the real cases and control cases to hospital areas [HOSPITAL]
11. Proximity of the real cases and control cases to public market areas [MARKET]
12. Proximity of the real cases and control cases to business areas [BUSINESS]
13. Proximity of the real cases and control cases to government institution areas [GOVERNMENT]
14. Population density [DENSITY]
15. Rainfall [RAIN]
16. Temperature [TEMPERATURE]

Table 1 shows the information on the actual distance of real cases and control cases to 16 independent variables.

CALCULATION OF DISEASE INCIDENCE

Based on the spatial distribution model developed for Georgetown area, the incidence of diseases for the whole of Penang area was estimated. For this reason, the Penang area was divided into square cells of area 250 meter width and 250 meter length. Overall, there were about 5 015 cells developed to represent the whole of Penang area. The centre (centroid) of each cell was used as the basis for calculating the distance of each cell to the location of the independent variables. Accordingly, with this information the disease risk for DF and DHF in each cell has been calculated.

RESULTS

Based on the stepwise logistic regression analysis, only 10 out of 16 variables significantly influence the probability of the incidence of DF and DHF (Table 2). The significant variables are proximity to cemeteries, proximity to swamp and forest areas, proximity to river bank areas, proximity to public infrastructure areas, proximity to parks, proximity to school areas, proximity to industrial areas, proximity

TABLE 1. Actual distance of real case and control cases to independent variables

Independent Variables	Mean (m)		Std. Deviation (m)	
	Real Cases	Control Cases	Real Cases	Control Cases
Business	79.9	193.3	95.2	236.1
Infra	141.4	229.4	106.8	226.4
Religous	167.6	256.0	114.0	210.5
Forest	202.6	260.8	118.6	181.0
Vacant	206.5	334.6	136.3	255.5
Industy	212.9	456.4	189.2	347.6
School	233.6	334.9	120.4	262.0
Utility	267.4	334.5	142.2	191.6
River	396.6	427.4	352.0	418.2
Cemetery	474.6	578.5	266.0	292.1
Government	489.0	629.6	347.0	416.0
Market	574.1	746.7	243.2	364.6
Hospital	1167.6	1045.7	614.6	637.6
Forest	1174.2	649.1	1486.3	1204.7
Rain	59.7	72.18	66.9	115.9
Temperature	31.5	31.1	0.9	1.0

to hospital areas, proximity to government institution areas and population density. Hence, the equation for the spatial distribution model for incidence of disease (Y) can be stated as follow:

$$Y = \frac{1}{1 + e^{-Z}}$$

where z = 3.0822 - 2.0175[INFRA] - 0.3391 [HOSPITAL] - 0.3830[GOVERNMENT] - 1.2208[SCHOOL] - 1.5375[VACANT] - 2.7262[INDUSTRY] - 1.0403[CEMETERY] - 0.9749[RIVER] + 0.0002[DENSITY] +0.8210[FOREST]

The accuracy percentage of the model is 70.3% with the level of significance set at 0.05 and the cut value are 0.5 (Table 3).

The method of spatial analysis using ArcGIS software was implemented to generate spatial data for ten significant variables in the process of developing a model for the whole of Penang area. Using the spatial model equation, the incidence of diseases for the whole area can be calculated and mapped (Figure 2). Locations with a high risk for the disease were generally clustered in certain areas. High risk for the disease is in the northeast, which is in Georgetown and Air Hitam. Georgetown was a center for business activities, services, administrative and residential areas with high concentrations. Air Hitam is located on the right side of Georgetown, which is also a densely populated area. Other areas are in the southeast of Penang, in the Jelutung-Gelugor-Sungai Dua, which is also a major residential area. High-risk areas are also found in the northernmost part of Penang, in Teluk Bahang and Batu Feringgi. Meanwhile,

TABLE 2. The result of spatial model using logistic regression

Variables		B	S.E.	Wald	df	Sig.	Exp(B)
Step 5(a)	River	-.974	.000	6.804	1	.009	.999
	Density	.000	.000	14.659	1	.000	.999
	Vacant	-1.537	.001	3.384	1	.006	1.001
	Hospital	-.033	.000	7.372	1	.007	1.000
	Industry	-2.726	.000	28.021	1	.000	.998
	Infra	-2.017	.001	4.374	1	.036	.998
	Government	-.383	.000	23.785	1	.000	1.002
	Cemetery	-1.040	.000	13.521	1	.000	.999
	Forest	0.821	.001	10.279	1	.001	.998
	School	-1.220	.000	4.083	1	.043	.999
	Constant	3.082	.026	39.952	1	.000	4.177

TABLE 3. The Accuracy of Spatial Model

Step	Observed		Predicted MODEL 0	Predicted MODEL 1	Predicted Percentage Correct
Step 1	MODEL	0	478	288	62.4
		1	171	595	77.7
	Overall Percentage				70.0
Step 2	MODEL	0	477	289	62.3
		1	169	597	77.9
	Overall Percentage				70.1
Step 3	MODEL	0	480	286	62.7
		1	170	596	77.8
	Overall Percentage				70.2
Step 4	MODEL	0	478	288	62.4
		1	169	597	77.9
	Overall Percentage				70.2
Step 5	MODEL	0	475	291	62.0
		1	164	602	78.6
	Overall Percentage				70.3

a The cut value is 0.5

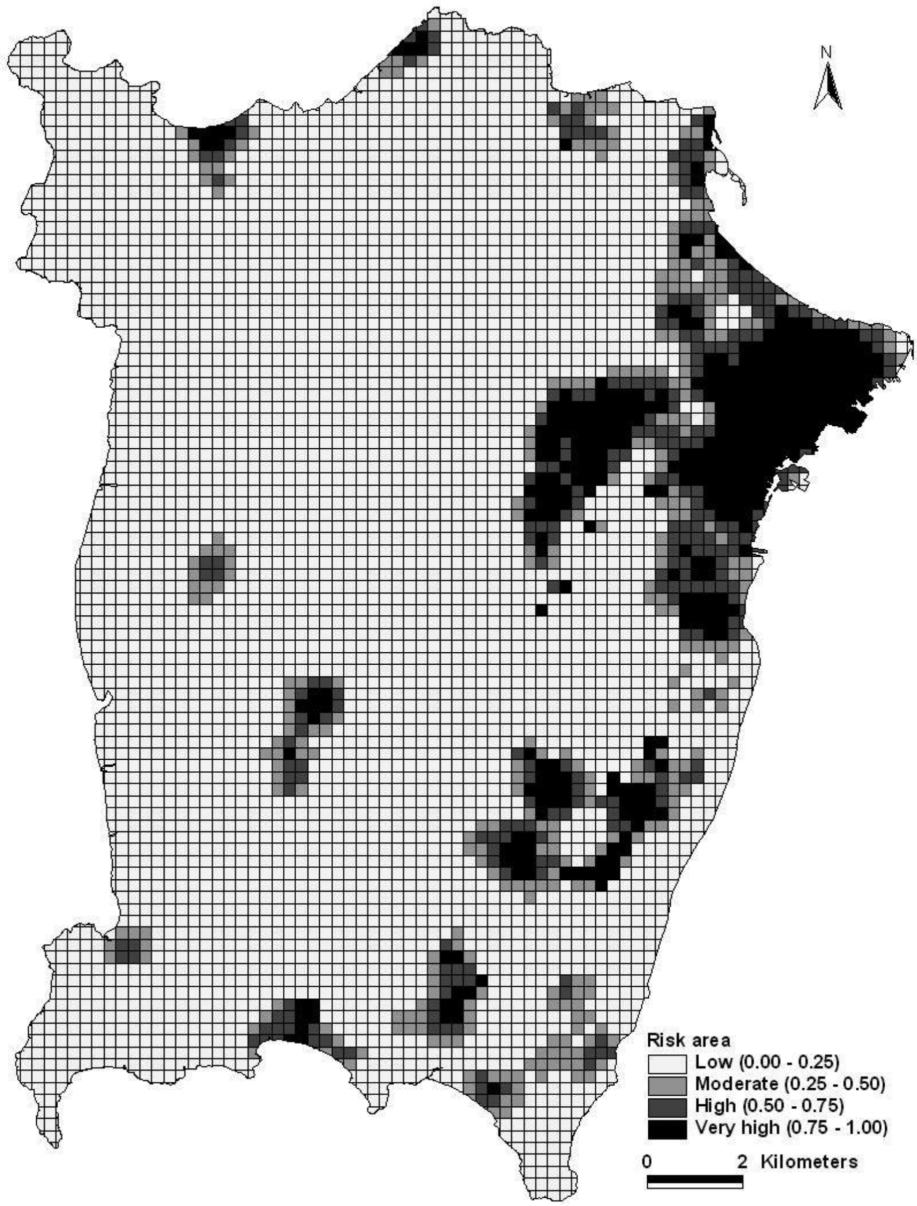

FIGURE 2. Risk of Disease Incidence

the area around the Bayan Lepas Industrial Zone is located in the south of Penang, an area with a high risk.

DISCUSSION

The mosquito breeding and the transmission of dengue fever are closely associated with human beings and the environment. Activities for the prevention and control of dengue carried out in Malaysia do not give special attention to environmental factors and demography. The environmental and demographic factors were rank high in influencing the mosquitoes breeding, the incidence and spread of dengue, especially in tropical countries (Kolivras 2006; Richards et al. 2010; Troyo et al. 2009). This study has fully utilized environmental and demographic elements in determining the risk for DF and DHF, taking into account the area guidelines prioritised by the MOH, Malaysia for planning and implementing preventive and control activities of these diseases. School areas, hospital areas, places of worship, forest and swamp areas, vacant land and parks, cemeteries, and other high human concentration areas are the independent variables utilised in this study. All these areas are priority areas for MOH, Malaysia to implement preventive and control activities for vector-borne diseases (Tham 2001).

Spatial model were developed is based on the occurrence of cases of disease in Georgetown. However, this spatial model still can also be applied to the whole study area. Environmental characteristics that influence the occurrence and spread of disease in the Georgetown area will also provide the same relationship for the entire study area. All independent variables used in modeling processes, such schools, forest areas, businesses areas and industrial areas were found throughout the study area. Based on the spatial relationship term, the model developed is expected to be able to predict the risk of disease cases throughout the island of Penang.

By integrating GIS application and statistical modeling, this study proved that it is possible to predict the risk areas without relying on the information about the density of mosquitoes or the occurrence of cases. Spatial relationship between the occurrence of disease cases that have occurred with the environmental elements related to the breeding of mosquitoes and disease transmission has been completely integrated in this study. It is sufficient to identify the risk areas for the whole study area. The risk areas are presented by developing a map showing the incidence of DF and DHF in the study area. This method is also used in several places to enhance the effectiveness of the activities of dengue prevention and control (Benjamin et al. 2010; Manguin et al. 2010; Mondini & Chiaravalloti, 2008; Pei-Chih Wua et al. 2009; Peterson et al. 2005; Schroder et al. 2006 & Tzai et al. 2010).

This map would provides useful information to health authorities and could assist in focusing and implementing control and preventive activities to monitor and control the incidence of dengue precisely and effectively, especially in the event when there is no report on dengue cases. Furthermore, this study provides a new dimension to the health authorities in Malaysia, specifically in the potential of using GIS application, GPS and other applications to develop strategies for the implementation of preventive and control activities, not only for DF and DHF, but also for other vector borne diseases in Malaysia.

REFERENCES

Aziz, S. 2008. Geographic Information System (GIS) Application to Identify High Risk Area of Dengue and Dengue Hemorrhagic Fever In Georgetown, Penang. PhD Thesis. Universiti Sains Malaysia, Geography Section. (Unpublished)

Barbazan, P., Yoksun, S. & Gonzalez, J.P. 2002. Dengue hemorrhagic fever epidemiology in Thailand. Description and forecasting of epidemics. *Microbes and Infectors* 4:669-705.

Beaty, J & Marquardt, C. 996. *The Biology of Disease Vectors*. Colorado: University Press of Colorado.

Benjamin, G.J., Joel, A.M., Erick, C., Daniel, A.G. & Robert, J.N. 2010. Geomapping generalize deigenvalue frequency distributions for predicting prolific *Aedes albopictus* and *Culex quinquefasciatus* habitats based on spatiotemporal field. *ActaTropica* doi:10.1016/j.actatropica.2010.10.002.

Bohra, A. & Andrianasolo, H. 2001. Application of GIS in Modeling of Dengue Risk Based On Sociocultural Data: Case of Jalore, Rajasthan, India. *Dengue Bulletin* 25: 92-102.

Chan, K.L. 1985. *Singapore's Dengue Hemorrhagic Fever Control Programme A case Study Of The Successful Control Of Aedes Aegypti And Aedesm Alboppictus Using Mainly Environmental Measures As A Part Of Integrated Vector Control*. Tokyo: Southeast Asian Medical Information Centre (SEAMIC).

Cheong, W.H. 1986. Dengue Hemorrhagic Fevers in Malaysia. In *Dengue Fever Studies in Malaysia*. In *Dengue Fever Studies in Malaysia*. Edited by Rudnick & Lim. Kuala Lumpur: Institute for Medical Research.

Field, A. 2005. *Discovering Statistics Using SPSS*. London: Sage Publication.

Gubler, D.J. & Kuno, G. 1997. *Dengue and Dengue Hemorrhagic Fever.* Cambridge: CAB International.

Haliza, M.R. 2001. Perception, Knowledge and Behavioral Aspects of Dengue Control in Urban Communities in Kuala Lumpur. *Proceedings of the Behavioral Interventions in Dengue Control in Malaysia,* Universiti Sains Malaysia, Pulau Pinang: 125-140.

Kolivras, K.N. 2006. Mosquito habitat and dengue risk potential in Hawaii: a conceptual framework and GIS application. *Professional Geographer* 58: 139-154.

Manguin, S., Bangs, M.J., Pothikasikorn, J. & Chareonviriyaphap, T. 2010. Review on global co-transmission of human Plasmodium species and *Wuchereria bancrofti* by *Anopheles* mosquitoes. *Infection, Genetics and Evolution* 10: 159-177.

Ministry of Health (MOH). 2009. *Current Situation of Dengue and Dengue Hemorrhagic Fever.* Kuala Lumpur: Ministry of Health.

Mohammad, A., Wagatsuma, Y., Emch, M. & Breiman, R. 2003. Use of A Geographic Information System For Defining Spatial Risk For Dengue Transmission in Bangladesh: Role For Aedes albopictus in An Urban Outbreak. *American Journal of Tropical Medicine and Hygiene* 69: 634-640.

Mondini, A. & Chiaravalloti-Neto, F. 2008. Spatial correlation of incidence of dengue with socioeconomic, demographic and environmental variables in a Brazilian city. *Science of the Total Environment* 393: 241-248.

Pang, T., Hamimah, H. & Shivaji, R. 1998. *Dengue and Dengue Hemorrhagic Fever*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Pedigo, L. 1999. *Entomology And Pest Management*. Upper Saddle River: Prentice Hall.

Pei-Chih Wua, Jinn-Guey Layb, How-Ran Guoc, Chuan-Yao Lind, Shih-Chun Lung & Huey-Jen Su 2009. Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan. *Science of the Total Environment* 407: 2224-2233.

Peterson, A.T., Campos, C.M., Yoshinori Nakazawac & Meyerd, E.M. 2005. Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 99: 647-655.

Richards, E.E., Masuoka, P., Major, D.B., Smith, M., Klein, T.A., Kim, H.C., Anyamba, A. & Grieco, J. (2010). The relationship between mosquito abundance and rice field density in the Republic of Korea. *International Journal of Health Geographic* 9: 32-42.

Rotela, C., Fouque, F., Lamfri, M., Sabatier, P., Introini, V., Zaidenberg, M. & Scavuzzo, C. 2007. Space–time analysis of the dengue spreading dynamics in the 2004 Tartagal outbreak, Northern Argentina. *Acta Tropica* 103: 1-13.

Schroder, W. 2006. GIS, geostatistics, metadata banking, and tree-based models for data analysis and mapping in environmental monitoring and epidemiology. *International Journal of Medical Microbiology* 296 (S1): 23-36

Tan, B.T. & Teo, B.T. 1986. *Modus Operandi in Aedes Surveillance and Control in Singapore*. Singapore: Institute of Environmental Epidemiology, Ministry of Environment.

Tham, A.S. 2001. Issues and Challenges in Aedes Surveillance and Control. *Proceedings of the Behavioral Interventions* *in Dengue Control in Malaysia*. Universiti Sains Malaysia, Penang: 15-26.

Thomson, M.C. & Connor, S.J. 2000. Environmental information systems for the control of arthropod vectors of disease. *Medical and Veterinary Entomology* 14: 227-244.

Troyo, A., Fuller, D.O., Arguedas, O.C., Solano, M.E. & Beier, J.C. 2009. Urban structure and dengue fever in Puntarenas, Costa Rica. *Singapore Journal of Tropical Geography* 30(2): 265-282.

Tzai, H.W., Neal, H.L., Chun, H.L., Chwan, C.K. & Ming, D.S. 2006. Spatial mapping of temporal risk characteristics to improve environmental health risk identification: A case study of a dengue epidemic in Taiwan. *Science of the Total Environment* 367: 631-640.

Tzai-Hung Wen, Neal, H.L., Day-Yu Chao, Kao-Pin Hwang, Chih-Chun Kan, Katherine Chun-Min Lin, Joseph Tsung-Shu Wu, Scott Yan-Jang Huang, I-Chun Fan & Chwan-Chuen King. (2010). Spatial–temporal patterns of dengue in areas at risk of dengue hemorrhagic fever in Kaohsiung, Taiwan, 2002. *International Journal of Infectious Diseases* 14: 334-343.

Wegbreit, K.C. & Reisen, W.K. 2000. Relationship Among Weather, Mosquito Abundance, and Encephalitis Virus Activity in California: Kern County 1990-1998. *Journal of the American Mosquito Control Association* 16: 22-27.

Wellmer, H. 1983. *Dengue Haemorrhagic Fever in Thailand. Geomedical Observations on Developments Over the Period 1970-1979*. Berlin: Springer-Verlag; 1983.

Department of Geography
Faculty of Arts and Social Sciences
University of Malaya
50603 Kuala Lumpur, Malaysia

*Corresponding author; email: azizs@um.edu.my

Received: 25 May 2010
Accepted: 29 November 2010