SPATIAL ANALYSIS OF DENGUE INCIDENCE IN TAIWAN

by

Jia-Yuh Chen

B.S. in Environmental Engineering, National Cheng Kung University, Taiwan, 2003

M.S. in Environmental Health, National Taiwan University, Taiwan, 2005

Submitted to the Graduate Faculty of

Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Jia-Yuh Chen

It was defended on

June 9, 2009

and approved by

Thesis Advisor:

Sati Mazumdar, PhD, Professor, Department of Biostatistics

Graduate School of Public Health, University of Pittsburgh

Committee Member:

Ravi K. Sharma, PhD, Assistant Professor,

Department of Behavior and Community Health Sciences

Graduate School of Public Health, University of Pittsburgh

Stewart Anderson, PhD, Professor, Department of Biostatistics

Graduate School of Public Health, University of Pittsburgh

Copyright © by Jia-Yuh Chen

2009

Sati Mazumdar, PhD.

SPATIAL ANALYSIS OF DENGUE INCIDENCE IN TAIWAN

Jia-Yuh Chen, MS

University of Pittsburgh, 2009

ABSTRACT

Dengue is an important mosquito-borne viral disease. In Taiwan, there are hundreds to thousands dengue cases each year, and dengue is considered one of the most important public health issues. The objective of this study was to use geographical information systems (GIS) methodology to map and analyze the spatial and temporal distribution of dengue in Taiwan during 2004 to 2007 and to elucidate the association of geographical and climatic risk factors with dengue incidence.

Dengue annually occurs starting in summer, peaking in fall and goes down in winter. The spatial distribution: Spatial autocorrelation of dengue was measured using Moran's I at the global level and LISA at the local level. The global spatial autocorrelation analysis revealed a significantly positive spatial autocorrelation of dengue for 2004 to 2007, with Moran's I=0.171, *p*-value=0.03. The local spatial autocorrelation analysis showed a significantly high dengue incidence around Tainan county and Kaohsiung county (*p*-value<0.05), which are located in the southern Taiwan. Based on the geographical features, dengue tended to occur in the southwestern cities/counties in Taiwan with plains and rivers spread. Temperature had a positive relationship with dengue incidence in summer and fall (r_s=0.74 and *p*-value=0.002 in summer, r_s=0.53 and *p*-value=0.003 in fall). Rainfall had a positive relationship with dengue incidence in summer (r_s=0.61 and *p*-value=0.017). However, there was no significant correlation between temperature or rainfall and dengue incidence in winter.

The public health importance of this study: Disease maps have been playing a key descriptive role in public health and epidemiology. By this study, areas of the current geographical distribution of the incidence of dengue in Taiwan were identified. Through spatial autocorrelation analyses, the identification of unusual concentration of dengue in Tainan county and Kaohsiung county has been defined. This could prompt health agencies and the government to take a critical look at these risk areas, and make appropriate health planning and resource allocation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 DENGUE

Dengue is an important mosquito-borne viral disease, affecting more than 1.5 billions people worldwide. It occurs in tropical and subtropical regions of the world. In Taiwan, there are hundreds to thousands dengue cases each year, and dengue is considered one of the most important public health issues. There are four Dengue virus serotypes, called DEN-1, DEN-2, DEN-3, and DEN-4. Dengue virus is transmitted to human by mosquitoes of *aedes* genus. *Aedes aegypti* and *aedes albopictus* mosquitoes are two principal vectors (transmitters) of dengue viruses in Taiwan. The former species is distributed in the south of Taiwan, whereas the latter is found throughout the island.

Many factors have been associated with dengue transmission. Marked spatial and seasonal diversity in dengue incidence reflects the influence of climate and geography on the dengue transmission. Increase in temperature and precipitation can accelerate the development rate of mosquitoes and then lead to increased mosquito breeding, which might result in an outbreak of dengue epidemic [1]. Taiwan is an island, which is located in the Western Pacific between Japan and the Philippines and lies 120 kilometers off the southeast coast of China. Crossed by the Tropic of Cancer, Taiwan has a subtropical climate and a tropical climate in its southern tip. Taiwan's high average annual temperature (22˚C) and precipitation (2,500mm in plains and basins, and 3,000mm in mountainous area) construct a perfect environment for mosquito breeding.

1.2 GEOGRAPHIC INFORMATION SYSTEMS [2, 3]

Geographic information systems (GIS) is a computer system that combines information with the geographic map in order to capture, search, check, integrate, manipulate, analyze and display these information related to positions on the Earth's surface. In the field of health, GIS is increasingly used to analyze the geography of disease, specifically the relationships between pathological factors (agents, vectors, and hosts) and their geographical environments. For example, GIS have been extensively used in describing the geographical distributions of disease agents, identifying regions in time and space where people may be exposed to disease agents, and mapping spatial and temporal patterns of diseases. Furthermore, contrast to mapping, which only provides a visual display of a disease, spatial statistics is widely used to find the relationship between cases and its geographical location and verify clustering of cases or spatial correlations. Because GIS makes it possible to map environmental factors associated with disease vectors, it is especially appropriate for the infectious and vector-borne diseases investigation, such as malaria or Lyme disease [4].

Many studies, some in which GIS was used, have been conducted to identify the mechanisms of dengue transmission in Taiwan. However, most of these studies were limited in certain cities or counties in southern Taiwan. Ciu.[5] investigated the relationship between dengue and climate change in Taiwan. It was found that the prevalence of dengue in Kaohsiung City seemed to be influenced by temperature and precipitation in the previous month. However, the accurate relationship between dengue infection and climate change was not found and this study was limited to a small area. In a study by using GIS, Wen *et al*.[6] identified spatial risk areas in Kaohsiung City through temporal characteristics of epidemic dynamic process, including frequency, duration and intensity. This study was limited to a small area as well.

The specific objective of the project work is to:

(1) map and analyze the spatial and temporal distribution of dengue incidence in Taiwan by using dengue case-incidence data, including onset dates of confirmed dengue cases and their geographic locations by city/county; and

(2) investigate the association between dengue incidence and its risk factors with geo-information, including geographical factors (geographical location, geomorphology, and hydrology), and climatic factors (temperature and rainfall)

2.0 MATERIALS AND METHODS

2.1 MATERIALS

A city or county, the biggest administrative unit in Taiwan, was used as the spatial mapping unit. The study involved 7 cities and 15 counties (Figure 1). The Center for Disease Control- Taiwan (CDC-Taiwan) summarized dengue cases on a weekly basis. Figure 2 illustrates the temporal progression of dengue cases for March 2004 to February 2008, and it is seen that most cases were concentrated in certain months (e.g., September, October, and November). In order to simplify the temporal mapping of a four-year data, we divided one year into four seasons: spring (March to May), summer (June to August), fall (September to November), and winter (December to February in the next year). Therefore, the season was used as temporal unit for better comparison on different dengue risk factors.

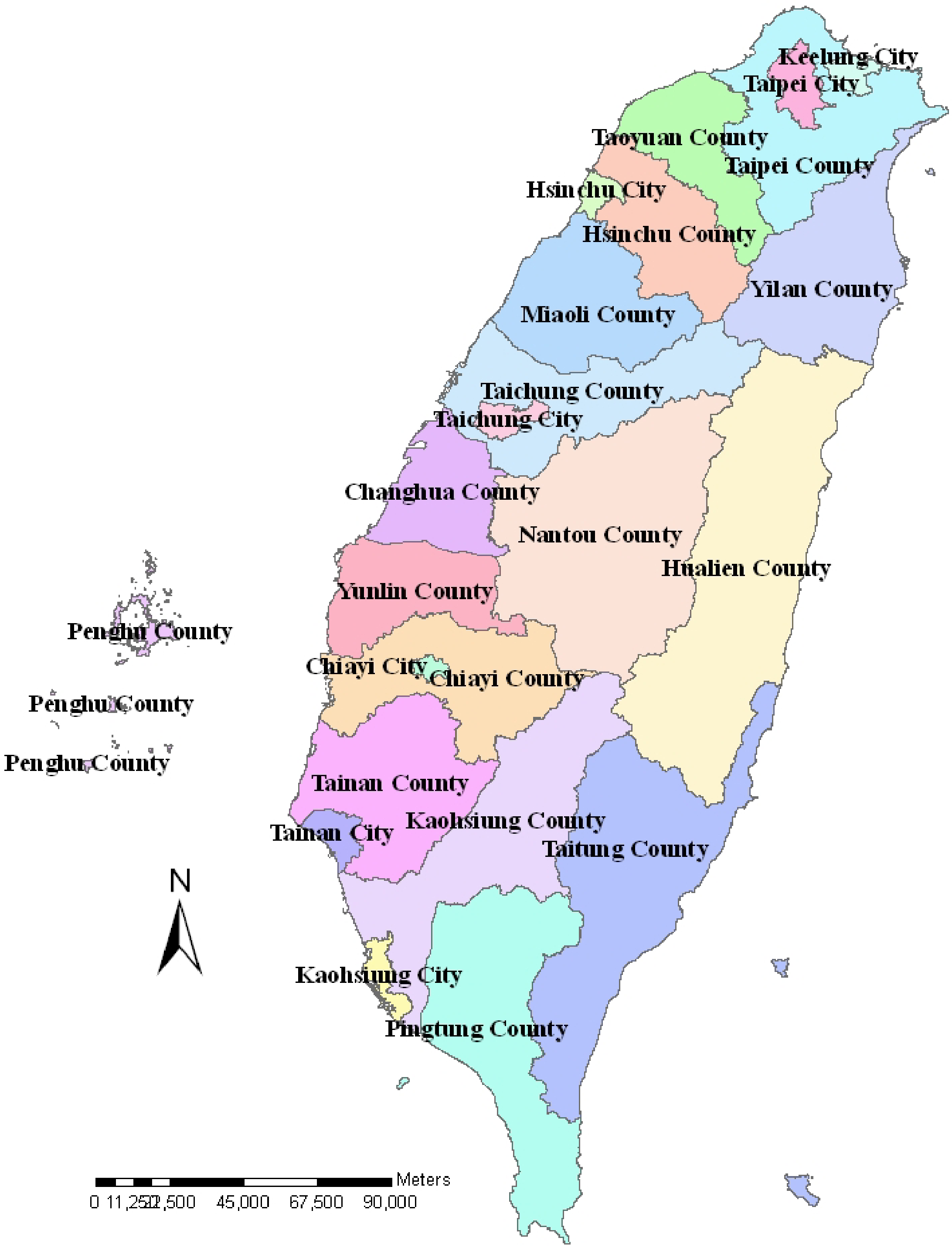


Figure 1. Geographical administrative distribution of Taiwan

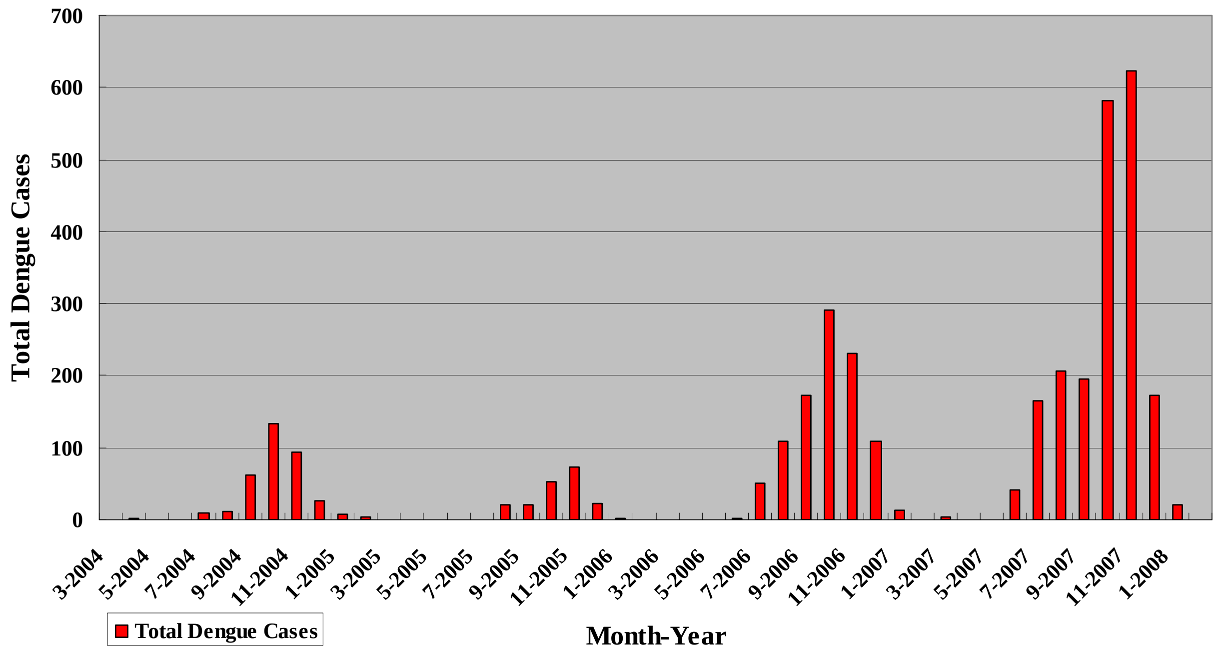


Figure 2. Monthly total dengue cases in Taiwan, March 2004 - February 2008

Dengue Incidence

The dengue case data in this study, including onset dates of a confirmed dengue cases and their locations (city/county), were collected from March 2004 to February 2008, provided by CDC-Taiwan. Total population size in each city/county was obtained from the Ministry of Interior-Taiwan. Dengue incidence is defined as the number of new dengue cases in each city/county that occur in each season divided by the total number of people in each city/county during that period of time.

Geographical and Climatic Data

Geographical data, including altitude distribution, river distribution, and geographical administrative distribution of Taiwan were obtained from the RITI Technology Inc.

Climatic data, including monthly average temperature and monthly average rainfall from March 2004 to February 2008 were obtained from Central Weather Bureau- Taiwan.

2.2 METHODS

2.2.1 Software

■ ArcGIS 9.0 Software [7]

ArcGIS is the name of a group of geographic information system software product lines produced by ESRI. ArcGIS consists of Desktop GIS products, which allows users to perform spatial analysis, model operational processes, and visualize results on a map. ArcGIS Desktop is available at different product levels, with increasing functionality: ArcReader, which allows users to view, print, and query maps; ArcView, which allows users to view and edit spatial data, create maps, and perform basic spatial analysis; ArcEditor which, in addition to the functionality of ArcView, includes more advanced tools for manipulation of shapefiles and geodatabases; ArcInfo, the most advanced version of ArcGIS, which allows users the most flexibility and control in all aspects of data building, modeling, analysis, and map display.

The spatial distribution of dengue incidence, climatic factors and geographical factors were mapped by the ArcGIS 9.0 software.

■ GeoDa 095i Software [8-10]

GeoDa is a free software package that conducts spatial data analysis, geovisualization, spatial autocorrelation and spatial modeling. The package was developed by the Spatial Analysis Laboratory of the University of Illinois at Urbana-Champaign under the direction of Luc Anselin.

GeoDa has powerful capabilities to perform spatial analysis, multivariate exploratory data analysis, and global and local spatial autocorrelational analyses. It also performs basic linear regression. As for spatial models, both the spatial lag model and the spatial error model, both estimated by maximum likelihood, are included.

In this study, all spatial autocorrelation analyses were computed using GeoDa 0.9.5i software [13, 14].

2.2.2 Empirical Bayes Smoothing (EBS) Method [11-14]

In recent years, a considerable amount of effort has been put into the mapping of mortality and disease rates to display the geographical variability of diseases. It is known that when rates are estimated from where populations across regions widely vary, the results are inherently unstable. Therefore, a number of techniques have been developed to correct for the intrinsic variance instability of rates by “smoothing” the rate estimate. The motivation for using these techniques comes from the Bayesian inference, which depends on a prior distribution of data to get a posterior distribution from which the parameters of interest are estimated. A specific Bayesian technique, the Empirical Bayes smoothing (EBS) method, is used in this study. This method mostly affects the rate for areas with a small sized population at risk.

The EBS method consists of estimating the moments of the prior distribution from the data and the raw rate is "shrunk" towards an overall mean. The amount of shrinkage is inversely proportional to the population size. In other words, areas with a relatively small population at risk will tend to have their rate adjusted considerably, whereas for areas with a relatively large population at risk the rate will barely change.

Suppose a region to be mapped is partitioned into N mutually exclusive areas, an individual area being notated as i (i=1, 2, …, N). Let θ_i be the event rate in area i; E_i be the count of events, which is distributed as a Poisson random variable; and n_i be the population at risk in area i. The maximum likelihood (ML) estimator of θ_i is $p_i=E_i/n_i$ (the "raw rate"). Adopting a Bayesian framework, suppose that θ_i has a prior distribution characterized by a mean $m_i=E_\theta(\theta_i)$ and variance $\phi_i=\mathrm{Var}(\theta_i)$. The Bayesian estimator for θ_i, $\hat{\theta}_i$, then becomes a weighted average of the raw rate p_i and m_i:

$$\hat{\theta}_i = m_i + w_i(p_i - m_i) \tag{1}$$

with

$$w_i = \frac{\phi_i}{\phi_i + (m_i / n_i)}$$

The model can be simplified by taking $m_i=m$ and $\phi_i=\phi$ for all i. There are three methods to estimate m and ϕ: ML method, method of moments, and a mixed method of ML and moments. The method of moments is used in the study as it is a distribution-free and non-iterative procedure and easy to compute compared to the ML method, which requires a full specification of the prior and subsequent derivation of the marginal likelihood values.

To estimate m: Since $E_p(p_i) = E_\theta\{(p_i|\theta_i)\} = E_\theta(\theta_i) = m$, any weighted mean of the p_i provides an unbiased estimate of m. So, the m can be estimated by the overall mean, $\tilde{m} = \frac{\sum_i E_i}{n}$, where $n = \sum_i n_i$ is the total population at risk.

To estimate ϕ ; consider a weighted sample variance, $s^2 = \frac{\sum_i n_i (p_i - \tilde{m})^2}{n}$, and $E_p(p_i - \tilde{m})^2 \cong \text{var}_p(p_i) = \phi + \frac{m}{n_i}$ by ignoring the error in using $\tilde{m}$ as an estimate of m. Thus, $E_p(s^2) = \frac{\sum_i n_i (\phi + m / n_i)}{n} = \phi + \frac{m}{\bar{n}}$, where $\bar{n} = \frac{n}{N}$ is the mean population at risk. Therefore, with m replaced by $\tilde{m}$, a moment estimate of ϕ is $\tilde{\phi} = s^2 - \tilde{m} / \bar{n}$, and $\tilde{\phi} = 0$ when $s^2 < \tilde{m} / \bar{n}$. With m and ϕ replaced by $\tilde{m}$ and $\tilde{\phi}$ in equation (1), the Empirical Bayes smoothed estimator of θ_i, $\tilde{\theta}_i$, is calculated by the following formula:

$$\tilde{\theta}_i = \tilde{m} + \tilde{w}_i (p_i - \tilde{m})$$

with

$$\tilde{w}_i = \frac{s^2 - \tilde{m} / \bar{n}}{s^2 - \tilde{m} / \bar{n} + \tilde{m} / n_i}$$

When the population at risk in area i, n_i, is small, the $\tilde{w}_i$ (shrinkage factor) is near zero, and the value of the Empirical Bayes smoothed estimator ($\tilde{\theta}_i$) is shrunk to the overall mean. When the n_i is large, the $\tilde{w}_i$ is near 1, and the $\tilde{\theta}_i$ remains essentially unchanged.

The raw dengue incidence, defined as the number of new dengue cases in each city/county that occur in each season divided by the total number of population in each city/county during that period of time, was smoothed using Empirical Bayes Smoothing (EBS) technique by the GeoDa 0.9.5i software in order to correct for the variance instability of dengue incidence as a result of heterogeneity in dengue cases and population. The EBS technique consists of computing a weighted average between the raw dengue incidence for each city/county and the whole cities/counties average, with weights proportional to the underlying population at risk. In effect, cities/counties with a relatively small population will tend to have their dengue incidence adjusted considerably, whereas for cities/counties with a relatively large population the dengue incidence will barely change.

2.2.3 Spatial Autocorrelation Analysis [15-18]

The correlation pattern among neighboring dengue incidence and the level of spatial autocorrelation among neighboring areas were measured by the spatial autocorrelation analyses (EB adjusted Moran's I tests and EB adjusted LISA tests). Spatial autocorrelation of dengue incidence was calculated for the summer, fall, and winter during 2004 to 2007. The spatial autocorrelation was not calculated for spring because of very few reported cases in this season.

Spatial autocorrelation is classified as either positive or negative. A positive spatial autocorrelation refers to a map pattern where geographic features of similar values tend to cluster on a map, whereas a negative spatial autocorrelation indicates a map pattern in which geographic features of similar values scatter throughout the map. When no statistically significant spatial autocorrelation exists, the pattern of spatial distribution is considered random.

There are two types of spatial autocorrelation analysis: theglobal spatial autocorrelation test (test for clustering) and the local spatial autocorrelation test (test for clusters). Global spatial autocorrelation statistics (Moran's I was used in this study) estimates the overall degree of spatial autocorrelation (clustering) for a dataset. It shows clustering but does not show where the clusters are. So, local spatial autocorrelation statistics (LISA was used in this study) is used to measure the degree to which a location is surrounded by like values, yielding a measure of spatial autocorrelation for each individual location.

The first step in the computation of spatial autocorrelation statistics is to construct a spatial weights file that defines the "neighborhood" structure for each location. There are many ways to construct a weights file. A contiguity-based spatial weights (w_{ij}) is considered in this study, where the definition of neighbor is based on sharing a common boundary. There are two types of contiguity-based spatial weights: rook contiguity and queen contiguity. Rook contiguity uses only common boundaries to define neighbors, while queen contiguity includes all common points (boundaries and vertices). We first used spatial weights based on queen contiguity as they always have a denser connectedness structure (more neighbors). We also used spatial weights based on rook contiguity to check robustness of results.

■ Global Spatial Autocorrelation

Suppose a region is partitioned into N mutually exclusive areas, an individual area being notated as i (i=1, 2, …, N)**.** Let x_i and n_i be the count of events and the population at risk in area i, respectively. The observed rate in area i is defined as $p_i = x_i/n_i$. Moran's I statistic, which is similar to the Pearson correlation coefficient, is calculated as:

$$I = \frac{N}{S_0}\sum_i \sum_j w_{ij} \frac{(p_i - \overline{p})(p_j - \overline{p})}{\sum_i (p_i - \overline{p})^2},$$

$$S_0 = \sum_{ij, i \neq j} w_{ij}$$

where w_{ij} is the contiguity-based spatial weights (w_{ij}=1 if areas i and j are contiguous, else w_{ij}=0), $\overline{p} = \sum_i p_i / N$ is the mean rate for the entire region. Moran's I ranges between −1 and 1 all the time, with large positive values indicating clustering, with large negative values indicating dispersion, and values close to zero indicating absence of spatial autocorrelation.

Moran's I statistics can be visualized in a graph, "Moran scatter plot", where the spatially lagged variable (a sum of spatial weights multiplied with values for observations at neighboring locations) is on the vertical axis and the original variable is on the horizontal axis. Moran's I statistics is obtained from the slope in the graph. The four quadrants in the graph provide four types of spatial autocorrelation: high-high (upper right), low-low (lower left), for positive spatial autocorrelation; high-low (lower right) and low-high (upper left), for negative spatial autocorrelation (see Figure 3).

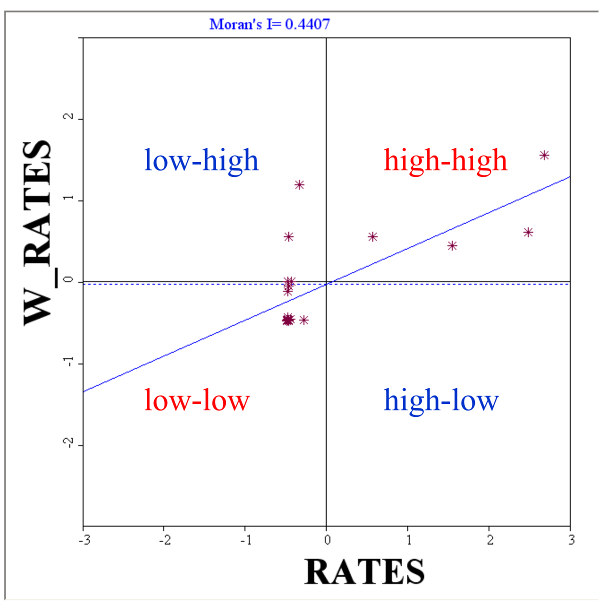


Figure 3. Moran scatter plot

■ **Local Spatial Autocorrelation**

Local autocorrelation analysis decomposes the global measure into the contributions for each location, detecting similarities or dissimilarities in values around a given rate of a certain event.

Suppose a region is partitioned into N mutually exclusive areas, an individual area being notated as i (i=1, 2, …, N). Let x_i and n_i be the count of events and the population at risk in area i, respectively. The observed rate in area i is defined as $p_i = x_i/n_i$. LISA (local indicator of spatial association) statistic, which can be seen as the local Moran's I, was calculated as:

$$I_i = \frac{(p_i - \overline{p})}{\sum_i (p_i - \overline{p})^2} \cdot \sum_j w_{ij}(p_j - \overline{p}), i \neq j$$

where I_i is the LISA statistic for each area i, w_{ij} is the contiguity-based spatial weights (w_{ij}=1 for contiguous areas i and j, else w_{ij}=0), p_i and p_j are the observed rates for area i and j, $\overline{p} = \sum_i p_i / N$ is the mean rate for the entire region.

For each location, LISA values allow for the computation of its similarity with its neighbors and also to test its significance. There are four types of the local spatial autocorrelation:

- Locations with high values with similar neighbors: high-high.
- Locations with low values with similar neighbors: low-low.
- Locations with high values with low-value neighbors: high-low.
- Locations with low values with high-value neighbors: low-high.
- Locations with no significant local autocorrelation.

The high-high and low-low locations (positive local spatial autocorrelation) are typically referred to as spatial clusters, while the high-low and low-high locations (negative local spatial autocorrelation) are referred to as spatial outliers.

LISA statistics can be visualized in a "LISA cluster map", which shows the locations with significant LISA statistics (p-value<0.05 is used in the study) with four colors coded by type of spatial autocorrelation: dark red for high-high, dark blue for low-low, pink for high-low, and light blue for low-high. These four categories correspond to the four quadrants in the Moran scatter plot. (See Figure 4)

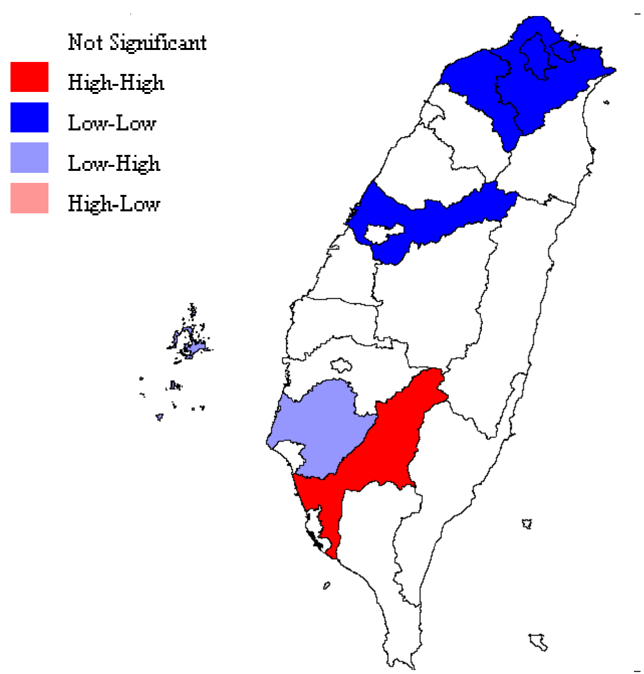


Figure 4. LISA cluster map

Moran's I and LISA statistics are calculated based on the assumption that the rates are independently and identically distributed (i.i.d.) random variables with a normal distribution. This implies that the expected value of the rates is constant in all areas. This assumption is usually violated when rate is varied in different areas. Therefore, EB adjusted Moran's I tests and EB adjusted LISA tests, proposed by Assuncao and Reis (1999) [18], were performed in this study to adjust for the inconstant variance. This adjustment procedure uses a variable transformation based on the Empirical Bayes (EB) principle.

Suppose a region is partitioned into N mutually exclusive areas, an individual area being notated as i (i=1, 2, …, N). Let x_i and n_i be the count of events and the population at risk in area i, respectively; and θ_1, …, θ_N be the unknown and possibly different underlying rate of the areas. Suppose x_i observed events during a reference period has a Poisson distribution with conditional mean $E(x_i|\theta_i) = n_i\theta_i$. The estimated rate p_i has conditional mean $E(p_i|\theta_i) = \theta_i$ and variance $var(p_i|\theta_i) = \theta_i/n_i$. Therefore, the estimated rates have different conditional means and variances.

Adopting a mixing approach, suppose the underlying rates θ_i have a *priori* with mean β and variance α. Hence, the marginal expectation of p_i is β and the marginal variance is $\alpha + \beta/n_i$. Now, only the variances differ among the areas and it increases as the population decreases. The moment estimates proposed by Marshall for the parameters α and β given by $a = s^2 - b/(n/N)$ and $b = x/n$, respectively, where $s^2 = \sum n_i (p_i - b)^2 / n$. Therefore, the marginal expectation and variance of p_i are estimated by b and $v_i = a + b/n_i$, respectively. If $v_i < 0$, v_i is set to be b/n_i.

Instead of using the rate p_i, EB adjusted Moran's I uses a deviation of the estimated marginal mean standardized by an estimate of its standard deviation:

$$z_i = \frac{p_i - b}{\sqrt{v_i}}$$

EB adjusted Moran’s I is defined as

$$EBI = \frac{N}{\sum W_{ij}} \frac{\sum W_{ij} z_i z_j}{\sum (z_i - \bar{z})^2}$$

Like Moran's I, EBI will tend to be positive if the rates are spatially correlated.

2.2.4 Correlation Analysis

The correlation between dengue incidence and climatic factors, including temperature and rainfall, in summer, fall, and winter from 2004 to 2007 was evaluated by the Spearsman’s Correlation test.

3.0 RESULTS AND DISCUSSIONS

3.1 SPATIAL AND TEMPORAL DESCRIPTION OF DENGUE INCIDENCE AND CLIMATIC FACTORS

Figures 5 through 7 demonstrate geographical variation in dengue incidence, temperature, and rainfall in four different seasons, respectively. During the dengue-occurring seasons, the dengue incidence was higher in southern Taiwan, including Tainan county, Tainan city, Kaohsiung county, Kaohsiung city, and Pingtung county. Dengue annually occurred from summer, peaked in fall and went down in winter. In the peak season fall, adjusted by EBS method, the highest dengue incidence was 0.272 per 1000 people in Pingtung county in 2004, 0.052 per 1000 people in Tainan city in 2005, 0.358 per 1000 people in Kaohsiung city in 2006, and 1.336 per 1000 people in Tainan city in 2007, respectively.

The average temperature in each city/county increased from the north to the south in four seasons except the Nantou county, which is the only landlocked and 83% of total area covered by hills and mountains (with the tallest mountain located) county in Taiwan. When Nantou county was not considered, the variance in temperature among cities/counties was small in summer (from 26˚C to 29˚C), but it was large in winter (from 15˚C to 22˚C). The proper temperature for mosquito breeding is between 25˚C and 27˚C and viruses can reproduce when the temperature is greater than 18˚C. Unfortunately, the average temperature in southern Taiwan was greater than 20°C in the whole year, which provides a potential mosquito and virus growth area (see Figure 6).

The distribution of rainfall varied strongly with the season. In spring, the average rainfall decreased from the north to the south while it increased from the north to the south in summer. In fall, the average rainfall increased from the north to the south and it was higher in the east than in the west. In winter, the average rainfall was low (less than 100mm, except Yilan County and Keelung City) and decreased from the north to the south. The rainfall was concentrated in summer, where the average rainfall was greater than 300 mm nationwide (see Figure 7).

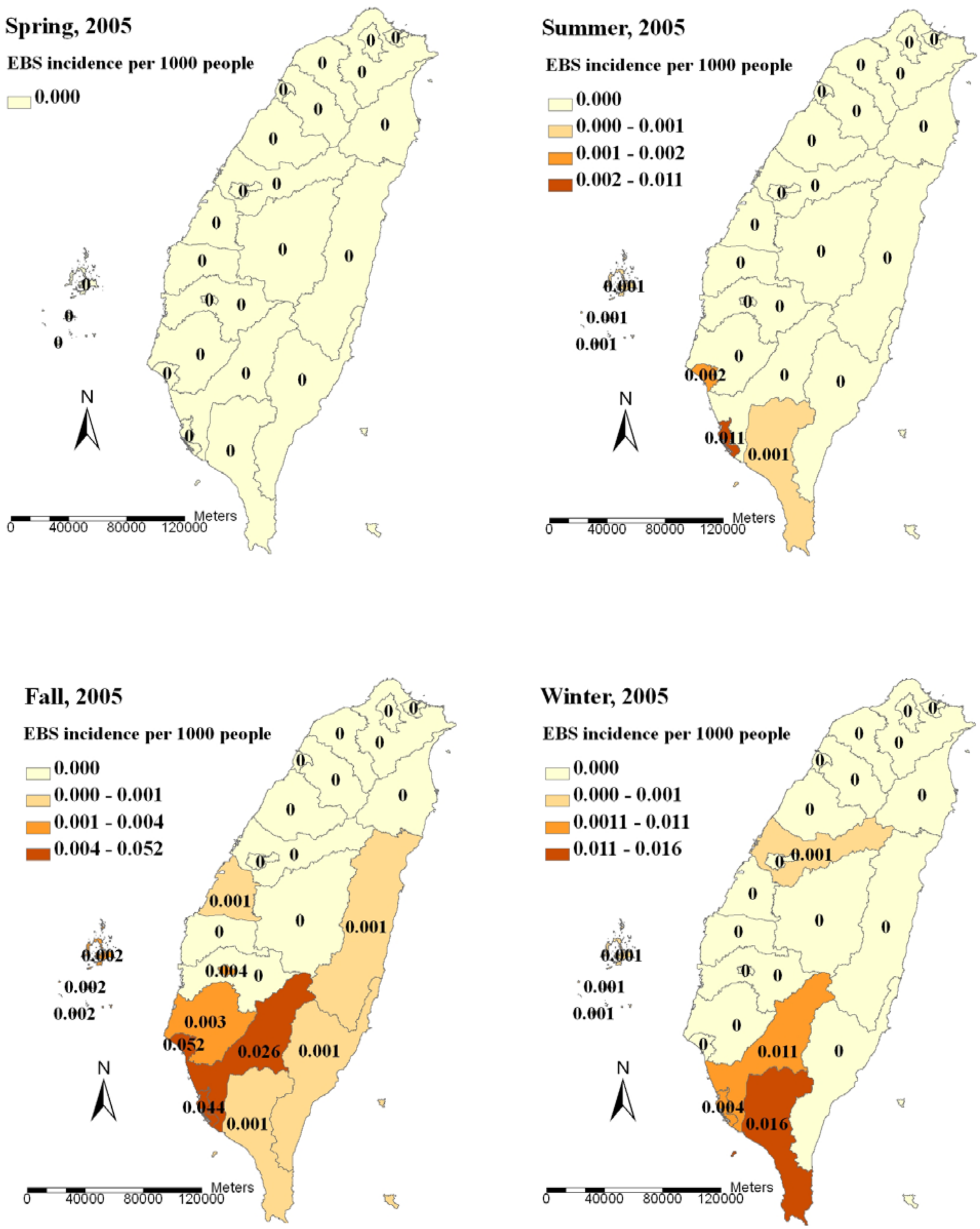


Figure 5. EBS dengue incidence in four seasons, 2005

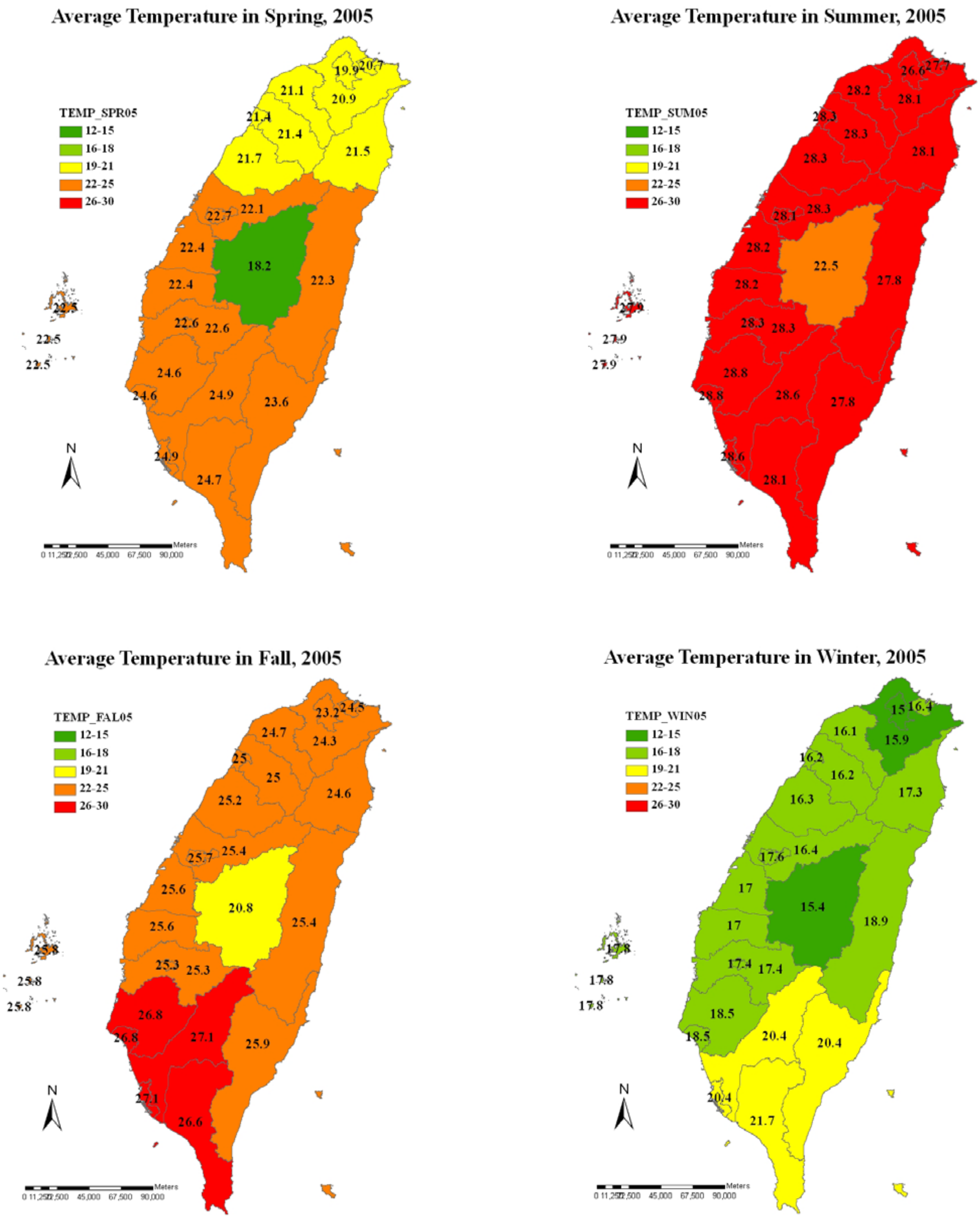


Figure 6. Average temperature in four seasons, 2005

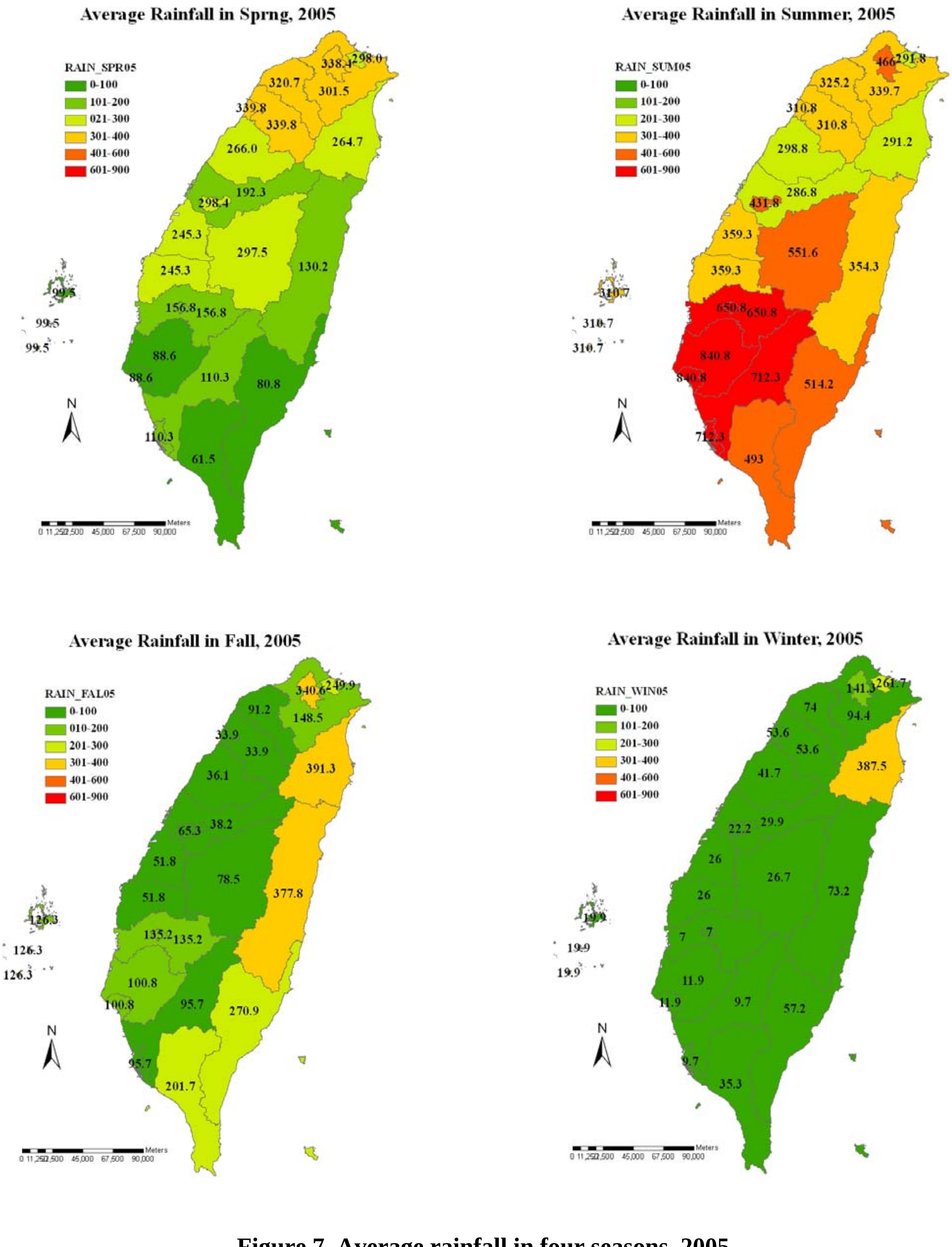


Figure 7. **Average rainfall in four seasons, 2005**

3.2 SPATIAL AUTOCORRELATIONAL ANALYSES OF DENGUE INCIDENCE

A spatial autocorrelation analysis was carried out to determine the clustering of dengue incidence. Based on global spatial autocorrelation analysis, there was a positive and statistically significantly spatial autocorrelation for cumulative dengue incidence from 2004 to 2007 (EB adjusted Moran's I=0.171, *p*-value=0.03). The result of a positive EB adjusted Moran's I indicated that dengue incidence of similar values tended to cluster on a map. Moreover, the EB adjusted Moran's I for the year 2005, 2006, and 2007 were statistically significant (*p*-value<0.05). However, there was no significant spatial autocorrelation of dengue in 2004. It was possible that only a very small number of cities/counties (seven) had dengue occurred in that year, which mad it hard to determine the spatial autocorrelation (see Table 1a and 1b). Detail Moran scatter plots are given in Appendix A. Further, according to the EB adjusted LISA cluster map of cumulative dengue incidence for 2004 to 2007 shown in Figure 8a and 8b, high dengue incidence were significantly clustered around Tainan county and Kaohsiung county (*p*-value<0.05). Kaohsiung county could be determined as a significant endemic place because it had a significant dengue incidence cluster in each year of 2004 to 2007 (*p*-value<0.05, see Figure 9a and 9b). The EB-adjusted Moran's I values are found to be robust for two types of spatial weights.

Table 1. EB-adjusted Moran's I for the cumulative dengue incidence (2004 – 2007) and the dengue incidence in the specific years.

1a. Use queen-contiguity spatial weights

Year	EB-adjusted Moran's I	*p*-value
2004-2007	0.171*	0.03
2004	0.003	0.09
2005	0.441*	0.001
2006	0.307*	0.001
2007	0.114*	0.016

* Significant at 0.05 level

1b. Use rook-contiguity spatial weights

Year	EB-adjusted Moran's I	*p*-value
2004-2007	0.171*	0.04
2004	0.003	0.09
2005	0.441*	0.003
2006	0.307*	0.001
2007	0.114*	0.003

* Significant at 0.05 level

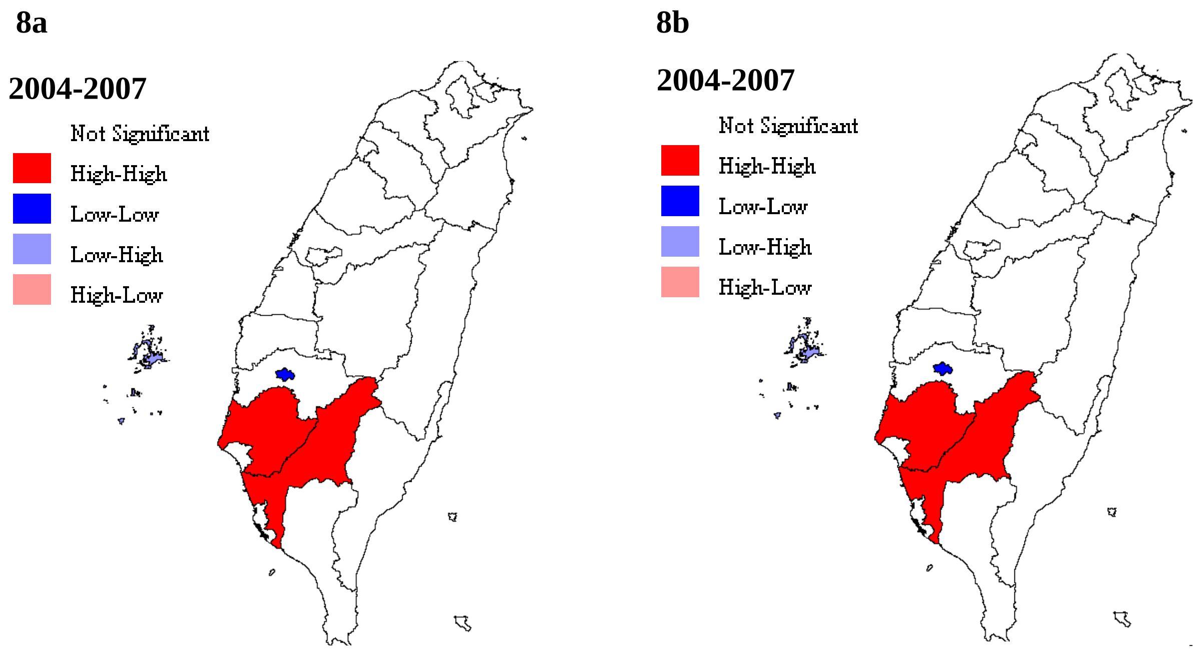
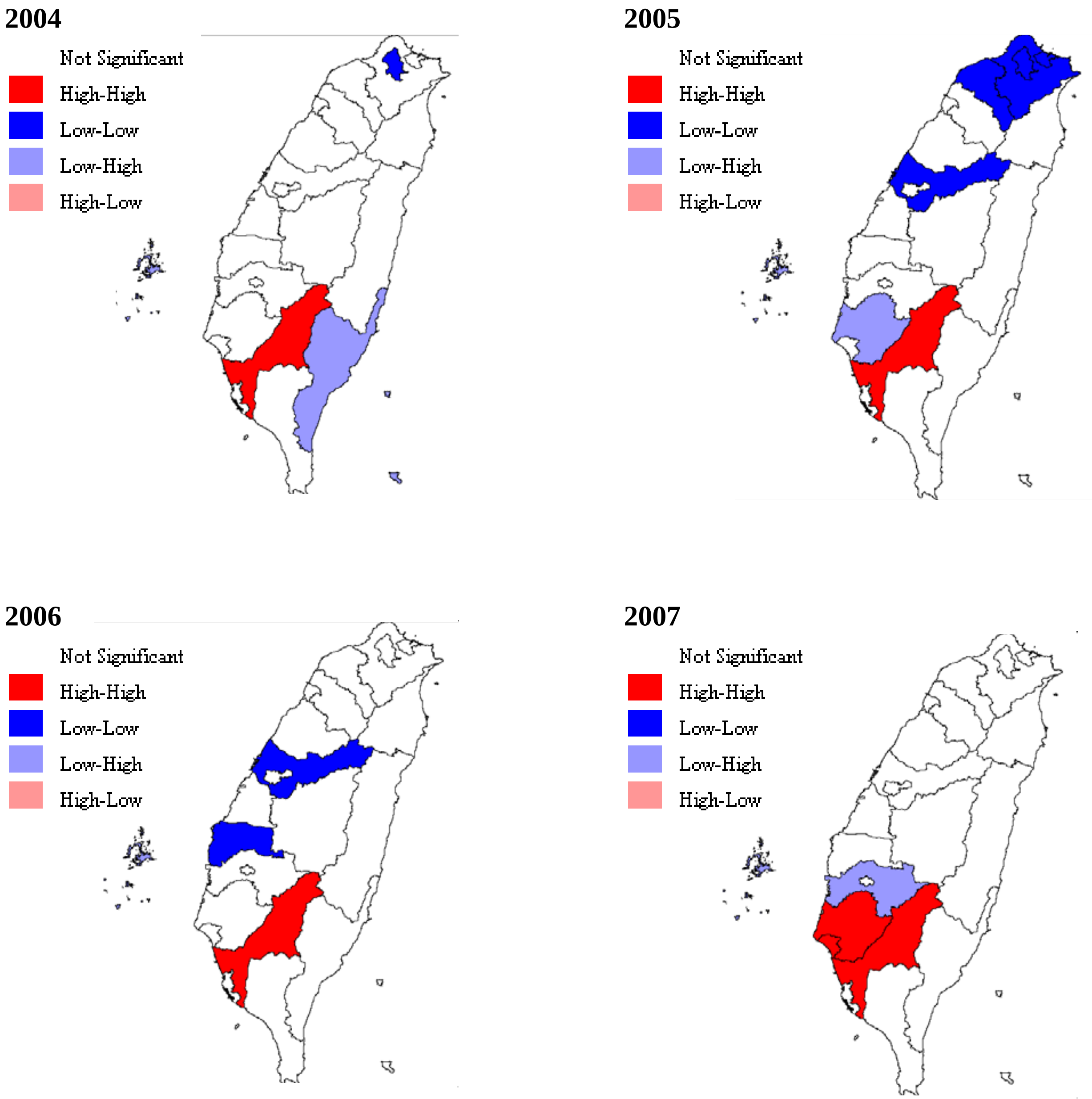


Figure 8. EB-adjusted LISA cluster map of cumulative dengue incidence for 2004-2007: 8a. using Queen-contiguity spatial weights and 8b. using Rook-contiguity spatial weights

9a


2004
Not Significant
High-High
Low-Low
Low-High
High-Low
2005
Not Significant
High-High
Low-Low
Low-High
High-Low
2006
Not Significant
High-High
Low-Low
Low-High
High-Low
2007
Not Significant
High-High
Low-Low
Low-High
High-Low

9b

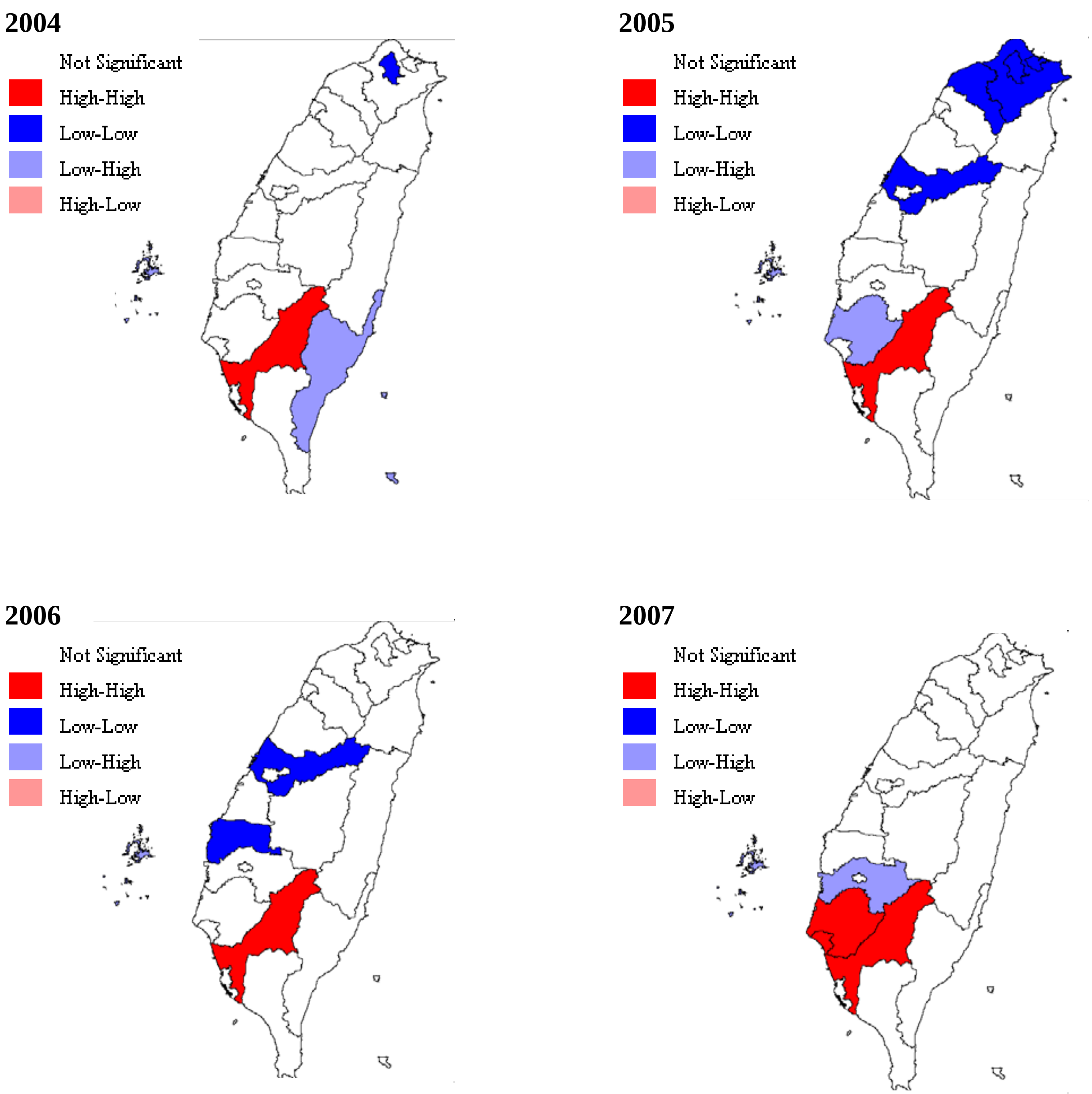


Figure 9. EB-adjusted LISA cluster map of dengue incidence for 2004, 2005, 2006, and 2007: 9a. using queen-contiguity spatial weights and 9b. using rook-contiguity spatial weights

3.3 RELATIONSHIP BETWEEN DENGUE INCIDENCE AND GEOGRAPHICAL AND CLIMATIC FACTORS

Figure 10 showed the relationship between dengue and the geomorphology and the hydrology of Taiwan. Taiwan is a mountainous island with 30% mountains area, 40% hills and plateaus area, and 30% plains area. Taiwan's mountainous area spreads from the east to the central part of the island. Most of the mountain ranges go from north to south. There are 129 rivers in Taiwan, most of which flow toward the east or west. Because the major watershed, Snow Mountain Range, has an eastward inclination, the drainage area of western Taiwan is larger than that in the east. Based on these geographical features, dengue tended to occur in the southwestern cities/counties in Taiwan with plains and rivers spread.

Table 2 showed the Spearman's correlation between dengue incidence and climatic factors, temperature and rainfall, in summer, fall, and winter from 2004 to 2007. Average temperature had a positive relationship with dengue incidence in summer and fall (r_s=0.75 and *p*-value <0.001 in summer, r_s=0.51 and *p*-value=0.002 in fall), which suggested that the higher the temperature, the higher the dengue incidence in summer and fall. This may be due to temperature's influence on the life cycle of a mosquito or viral replication rates. However, there was no significant correlation between temperature and dengue incidence in winter (*p*-value =0.099). This may be due to a low temperature (less then 20 °C) in most cities/counties in Taiwan during the winter, which was not proper for mosquito breeding (25°C to 27 °C) and virus reproduction (greater than 18 °C).

Average rainfall had a positive relationship with dengue incidence in summer (r_s=0.61 and *p*-value=0.002) However, there was no significant correlation between rainfall and dengue incidence in fall and winter.

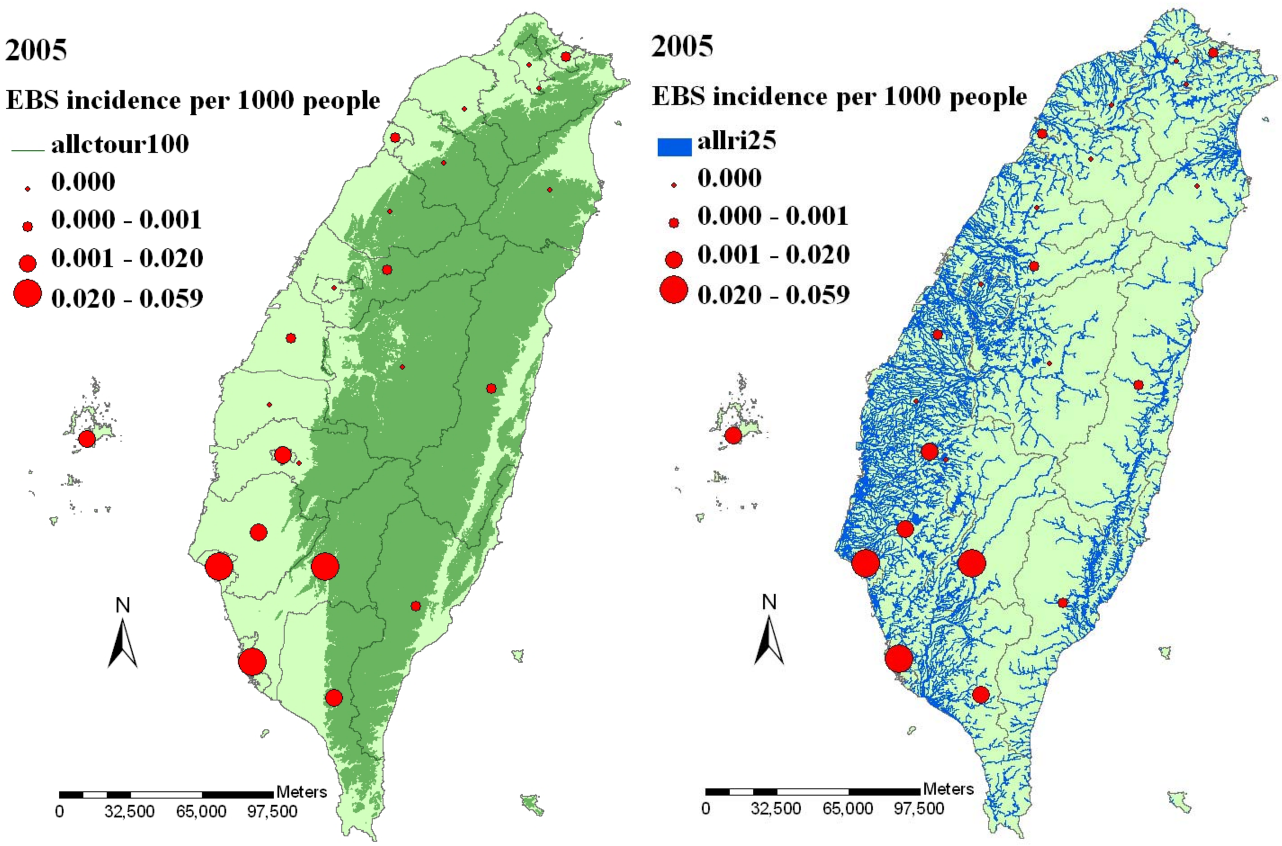


Figure 10. EBS dengue incidence in 2005 with the geomorphology (left) and the hydrology (right) of Taiwan

Table 2. Spearman's correlation between dengue incidence and climatic factors, 2004-2007

Variables / season	Spearman's correlation (r_s)	*p*-value
Temperature and Dengue incidence		
Summer	0.75*	<0.000
Fall	0.51*	0.002
Winter	0.38	0.099
Rainfall and Dengue incidence		
Summer	0.61*	0.002
Fall	-0.38**	0.047
Winter	-0.01	0.956

* Significant at 0.05 level

** Barely borderline significant

4.0 CONCLUSION

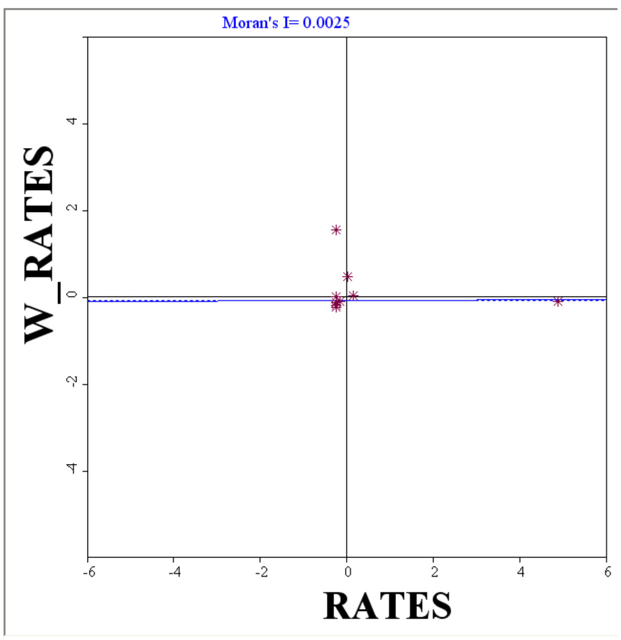
This study demonstrated the use of GIS methodology to map and analyze the spatial and temporal distribution of dengue in Taiwan and elucidate the association of geographical and climatic risk factors with dengue incidence. The result of the study suggested that: 1) Dengue had a temporal pattern, where it annually occurred from summer, peaked in fall and went down in winter; 2) Through spatial autocorrelation analyses, non-randomness in the distribution of dengue and the identification of unusual concentration of dengue in Tainan county and Kaohsiung county (the southern Taiwan) has been defined. This could prompt health planners in the county/city to take a critical look at these risk areas, and make appropriate health planning and resource allocation. 3) Based on the geographical features, dengue tended to occur in the southwestern cities/counties in Taiwan with plains and rivers spread. 4) High temperature and high rainfall are impartment risk factors of dengue in Taiwan. Due to the limited available resource, there are many other possible risk factors of dengue which were not included in the study. Thus, a more detailed research is required to consider factors like urban development and housing construction to thoroughly evaluate the risk of dengue in Taiwan.

APPENDIX A

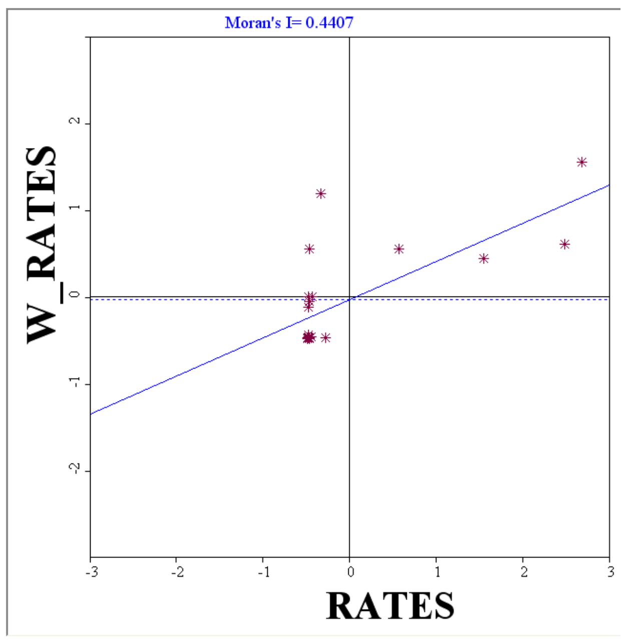
MORAN SCATTER PLOTS OF DENGUE INCIDENCE

1a

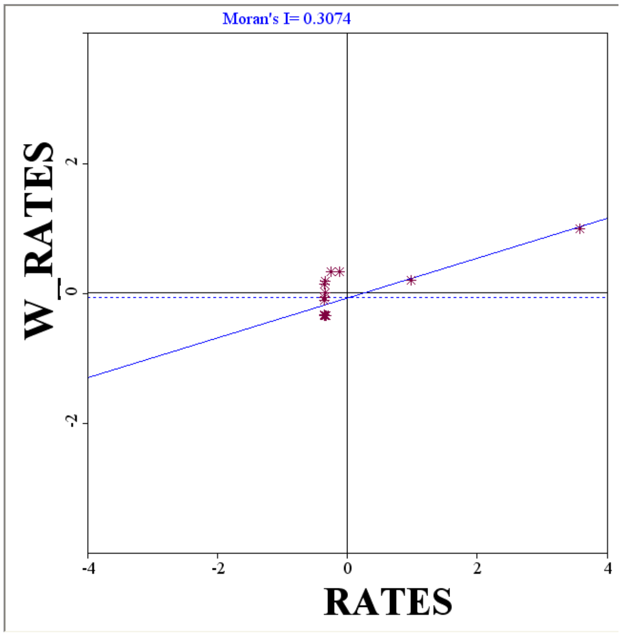
2004



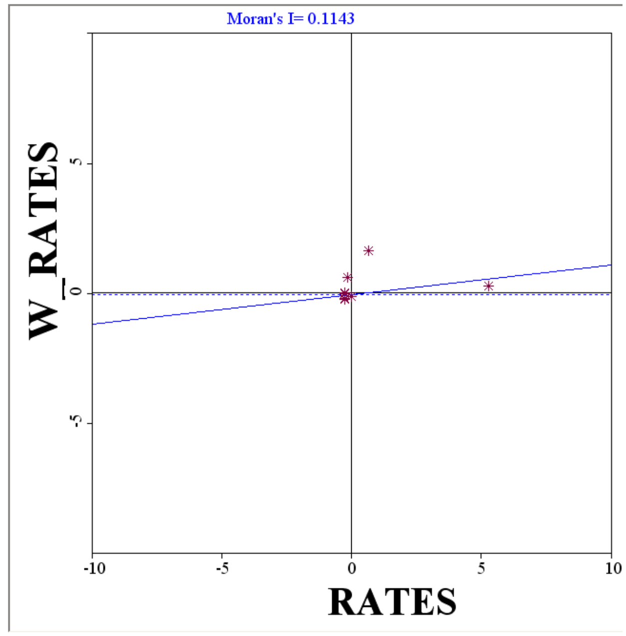
2005



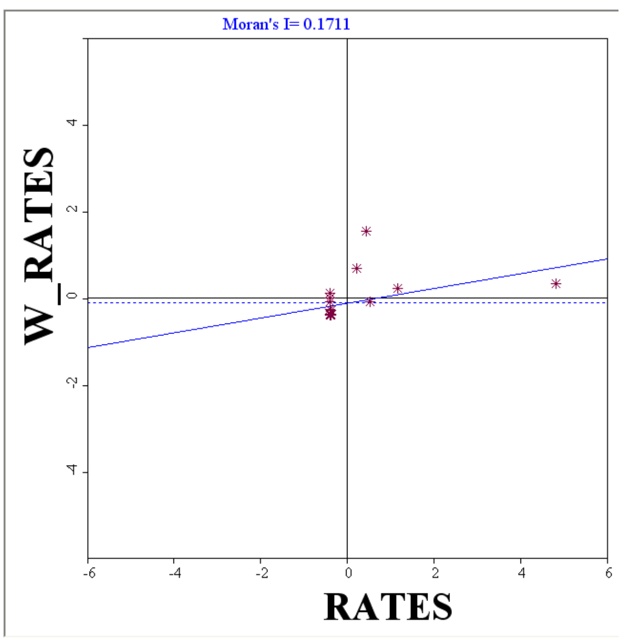
2006



2007

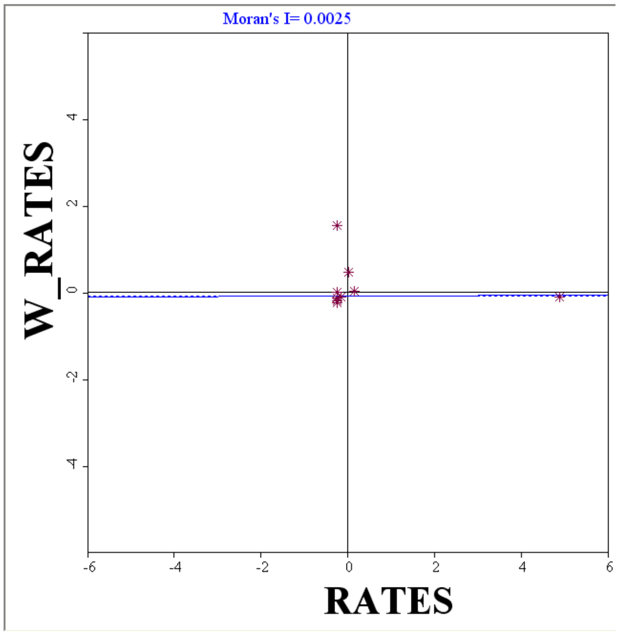


2004-2007

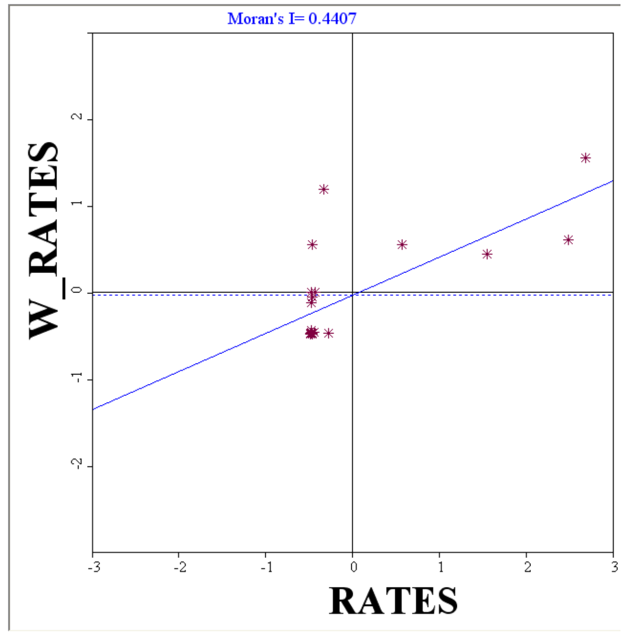


2a

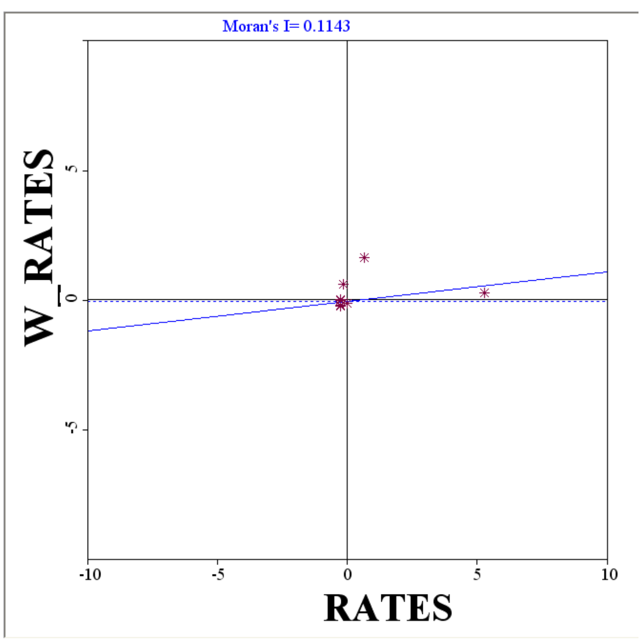
2004



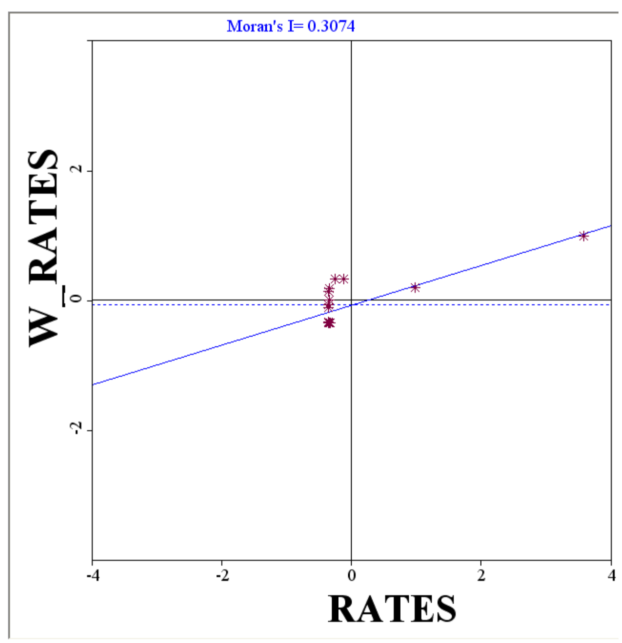
2005



2006



2007



2004-2007

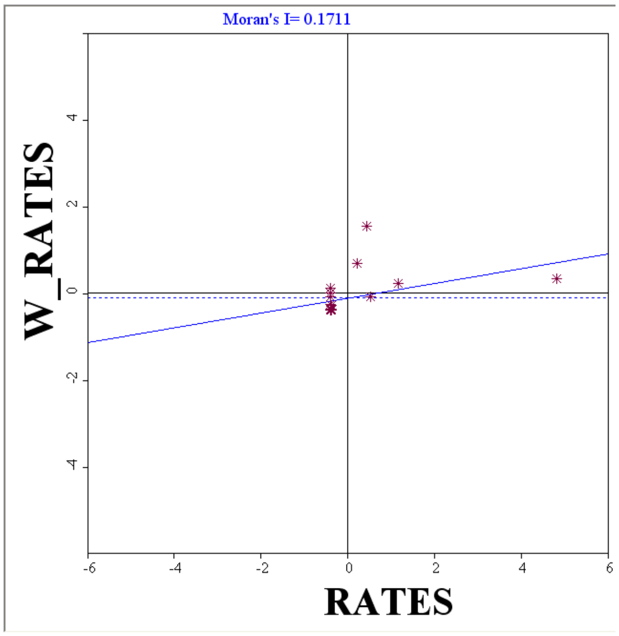


Figure 11. Moran scatter plot of dengue incidence for 2004, 2005, 2006, 2007, and 2004-2007: 11a. using queen-contiguity spatial weights and 11b. using rook-contiguity spatial weights

BIBLIOGRAPHY

1. Johansson MA, D.F., Glass GE, *Local and Global Effects of Climate on Dengue Transmission in Puerto Rico. PLoS Negl Trop Dis*, 2009 Feb. **3**(2): p. e382.
2. Ellen K. Cromley, S.L.M., ed. *GIS and Public Health*. 2002, The Guilford Press: New York. 340.
3. Jacquez., G.M., *Spatial analysis in epidemiology: Nascent science or a failure of GIS? J Geograph Syst.*, 2000. **2**: p. 91-97.
4. Annelise Tran, X.D., Philippe Dussart, Jacques Morvan, Patrick Rabarison, Franck Remy, Laurent Polidori, and Jacques Gardon, *Dengue Spatial and Temporal Patterns, French Guiana, 2001. Emerg Infect Dis.*, 2004 April. **10**(4): p. 615-21.
5. Ciu, H.-M., *Dengue in the Southern Taiwan Report*. 1993.
6. Wen TH, L.N., Lin CH, King CC, Su MD., *Spatial mapping of temporal risk characteristics to improve environmental health risk identification: a case study of a dengue epidemic in Taiwan. Sci Total Environ*., 2006 Aug. **31**(267): p. 631-40.
7. *ArcGIS: A Complete Integrated System*. 1969; Accessed on August 2009. http://www.esri.com/software/arcgis/.
8. *GeoDa Center For Geospatial Analysis and Computation*. Accessed on August 2009. http://geodacenter.asu.edu/.
9. Luc, A. *Exploring spatial data with GeoDa: A workbook (Revised version)*. 2005; Accessed on August 2009. http://www.csiss.org.
10. Luc, A. *GeoDa 0.9 User's Guide*. 2003; Accessed on August 2009. http://www.csiss.org.
11. Anselin, L., Y-W Kim and I. Syabri, *Web-Based Spatial Analysis Tools for the Exploration of Spatial Outliers.* J Geograph Syst., 2004. **6**: p. 197–218.
12. Clayton D., K.J., *Empirical Bayes estimates of age-standardized relative risks for use in disease mapping.* Biometrics, 1987 Sep. **43**: p. 671-81.
13. Kafadar, K., *Smoothing geographical data, particularly rates of disease. Stat Med.*, 1996 Dec. **15**(23): p. 2539-60.
14. Marshall, R.J., *Mapping disease and mortality rates using Empirical Bayes estimators. Applied Statistics*, 1991. **40**: p. 283-294
15. Osei FB, D.A., *Spatial and demographic patterns of Cholera in Ashanti region-Ghana. Int J Health Geogr.*, 2008 Aug. **7**(44).
16. Jepsen MR, S.J., Ethelberg S, *Spatio-temporal cluster analysis of the incidence of Campylobacter cases and patients with general diarrhea in a Danish county, 1995-2004. Int J Health Geogr.*, 2009 Feb. **8**(11).
17. Vanwambeke SO, v.B.B., Khantikul N, Burghoorn-Maas C, Panart K, Oskam L, Lambin EF, Somboon P., *Multi-level analyses of spatial and temporal determinants for dengue infection. Int J Health Geogr.*, 2006 Jan. **5**(5).

18. Assunção RM, R.E., *A new proposal to adjust Moran's I for population density.* Stat Med. , 1999 Aug. **18**(16): p. 2147-62.