Spatial Diffusion of Diarrhea Outbreak in Chiang Mai, Thailand

Chaikaew, N.,[1] and Tripathi, N. K.,[2]
[1]Geographic Information Science field of study, School of Information and Communication Technology University of Phayao, Phayao, Thailand, E-mail: nakarin.ch@up.ac.th
[2]Remote Sensing and GIS field of study, Asian Institute of Technology, Pathumthani, Thailand E-mail: nitinkt@ait.ac.th

Abstract

The objective of this study was to utilize the space-time permutation scan statistic and GIS for investigating and visualizing the spatial diffusion of diarrhea outbreaks in Chiang Mai province, Thailand. A space-time retrospective analysis was conducted to detect the outbreak signals of diarrhea from 1 November 2003 to 31 October 2006, in order to include all confirmed diarrhea cases of patient less than five years of age. A total of 37,536 diarrhea cases were reported in Chiang Mai from November 1, 2003 to October 31, 2006. In the study period had more village outbreaks and the diarrhea incidence was so high. The strongest signals of outbreak (13 villages) was on 17 January 2006 (cool season) and occurred to the conurbations of Mae Chaem district (Chang Khoeng and Tha Pha sub-districts), which is located about 100 km to the south-west of Chiang Mai City. This signal had 44 cases observed over 8 days when 2.98 were expected (relative risk= 14.79). The weaker signal was on 29 June 2005 (begin of rainy season) that were located in southern Chiang Mai (Chom Thong district), there had 16 cases observed over 11 days when 1.21 were expected theoretically (relative risk= 13.24). The daily spatial spread or diffusion of diarrhea outbreak from the first day to the end day of incidence (17-24 January 2006) in the study site of Mae Chaem district was follow a pattern of contagious diffusion. They spread outward from a village of origin (Ban Rai) with high incidence to nearby villages, which related to the area of high population densities, within the urban and agricultural areas along the main course of Mae Chaem River.

1. Introduction

In Thailand, diarrhea has been a major public health problem for many years. The Bureau of Epidemiology, Ministry of Public Health, estimated nearly 1 million cases every year (in the period 2001-2005: 1,020,377, 1,055,393, 966,760, 1,161,877 and 1,142,581 respectively, with corresponding deaths: 176, 160, 124, 93 and 77). In 2006, the diarrhea incidence was estimated to be 1,245,022 cases and 9 deaths, with the highest incidences occurring in Chiang Mai, Chiang Rai, Khon Kaen and Roi Et provinces, all in the northern and north-eastern region of Thailand. Diarrhea is most commonly caused by gastrointestinal infections (bacterial, viral and parasitic organisms). The infection is spread through contaminated food or drinking-water, or from person to person as a result of poor hygiene (WHO, 2005). The disease incidence in children under five years of age was also reported high (FAO, 2004). Diarrhea outbreak is defined as the occurrence of a large number of diarrhea cases in a restricted geographical area over a short period of time (FAO, 2004 and Pande et al., 2008). For disease outbreak detection, the space-time scan statistics (i.e. space-time permutation scan statistic), was developed by Martin Kulldorff, is widely used to detect and evaluate disease outbreaks in space and time such as cancer (Michelozzi et al., 2002 and Viel et al., 2000), diarrhea (Kulldorff et al., 2005), malaria (Coleman et al., 2009) and many others (Kulldorff, 2003). The basic of these statistics is defined by a scanning cylindrical window with a circular geographic base and with height corresponding to time that moves across space and time. For each window the numbers of diseases cases inside and outside the cylinder are noted, together with the expected number of cases reflecting the population at risk and relevant covariates. On the basis of these numbers, the likelihood is calculated for each cylinder. The cylinder with the maximum likelihood, and with more than its expected number of cases, is denotes the "most likely cluster" or "disease outbreak" (Kulldorff, 1997). Chiang Mai province is the endemic area of diarrhea in the north of Thailand. The epidemic pattern of diarrhea in this area has fluctuated every year from 2001 to 2006. The highest number of cases was recorded in 2004, particularly in Doi Tao, Samoeng, and Hot districts,

ISSN 1686-6576 / © Geoinformatics International

with incidence rates of 6,345, 4,905 and 4,493 per 100,000 people, respectively. The incidence of diarrhea marks high variability at district, sub-district and village levels. Socio-demographic, environmental and sanitation, and climate factors are thought to be related to the outbreak of diarrhea. The burden of diarrhea is well described in this area as mandatory reporting of the patient cases, which is recorded from the Chiang Mai Provincial Public Health Office (CMPHO). All confirmed diarrhea cases are entered into a computerized public health surveillance system of CMPHO. The attempt to apply the spatial statistical methods and GIS for understanding and identifying the disease outbreaks in Thailand has become the important tool in epidemiological studies (Ruankaew, 2005). This chapter describes the space–time analysis to detect diarrhea outbreaks when available signal (i.e. number of cases) for a particular area is being monitored. The daily notification data from November 1, 2003 to October 30, 2006 of Chiang Mai was analyzed to determine local clustering of cases (the village outbreak identification) by using space-time permutation scan statistic (Kuldorff et al., 2004). Moreover, the spatial prediction method of inverse distance weighed (IDW) was then used to produce the surface maps, which highlight the risk of diarrhea at certain places in the outbreak area to dye in each day from the first day to the end day of outbreak, for visualizing the spatial spread or diffusion of diarrhea outbreaks.

2. Methodology

Diarrhea outbreak detection and determination of spatial diffusion are summarized in the flowchart shown in Figure 1.

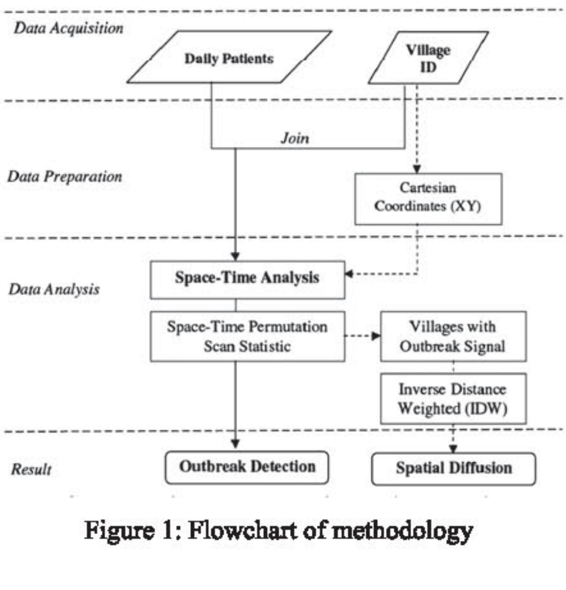

Figure 1: Flowchart of methodology

2.1 Data Acquisition and Preparation

The daily diarrhea patients less than five years of age at village level were obtained from the Chiang Mai Provincial Public Health Office (CMPHO) during November 1, 2003 to October 31, 2006 (Table 1). The spatial data related to the village location were collected from The Geo-Informatics and Space Technology Centre (Northern Region) (GISTC), Thailand. All these data were incorporated into a geographic information system (GIS).

Table 1: Summary of the data set

Age	Locations	Period	Total cases
< 5yr	2005	Nov 2003 - Oct 2006	37,536

2.2 Data Analysis

2.2.1 Space-time permutation scan statistic

Space-time permutation scan statistic (Kulldorff, 1997) is used to identify diarrhea outbreaks by considering both spatial and temporal aspects of the signal. This is done by modifying the scanning window so that instead of circles across space, cylinders are tested, with time presenting as height of the cylinder. Thousands or millions of overlapping cylinders are utilized to define the scanning window, each being a possible candidate for disease outbreak. The circular base represents the graphical area of the potential outbreak. The spatial statistic is implemented in the SaTScan software (Kulldorff, 2003). For diarrhea outbreak detection, a space-time retrospective analysis is used, 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 5.2 km radius (mean disperse distance of villages), and the maximum temporal length is set to 14 days. To ensure sufficient statistical power, the number of Monte Carlo replications is set to 999, and outbreaks with statistical significance of $p \leq 0.001$ are reported.

2.3 Inverse Distance Weighted (IDW)

Inverse distance weighted is a technique of interpolation that estimates grid cell values by averaging the values of sample data points in the neighborhood of each processing cell. This method is based on the assumption that the interpolating surface should be powered most by the nearby points and less by the more distant points. The closer a point is to the center of the cell being calculated, the more power, it has in the averaging process (Anonym, 1999).

3. Results

3.1 Diarrhea Outbreak Detection

The daily diarrhea patients less than five years of age in each village from 1 November 2003 to 31 October 2006 were obtained to detect the outbreak signals by using daily retrospective analysis in SaTScan software. As the result, outbreak signals with p-value ≤ 0.001 are listed in Table 2 and presented on the map in Figure 2. The strongest signal (A) was on 17 January 2006 (cool season), covering 13 villages in Mae Chaem district. This signal had 44 cases observed over 8 days when 2.98 were expected (relative risk= 14.79). The signal immediately preceded a sharp increase in provincewide cases from 18-23 January (Figure 3). In both the localized 17 January cluster and the provincewide outbreak, the increase was most notable among children less than two years of age. The weaker signal (K) was on 29 June 2005 (begin of rainy season) that were located in southern Chiang Mai (Chom Thong district), there had 16 cases observed over 11 days when 1.21 were expected theoretically (relative risk= 13.24). While top two outbreak signals (A and B) were found in cool season month (January), almost signals (C to K) were occurred in hot and rainy seasons (March-October). The large area of outbreak (H) covered 58 villages in San Kamphaeng and Doi Saket districts. There had 25 cases observed over 2 weeks (24 September to 7 July 2006), when 2.89 were expected (relative risk= 8.64). The small area of outbreak (G) was located in Mae Chaem district during 3-6 August 2005. They found only one village that had 10 observed cases over 4 days when 0.15 were expected (relative risk= 68.41).

3.2 Spatial Diffusion Analysis

3.2.1 Study site

The study site, which is selected to analyze the spatial spread or diffusion of diarrhea, is the area with strongest signal (A) of diarrhea outbreak of Chiang Mai province. It covers 13 villages (Table 3) that locate to the conurbations of Mae Chaem district and about 100 km to the south-west of Chiang Mai City (Figure 4). These outbreak villages occur on the alluvial plain of the Mae Chaem River and its tributaries, which are clustered in Chang Khoeng and Tha Pha sub-districts. 13 villages over 8 days (17-24 January 2006) are considered for analyzing the spatial spread or diffusion of diarrhea outbreak. The related factors (i.e. population density, elevation, land cover and stream network etc.) are overlaid with the outbreak villages to know the relationship between these factors and the outbreak areas.

3.3 The Distribution of Diarrhea Outbreak

The distribution of villages with diarrhea outbreak over 8 days in the study site is illustrated by overlaying the topographic, population density and land cover maps (Figures 5, 6 and 7, respectively). Almost infected villages occurred within agricultural, urban and built-up areas (approximate elevation as 500 meters above mean sea level), which had high population densities, along the Mae Chaem River. Villages with diarrhea outbreak were located within densely populated areas, more than 1,000 persons per 1 sq. km, and nearby main river of the study site. Over 70% of infected villages distributed within 1 km from the Mae Chaem River while 70% of non-infected villages occurred more than 1 km from the river.

Table 2: Diarrhea outbreaks from November 1, 2003 to October 31, 2006 (p-value ≤ 0.001)

Signal ID	Outbreak signal date	No. Days in signal	No. Villages in signal	Sub-district	District	Observed cases	Expected cases	Relative risk	P-value
A	Jan 17, 2006	8	13	- Chang Khoeng - Tha Pha	Mae Chaem	44	2.98	14.79	0.001
B	Jan 16, 2006	14	2	- Mae Suek	Mae Chaem	44	5.46	8.05	0.001
C	April 29, 2006	12	3	- Sop Khong	Om Koi	26	1.77	14.67	0.001
D	Jun 10, 2004	8	4	- Hot - Na Kho Ruea	Hot	18	0.83	21.59	0.001
E	May 28, 2006	3	4	- Ban Chan	Mae Chaem	10	0.10	99.30	0.001
F	Sep 17, 2005	13	10	- Thung Satok - Ban Mae - Thung Ruang Thong - Ban Kat	San Pa Tong Mae Wang	21	1.61	13.03	0.001
G	Aug 3, 2005	4	1	- Mae Na Chon	Mae Chaem	10	0.15	68.41	0.001
H	Sep 24, 2006	14	58	- Buak Khang - Chae Chang - Mae Pu Kha - Rong Wua Daeng - Sai Mun - San Kamphaeng - Ton Pao - Mae Khue	San Kamphaeng Doi Saket	25	2.89	8.64	0.001
I	Mar 13, 2005	8	1	- Mae Soi	Chom Thong	17	1.11	15.36	0.001
J	Mar 1, 2005	14	39	- Mae Kha - Mae Sun - Mon Pin - San Sai	Fang	71	24.81	2.86	0.001
K	Jun 29, 2005	11	3	- Doi Kaeo	Chom Thong	16	1.21	13.24	0.001

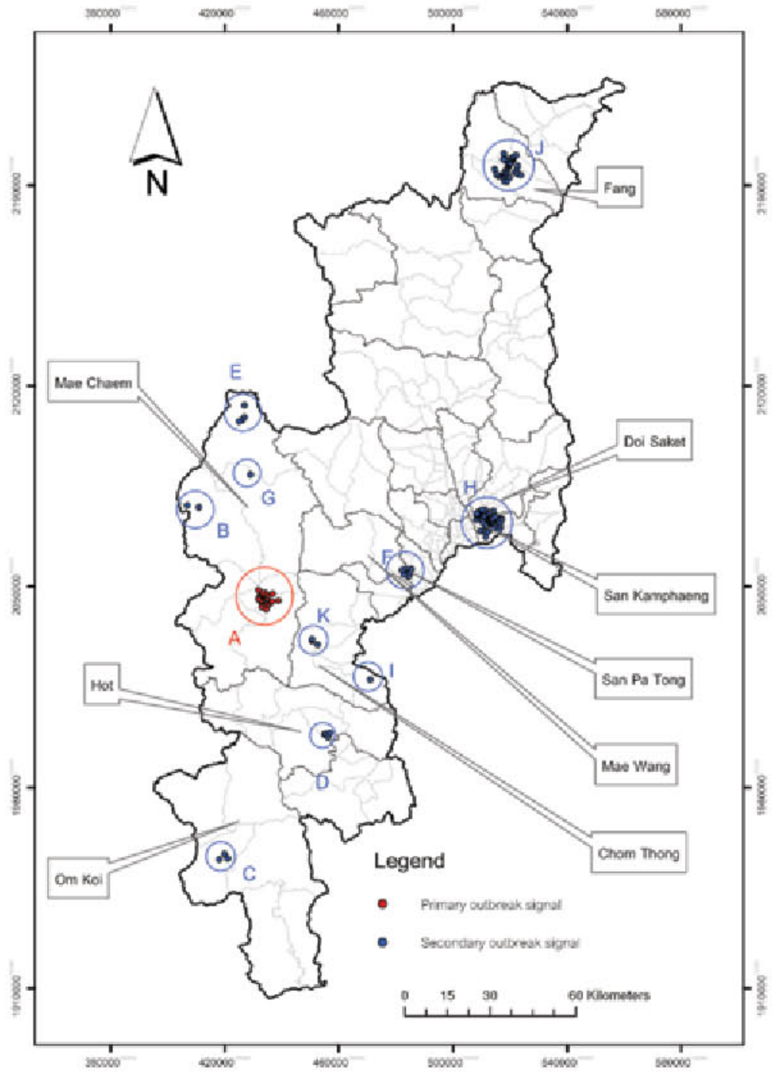

Figure 2: Location of detected diarrhea outbreak signal, using historical data from 1 November 2003 to 31 October 2006

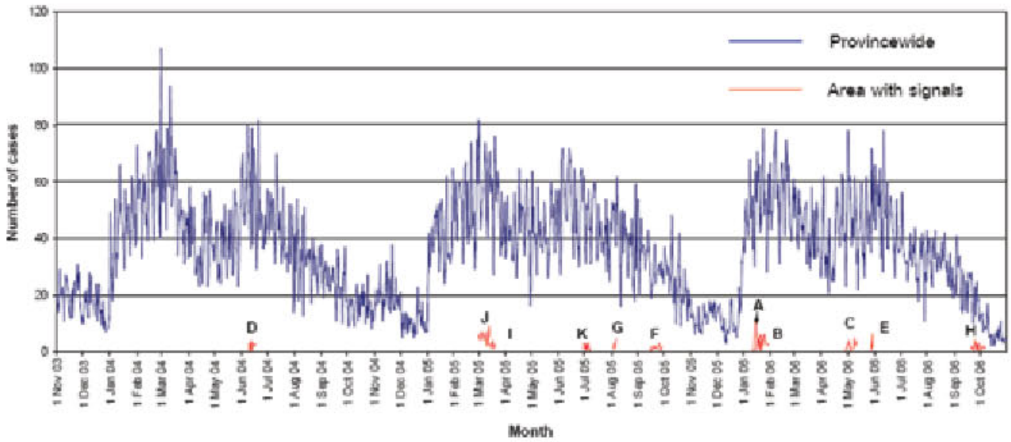

Figure 3: The daily temporal pattern of diarrhea incidence cases in Chiang Mai province, 1 November 2003 to 31 October 2006

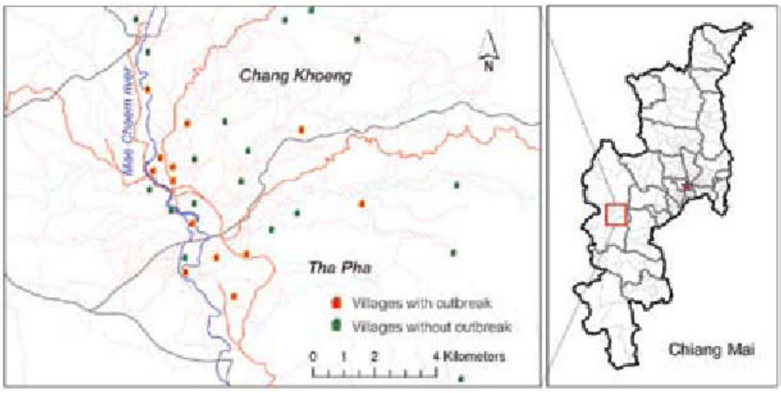

Figure 4: The study site of Mae Chaem district

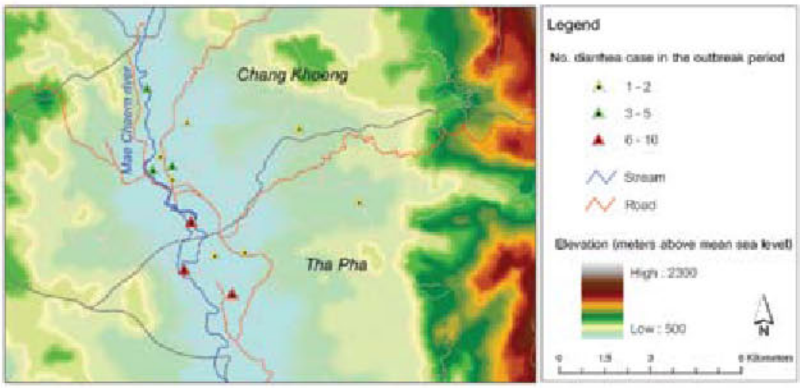

Figure 5: Topography over villages with diarrhea outbreak

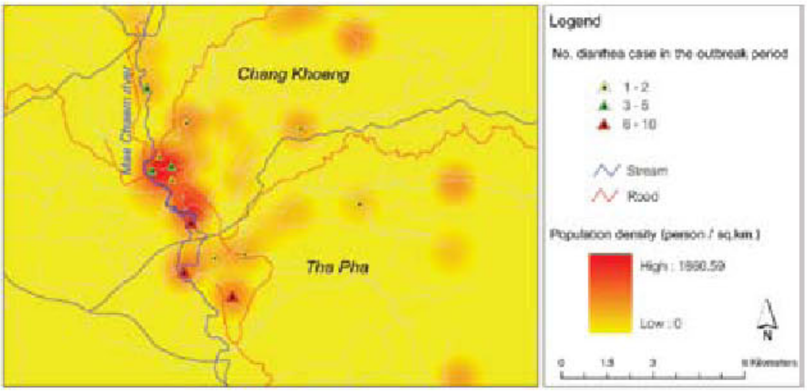

Figure 6: Population densities over villages with diarrhea outbreak

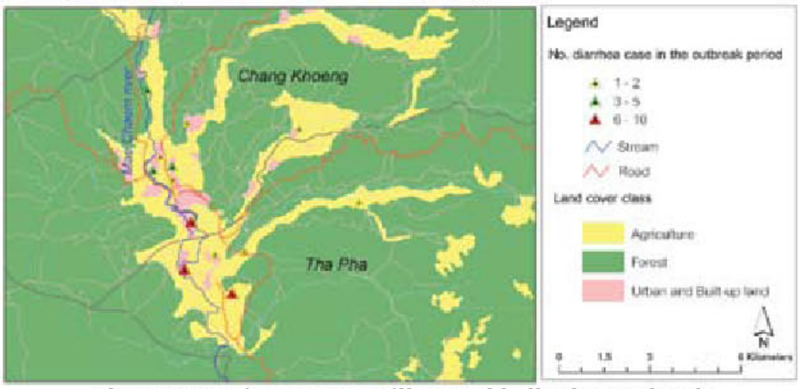

Figure 7: Land cover over villages with diarrhea outbreak

Table 3: The numbers of village with diarrhea outbreak in the study site

No. of village	Village with diarrhea outbreak	Sub-district	District
13	- Thong Fai - San Nong - Chiang - To Ruea - Mae Pan - Phrao Num - Mae King (Bon Na) - Ko	Chang Khoeng	Mae Chaem
	- Pa Daet - Huai Hai - Rai - Yang Luang - Pha Nang	Tha Pha	

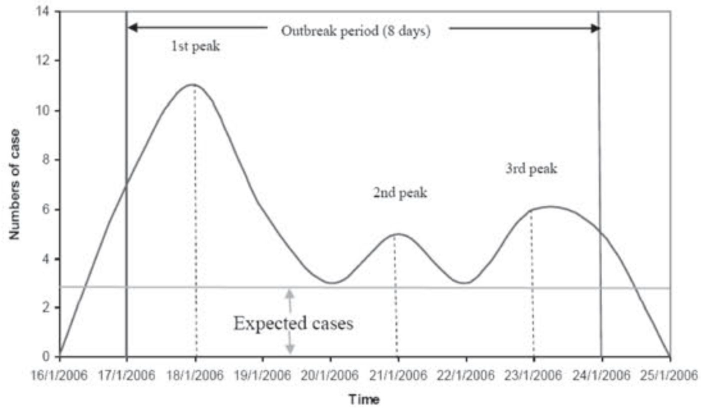

Figure 8: The numbers of diarrhea case over 8 days of outbreak in the study site

The land cover in 2006 around the villages with diarrhea outbreak and villages without diarrhea outbreak is obtained for 250 m, 500 m and 1,000 m buffer zone. The percentages of agricultural, urban and built-up areas around the villages with outbreak are higher than the villages without outbreak in the all of buffer zone. On the other hand, the villages without outbreak are distributed around the forest area of the study site. The percentages of forest areas around the non-infected villages are higher than the infected villages in the all of buffer zone. Especially, the percentage of forest areas increased over 60 % within the 1,000 m buffer zone from the villages without diarrhea outbreak.

3.4 Spatial diffusion of diarrhea outbreak

The spatial spread or diffusion pattern is analyzed by daily from the first day to the end day of outbreak (17-24 January 2006). Figure 8 represents the epidemics of diarrhea in the study site that started quickly, the number of infected increased and first peaked over than the expected cases (2.98) of outbreak period (see Appendix) in 2nd day from the first day of outbreak. In the outbreak period, the numbers of diarrhea case peaked at the 2nd, 5th and 7th day (18, 21 and 23 January, respectively) then decreased to 0 in 9th day (25 January 2006) from the first day of incidence. Figure 9 shows the movement of diarrhea through space and time. Spatial diffusion of diarrhea in the study site occurs when a disease is transmitted to new villages within the radius of 2.5 km (mean nearest neighbor distance of villages in the study site). It follows a pattern of contagious diffusion, which is spread outward from a village of origin to nearby villages (Cliff et al., 1981). From the daily diffusion analysis over 8 days of outbreak, the first outbreak of diarrhea (17 January 2006) occurs in the villages along the Mae Chaem River (Rai, Ko and Thong Fai) then transmits to neighbor villages within 2.5 km from these infected villages. It is clearly understood that there are strong spatial spread in 2nd, 3rd, 4th, 6th and 7th day from the first day of incidence within the areas of high population density and around the main river.

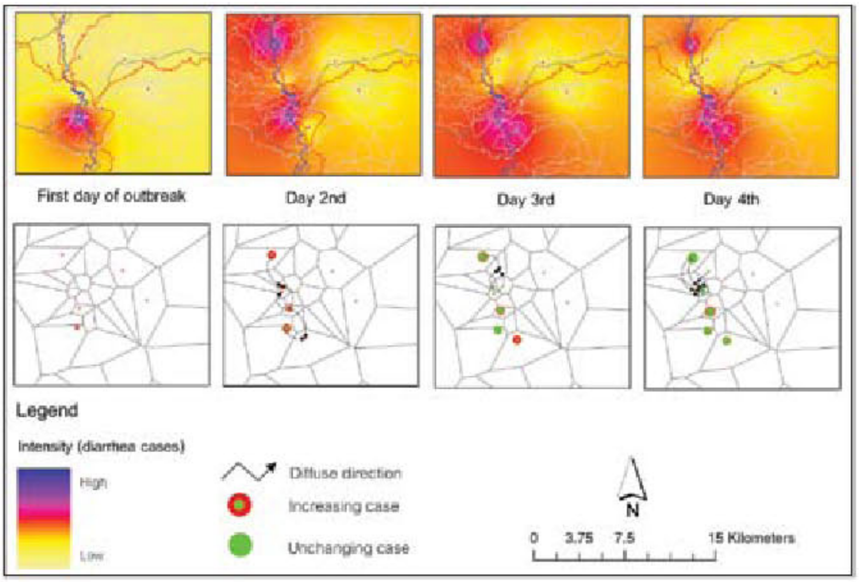

Figure 9: Spatial diffusion of diarrhea outbreak from 17 to 24 January 2006

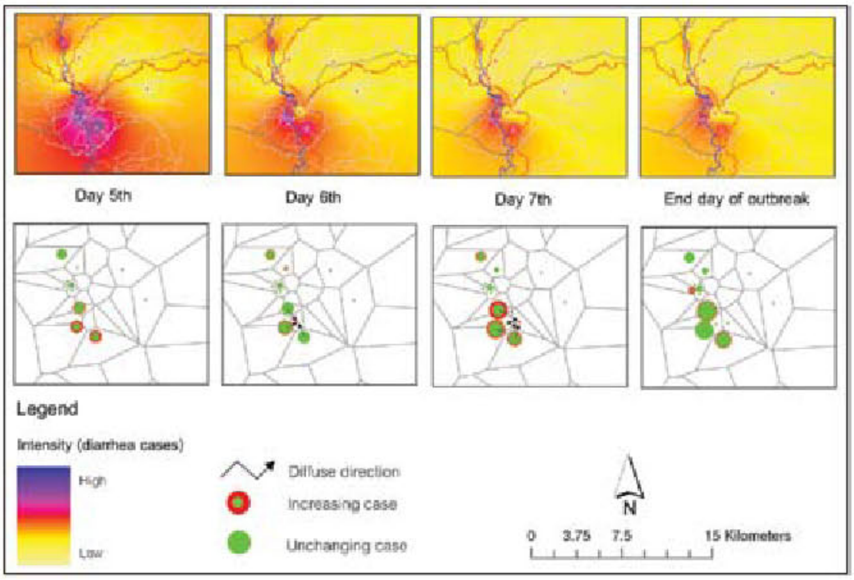

Figure 10: Spatial diffusion of diarrhea outbreak from 17 to 24 January 2006 (continue)

In the end day of outbreak (24 January 2006), the cumulative incidence of diarrhea are highly increased in the downstream area of the study site, it is clearly understandable that the villages in this area get diarrhea infection easily. As the result, it can be concluded that the diarrhea infection may be transmitted through water from upstream villages to downstream villages in the cool season.

4. Conclusion

Space-time permutation scan statistics was successfully used to detect diarrhea outbreaks in Chiang Mai province when the numbers of diarrhea patients less than five years of age are highly increased over a short period of time. The daily analysis found that the strongest signal of diarrhea outbreak occurred within agricultural, urban and built-up areas, which had high population densities, along the main river of Mae Chaem district on cool season (January 2006). The signal covered 13 villages, which located in 2 sub-districts of Mae Chaem district, over 8 days of outbreak (17-24 January 2006). The spatial prediction method of IDW can be illustrated clear spatial spread or diffusion of these outbreak signals that spread in the areas of high population density and diffused from upstream to downstream villages in a short time. This study reveals that the spatial methods can be useful for diarrhea surveillance for public health officials and other healthcare workers. It represented that cooperated the daily data and spatial analysis with GIS to identify the geographical area for outbreaks in any location and any size and investigate the spatial diffusion of diarrhea during the period of outbreak. The use of these analysis and tools should be beneficial for developing the early diarrhea outbreak detection system of CMPHO in the future.

Acknowledgements

Special thanks to the Chiang Mai Provincial Public Health Office and Geo-Informatics and Space Technology Centre (Northern Region) for providing data.

References

Anonym, 1999, Golden Software, Surfer 8. *User's Guide: Contouring and 3D Surface Mapping for Scientist and Engineers*. Colorado, USA

Chiang Mai Provincial Public Health Office, 2006. *Epidemiology Surveillance Report* [online]. Available from: http://61.19.145.137/cmpho_web/epe_web/ [Accessed 2 November 2006].

Cliff, A. D., Haggett, P., Ord, J. K., and Versey, G., 1981, *Spatial Diffusion: An Historical Geography of Epidemics in an Island Community*. Cambridge, United Kingdom: Cambridge University Press.

Coleman, M., 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*(1), 68.

Ellen, K. C., and Sara, L. M., 2002, Analyzing the Risk and Spread of Infectious Disease. *GIS and Public Health*, New York: Guilford Press.

Food and Agriculture Organization of the United Nations, 2004, *Foodborne Diseases: Situation of Diarrheal Diseases in Thailand*. In FAO/WHO Regional Conference on Food Safety for Asia and the Pacific, FAO.

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., 2003, Methods Comparisons. Re: 'Local Clustering in Breast, Lung and Colorectal Cancer in Long Island, New York', by Jacquez GM, Greiling DA [commentary]. *International Journal of Health Geographics, 2*(2003), 3.

Kulldorff, M,, Zhang, Z., Hartman, J., Heffernan, R., Huang, L., and Mostashari, F., 2004, Benchmark data and power calculations for evaluating disease outbreak detection methods. *Morbidity and Mortality Weekly Report, 53*(2004), 144-151.

Kulldorff, M., Heffernan, R., Hartman, J., Assuncao, R. and Mostashari, F. (2005). Aspace-Time Permutation Scan Statistic for Disease Outbreak Detection. *Public Library of Science Medicine, 2*(3), 216-224.

Kuo, C. L., 2005, *GIS and Statistical Analyses for the Spatial Diffusion of Cholera in Modern Japan: The Case Study of Fukushima Prefecture in Epidemic 1882 and 1895*. Global Security Research Institute (G-SEC), Keio University, Japan.

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*(12), 1096-1103.

Pyle, G. F., 1969, The Diffusion of Cholera in the United States in the Nineteenth Century. *Geographical Analysis, 1*(1969), 59-79.

Ruankaew, N., 2005, GIS and Epidemiology. *Journal of the Medical Association of Thailand, 88*(11), 1735.

Stock, R. F., 1976, Cholera in Africa-Diffusion of the Disease 1970–1975 with Particular Emphasis on West Africa. *African Environment Special Report 3*. London, International African Institute.

United States National Institutes of Health, 2007, *Fecal Calprotectin: Cheap Marker for Diagnosing Acute Infectious Diarrhea* [online]. Available from: http://clinicaltrials.gov/ct2/show/ NCT00429325 [Accessed 21 January 2007].

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*(1), 13-19.

World Health Organization, 2005, *The treatment of diarrhea* [online]. Available from: http://whqlibdoc.who.int/publications/2005/9241593180.pdf [Accessed 21 January 2007].