Spatio-Temporal Outbreak Detection of Dengue Hemorrhagic Fever in Phayao, Thailand

Chaikaew, N.
Geographic Information Science, School of Information and Communication Technology, University of Phayao, Phayao, Thailand, E-mail: Nakarin.ch@up.ac.th

Abstract
The objective of this study was to detect the spatio-temporal outbreak of Dengue Hemorrhagic Fever (DHF) in Phayao province, Thailand. This study utilized data collected from DHF patients at village level in 2010. A geographic information system (GIS) was used to prepare the spatial data for analyzing the DHF outbreaks. In data analysis, the space-time permutation scan statistic was used to determine whether the DHF signals occurred by chance in space and time. In the study, the detected DHF outbreaks were observed at statistically significant levels ($p \leq 0.05$), and were found in different places and times of the study area.

1. Introduction

Dengue hemorrhagic fever (DHF) is a major public health concern in Thailand. The Ministry of Public Health has been attempting to monitor and control this disease for many years. The objective of this study was to detect DHF outbreaks in Phayao province, northern Thailand when available signals for a particular area are being monitored. The methodology and the results could be useful for public health officers to develop a system to monitor and prevent DHF outbreaks. DHF is a systemic and dynamic disease, with potentially serious infection, spread by the *Aedes aegypti* (and, less commonly, *Ae. albopictus or Ae. polynesiensis*) mosquito. The infection occurs when a person is bitten by a mosquito that is infected with dengue viruses (DEN 1, 2, 3 and 4) (WHO, 2009). DHF is an important cause of morbidity among children in several countries of the world; with an estimate of more than 500,000 cases occurring annually (WHO, 2006). It is widespread especially in tropical and sub-tropical regions such as Africa, South East Asia, Western Pacific, Latin and Central America, with rapid population growth, increased urbanization, and the lack of effective mosquito control and limited of public health systems identified as significant issues (WHO, 2009). In Thailand, DHF has been a major public health problem for many years (Bureau of Epidemiology, 2010). The Division of Epidemiology, Ministry of Public Health, reported that the number of DHF patients increased yearly in the period 2006-2009. There were 46,829, 65,581, 89,626 and 56,651 reported cases respectively, with 59, 95, 102 and 50 corresponding deaths. In 2010, the DHF incidence was estimated to be 112,992 cases and 139 deaths, with the highest incidence rates occurring in the southern and northern region of Thailand (338.53 and 179.53 per 100,000 individuals, respectively). Phayao province is the endemic area of DHF in the north of Thailand (Figure 1), covering an area of 6,335.06 sq. km. It is mostly covered with forested mountain with an approximate elevation at 300 meters above mean sea level. It is divided administratively into nine districts (amphoe), 68 sub-districts (tambon), and 818 villages (mooban). The epidemic pattern of DHF in Phayao has alternated by year, from 2006 to 2010. The highest number of cases was recorded in 2010, particularly in Muang Phayao, Dok Khamtai, and Phu Kamyao districts, with incidence rates of 637.84, 577.17 and 298.24 per 100,000 individuals, respectively. The incidence of DHF is marked by high variability at district, sub-district and village levels. DHF incidence in this area has increased in some time periods and specific areas in recent years. It was most common during the rainy season (May-October) as disease vectors need clean, standing water to reproduce. The peak of the season was usually from June through August (Bureau of Epidemiology, 2010). The space-time analytical methods in use for early detection of disease outbreaks will help to detect and evaluate DHF outbreaks when an available signal for a particular area is being monitored and to assist public health officers in the surveillance, planning, control and prevention of DHF outbreaks.

ISSN 1686-6576 / © Geoinformatics International

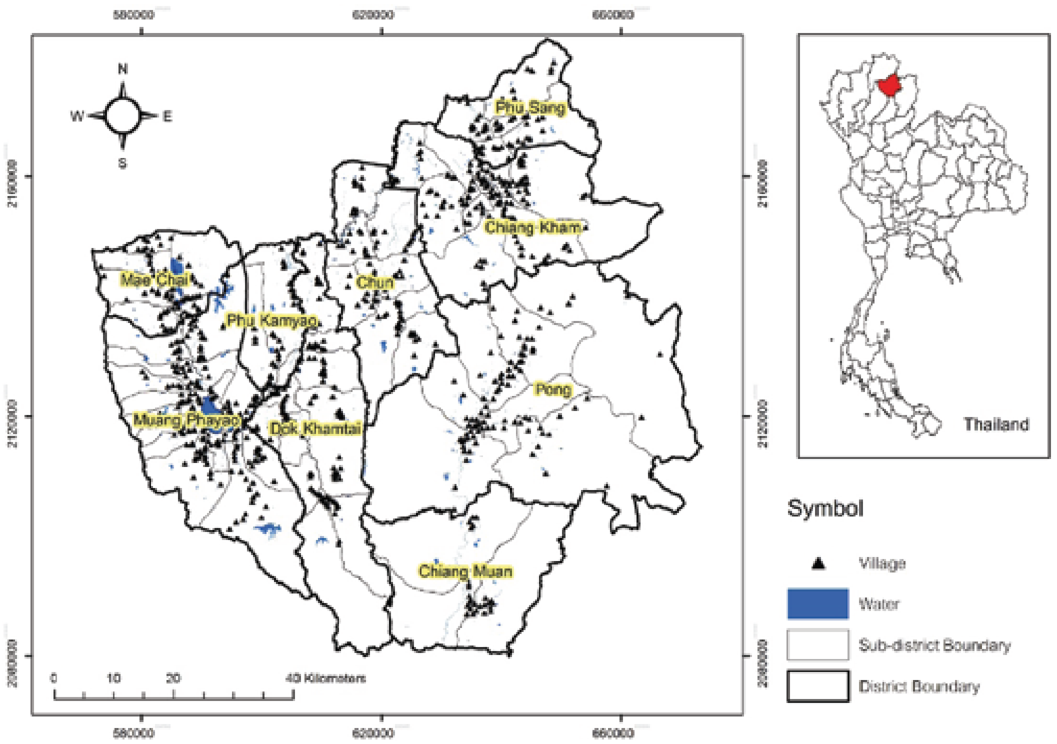

Figure 1: Phayao, Thailand

The space-time permutation scan statistic is widely used to detect and evaluate disease outbreaks in space and time such as cancer, diarrhea, malaria and many others (Michelozzi et al., 2002, Hsu et al., 2004, Wheeler, 2007, Viel et al., 2000, Kulldorff et al., 2005, Coleman et al., 2009, Jacquez and Greiling, 2003 and Sugumaran et al., 2009). The statistic is based on a scanning cylindrical window with a circular geographic base and with height corresponding to time that moves to regularly scan a geographical region for rapidly rising outbreaks of any location and any size (Kulldorff et al., 2005). This study utilized a space-time permutation scan statistic to detect the early outbreaks of DHF in the study area. The daily notification data from January 1, 2010 to December 31, 2010 was investigated to determine local clustering of DHF cases (the village outbreak identification) by replicating the realistic surveillance processes that arise in Phayao Provincial Public Health Office (PPHO), Thailand.

2. Methods

The methodology is summarized in the flowchart shown in Figure 2.

2.1 Data Acquisition and Preparation

The study of DHF outbreak detection covers the 818 villages of Phayao province. Data on 1,367 laboratory-confirmed DHF cases were obtained from PPHO and included all DHF cases reported between January 1, 2010 and December 31, 2010. Case data contained information on names, gender, age, village code, date of sickness, date reported to PPHO and laboratory date.

2.2 Data Analysis

A space-time permutation scan statistic using SaTScan Software (Kulldorff, 2006) was used to test for the occurrence and location of DHF clusters using time-periodic prospective surveillance and a Poisson probability model (Kulldorff et al., 2005 and Coleman et al., 2009). It was done by modifying the scanning window so that instead of circles across space, cylinders are tested, with time presenting as height of the cylinder. In searching for DHF clusters, thousands or millions of overlapping cylinders were generated to define the scanning window located across the Phayao area. Each cylinder invested a likelihood ratio based on whether the observed number of DHF cases within a cylinder was greater than would be expected if the DHF cases were distributed randomly in space and time with reference to the underlying population. The cylinder with the maximum likelihood was represented as the "most likely cluster" or "DHF outbreak" (Kulldorff, 1997).

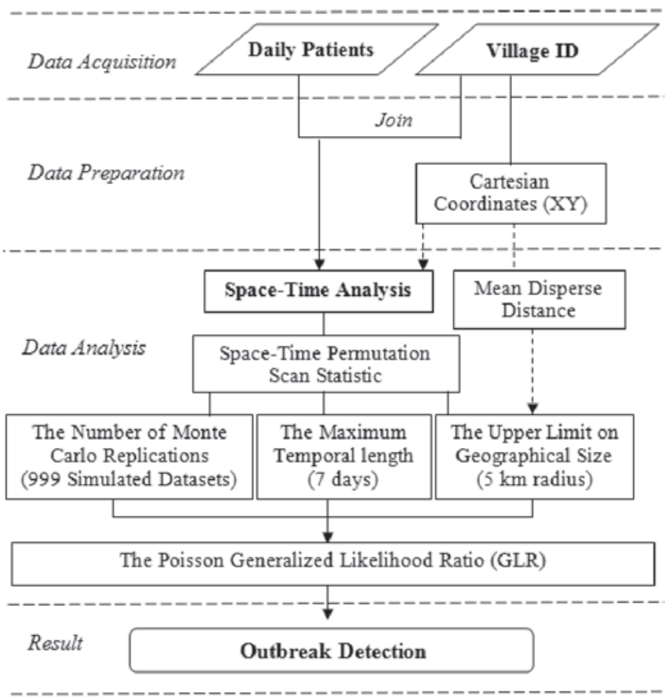

Figure 2: Flowchart of methodology

In this study, daily patient cases of DHF were counted for village-locations, where c_{vd} is the observed number of cases in the village location v and during the day d. C is the total number of observed cases.

$$C = \sum_v \sum_d c_{vd}$$

Equation 1

For each village and day, it can be calculated the expected number of cases in village locations v and day d conditioning on the observed marginal,

$$\mu_{vd} = \frac{1}{C}\left(\sum_v c_{vd}\right)\left(\sum_d c_{vd}\right)$$

Equation 2

The next equation is the expected number of cases μ_{vd} in a particular cylinder. A is the summation of these expectations over all the village-id-days within that cylinder,

$$\mu_A = \sum_{(v,d)\in A} \mu_{vd}$$

Equation 3

The assumption when calculating these expected numbers is that the probability of a case being in a village location v, given that it was observed on day d, is the same for all days d. Let c_A is the observed number of cases in the cylinder. Conditioned on the marginal, and when there is no space-time interaction, c_A is distributed according to the hypergeometric distribution with mean μ_A and probability function,

$$P(c_A) = \frac{\binom{\sum_{v\in A} c_{vd}}{c_A}\binom{C - \sum_{v\in A} c_{vd}}{\sum_{d\in A} c_{vd} - c_A}}{\binom{C}{\sum_{d\in A} c_{vd}}}$$

Equation 4

When both $\sum_{v\in A} c_{vd}$ and $\sum_{d\in A} c_{vd}$ are small compared to C, c_A is approximately Poisson distributed with mean μ_A. Based on this approximation, the Poisson Generalized Likelihood Ratio (GLR) can be measured for an outbreak in each cylinder A.

$$\left(\frac{c_A}{\mu_A}\right)^{cA}\left(\frac{C - c_A}{C - \mu_A}\right)^{(C - c_A)}$$

Equation 5

This is the observed divided by the expected to the power of the observed inside the cylinder, multiplied by the observed divided by the expected to the power of the observed outside the cylinder. Among the many cylinders evaluated, the one with the maximum GLR composes the space-time cluster of cases that is least likely to be a chance occurrence and, hence, is the primary candidate for a true DHF outbreak. Relative risk (RR) for the significant cluster is calculated by dividing the observed number of cases to its expected. RR value > 1 is used to adjust for an increased risk and RR < 1 to adjust for lower risk. RR = 0 is used to adjust for missing data for that particular location (Kulldorff, 2006). For DHF outbreak detection, a space-time prospective analysis is used (Kulldorff, 2001) in order to include all confirmed cases during the day.

In this analysis, the upper limit on the geographical size of the outbreak is set to be a circle with a 5 km radius (mean disperse distance of villages), and the maximum temporal length is set to seven days. To ensure sufficient statistical power, the number of Monte Carlo replications is set to 999, and outbreaks with statistical significance of $p \leq 0.05$ are reported (Kulldorff, 2005).

3. Results

The daily DHF patients in each village from January 1, 2010 to December 31, 2010 were obtained to detect the outbreak signals by using daily prospective analysis in SaTScan software. As a result, outbreak signals with a p-value ≤ 0.05 were listed in Table 1 and presented on the map in Figure 3.

Table 1: DHF outbreaks in 2010 (Weeks 23-34, P-Value ≤ 0.05)

Clusters	Signal (Week)	Signal Date	Number of Days in Signal	Number of Villages in Cluster	Total Number of Cases in Time Step	Observed Cases in the Cluster	Expected Cases	Relative Risk (RR)	*p-value*	Recurrence Interval
Most Likely Cluster	23	8/6/2010	3	12	22	3	0.81	3.68	0.048	Every 1 days
Most Likely Cluster	24	16/6/2010	2	9	24	7	2.31	3.03	0.026	Every 4 days
Most Likely Cluster	25	22/6/2010	3	1	50	5	2.31	3.09	0.019	Every 1 days
Most Likely Cluster	26	30/6/2010	2	6	64	4	0.63	6.40	0.016	Every 6 days
Most Likely Cluster	27 (a)	8/7/2010	1	6	86	3	0.24	12.64	0.012	Every 9 days
Secondary Cluster	27 (b)	5/7/2010	4	17	86	5	1.01	4.93	0.035	Every 3 days
Most Likely Cluster	28 (a)	15/7/2010	1	7	95	2	0.07	28.69	0.015	Every 7 days
Secondary Cluster	28 (b)	15/7/2010	1	18	95	2	0.10	19.13	0.042	Every 2 days
Tertiary Cluster	28 (c)	13/7/2010	3	4	95	4	0.66	6.07	0.045	Every 2 days
Most Likely Cluster	29 (a)	20/7/2010	3	12	109	3	0.22	13.39	0.022	Every 5 days
Secondary Cluster	29 (b)	18/7/2010	5	20	109	11	3.64	3.02	0.029	Every 3 days
Tertiary Cluster	29 (c)	16/7/2010	7	17	109	12	5.20	2.31	0.041	Every 1 days
Most Likely Cluster	30 (a)	28/7/2010	2	27	73	6	1.12	5.34	0.021	Every 5 days
Secondary Cluster	30 (b)	26/7/2010	4	12	73	4	0.57	7.05	0.047	Every 2 days
Most Likely Cluster	31 (a)	30/7/2010	7	15	113	13	4.4	2.96	0.012	Every 8 days
Secondary Cluster	31 (b)	1/8/2010	5	19	113	9	2.51	3.58	0.022	Every 5 days
Most Likely Cluster	32	8/8/2010	5	2	110	4	0.48	8.41	0.033	Every 3 days
Most Likely Cluster	33	18/8/2010	2	14	121	7	1.03	6.81	0.017	Every 59 days
Most Likely Cluster	34	25/8/2010	2	1	122	3	0.14	21.43	0.050	Every 11 days

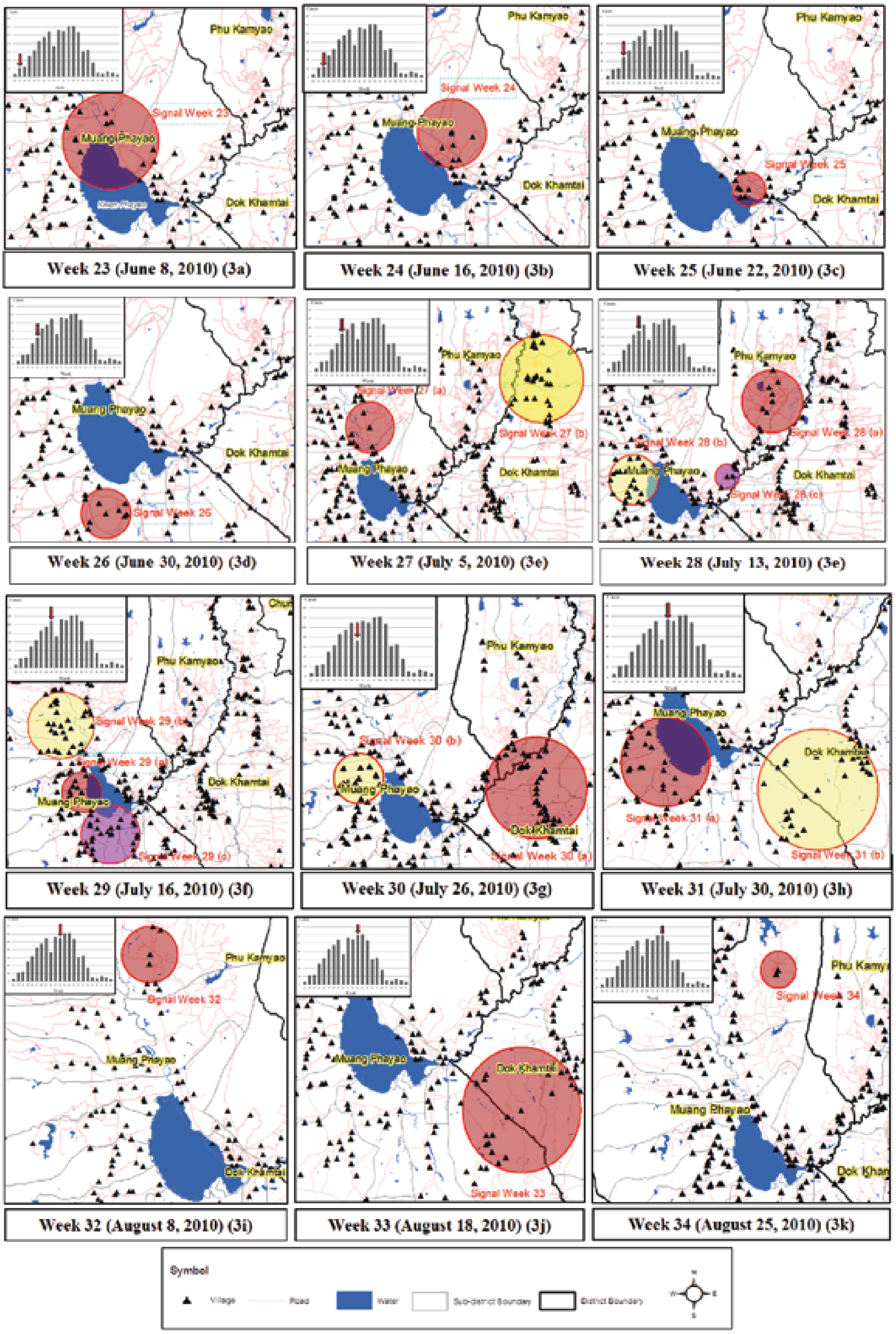

Figure 3: Locations and dates of detected DHF outbreak signals

From the study period, 19 space-time clusters were detected. The first confirmed DHF case was reported to PPHO on January 22, and the first significant signal was detected 145 days later on June 8 (week 23; the beginning of the rainy season). This signal covered 12 villages on the north side of Phayao Lake, Muang Phayao district. It had three cases observed over three days when 0.81 were expected (RR = 3.68). The signal preceded a sharp increase in province wide cases from June 4-10 (Figure 3a). During the early stages of the DHF outbreak (week 23-26) significant space-time clusters occurred around the lake (Figure 3a-3d). On July 5, a new significant cluster was detected in the north of Dok Khamtai district. The outbreak signal extended across 17 villages, and included five cases observed over four days, which was 4.93 times more cases than expected (signal week 27 (b) in Table 1). On week 28, three significant signals were detected (Figure 3e). During this time step there were a total of 95 cases. The most likely cluster occurred in the south of Phu Kamyao district, extended across seven villages, and included two cases observed in one day, which was 28.69 times greater than expected (signal week 28 (a) in Table 1). The secondary cluster occurred on the west side of the lake, extended across 18 villages, and included two cases observed in one day, which was 19.13 times more cases than expected (signal week 28 (b) in Table 1). The tertiary cluster occurred near the eastern edge of Muang Phayao district, extended across four villages, and included four cases observed over three days, which was 6.07 times more cases than expected (signal week 28 (c) in Table 1). Figure 3f shows three significant clusters of DHF outbreaks that were distributed in the north (Ban Mai, Tha Champi, Mae Puem and Ban Tam sub-districts), west (San Pa Muang and Ban Sang sub-districts) and south (Mae Sai, Mae Tam, Mae Na Ruea and Mae Ka sub-districts) of the lake. During this time step there were a total of 109 cases. The most likely cluster extended across 12 villages, and included three cases observed over three days, which was 13.39 times more cases than expected (Signal week 29 (a) in Table 1). The secondary cluster extended across 20 villages and included 11 cases observed over five days, which was 3.02 times greater than expected (signal week 29 (b) in Table 1). The tertiary cluster extended across 17 villages and included 12 cases observed over seven days, which was 5.20 times more cases than expected (signal week 29 (c) in Table 1). On August 18, a large space-time cluster was detected near the border between Muang Phayao and Dok Khamtai districts. This cluster extended across 14 villages and included 121 cases in the time step. The outbreak cluster covered seven cases observed over 14 days when the expected number of cases was 1.03, which produced a relative risk of approximately 6.81 (signal week 33 in Table 1). The last signals occurred on August 25, when the epidemic had diminished (Figure 3k). During this time step there were a total of 122 cases. The small cluster appeared on the north side of Muang Phayao district and included three cases observed over two days, which were 19.13 times more cases than expected (signal week 34 in Table 1).

4. Discussion and Conclusions

The study showed a method for prospective DHF outbreak surveillance that used only case data at the village level. The space-time permutation scan statistic illustrated the disease outbreaks in the grouping of DHF locations across the study area, and strongly confirms the surveillance areas on the point location map. When using DHF 2010 historical data to simulate a surveillance system with daily analysis, it may have been possible to detect the outbreak signals that ware emerging from the DHF cases being reported on a daily basis. The significant outbreak signals with the relative risk greater than ten were detected on week 28-34 (July-August). The outbreak areas occurred around Phayao Lake (the biggest freshwater lake in Muang Phayao district) and found to cover the conurbations in the north of Dok Khamtai district, which were associated with the DHF surveillance announcement from PPHO on July-August 2010 (Bureau of Epidemiology, 2010).

However, there were some limitations in the study:

1) The DHF cases were classified by laboratory date and by village, but not by climate, environment, sanitation and socio-demographic characteristics. Although, these characteristics can be an important determinant for DHF diseases (for example, DHF occurs mainly in children and in the rainy season) (Nakhapakorn and Tripathi, 2005 and Vanwambeke et al., 2006), they were not available in the epidemiological reports and were not included as a determinant in this study.

2) Circle with a 5 km radius (the maximum spatial cluster size parameter) was used to detect the significant clusters, which may or may not accord to the true boundaries of outbreak areas. The impact of parameter settings should be considered to analyze for the next study.

This study reveals that space-time statistical methods can be useful for DHF surveillance for public health officials and other healthcare workers. It represents the coordinating of daily data and spatial analysis with GIS to identify the geographical area for outbreaks in any location and any size during periods of outbreaks. The use of this analysis and tools should be beneficial for developing an early DHF outbreak detection system of PPHO in the future.

Acknowledgements
This study was supported by grants funded by the University of Phayao. The authors thank Phayao Provincial Public Health Office (PPHO) and Geo-Informatics and Space Technology Centre (Northern Region) for providing data.

References

Bureau of Epidemiology, 2010, Annual Epidemiological Surveillance Report 2010, Phayao Province. Bangkok: Department of Disease Control, Ministry of Public Health, Thailand.

Coleman, M., Mabuza, A. M., Kok, G., Coetzee, M. and Durrheim, D. N., 2009, Using the SaTScan Method to Detect Local Malaria Clusters for Guiding Malaria Control Programmes. *Malaria Journal*, 8, 68.

Hsu, C, E., Jacobson, H. and Mas, F. S., 2004, Evaluating the Disparity of Female Breast Cancer Mortality among Racial Groups - A Spatiotemporal Analysis. *International Journal of Health Geographics*, 3, 4.

Jacquez, G. M. and Greiling, D. A., 2003, Local Clustering in Breast, Lung and Colorectal Cancer in Long Island, New York. *International Journal of Health Geographics*, 2, 3.

Kulldorff, M., 1997, A spatial scan statistic. *Communications in Statistics: Theory and Methods*, 26(6), 1481-1496.

Kulldorff, M., 2001, Prospective Time-Periodic Geographical Disease Surveillance using a Scan Statistic. *Journal of the Royal. Statistical Society*, A(164), 61-72.

Kulldorff, M., Heffernan, R., Hartman, J., Assuncao, R. and Mostashari, F., 2005, A Space-Time Permutation Scan Statistic for Disease Outbreak Detection. *Public Library of Science Medicine*, 2, 216-224.

Kulldorff, M., 2006, SaTScanTM User Guide for version 7.0. Available: http://satscan.org/.

Michelozzi, P., Capon, A., Kirchmayer, U., Forastiere, F., Biggeri, A., Barca, A. and Perucci, C. A., 2002, Adult and Childhood Leukemia near a High-Power Radio Station in Rome, Italy. *American Journal of Epidemiology*, 155, 1096-1103.

Nakhapakorn, K. and Tripathi, N. K., 2005, An Information Value Based Analysis of Physical and Climatic Factors Affecting Dengue Fever and Dengue Haemorrhagic Fever Incidence. *International Journal of Health Geographics*, 4, 13.

Phayao Provincial Public Health Office, PROVIS datacenter. Available: http:// 61.7.231.104-/provis/main/index.php

Sugumaran, R., Larson, S. R. and DeGroote, J. P., 2009, Spatio-Temporal Cluster Analysis of County-Based Human West Nile Virus Incidence in the Continental United States. *International Journal of Health Geographics*, 8, 43.

Vanwambeke, S. O., Benthem, B. H. V., Khantikul, N., Maas, C. B., Panart, K., Oskam, L., Lambin, E. F. and Somboon, P., 2006, Multi-level Analyses of Spatial and Temporal Determinants For Dengue Infection. *International Journal of Health Geographics*, 5, 5.

Viel, J. F., Arveux, P., Baverel, J. and Cahn, J. Y., 2000, Soft-Tissue Sarcoma and Non-Hodgkin's Lymphoma Clusters around a Municipal Solid Waste Incinerator with High Dioxin Emission Levels. *American Journal of Epidemiology*, 152, 13-19.

Wheeler, D. C., 2007, A Comparison Of Spatial Clustering And Cluster Detection Techniques For Childhood Leukemia Incidence in Ohio, 1996 – 2003. *International Journal of Health Geographics*, 6, 13.

World Health Organization, 2006, Dengue Haemorrhagic Fever: Early Recognition, Diagnosis and Hospital Management. Available: http://www.who.int/csr/don/archive/disease/dengue_fever/dengue. pdf.

World Health Organization, 2009, Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. Available: http://whqlibdoc.who.int/publications/2009/9789241547871_eng.pdf.