ORIGINAL PAPER

Spatial risk assessment of typhoon cumulated rainfall: a case study in Taipei area

Yun-Huan Lee · Hong-Ding Yang · Chun-Shu Chen

Abstract Typhoon is one of the most destructive disasters in Taiwan, which usually causes many floods and mudslides and prevents the electrical and water supply. Prior to its arrival, how to accurately forecast the path and rainfall of typhoon are important issues. In the past, a regression-based model was the most applied statistical method to evaluate the associated problems. However, it generally ignored the spatial dependence in the data, resulting in less accurate estimation and prediction, and the importance of particular explanatory variables may not be apparent. Therefore, in this paper we focus on assessing the spatial risk variations regarding the typhoon cumulated rainfall at Taipei with respect to typhoon locations by using the spatial hierarchical Bayesian model combined with the spatial conditional autoregressive model, where the model parameters are estimated by designing a family of stochastic algorithms based on a Markov chain Monte Carlo technique. The proposed method is applied to a real data set of Taiwan for illustration. Also, some important explanatory variables regarding the typhoon cumulated rainfall at Taipei are indicated as well.

Y.-H. Lee
Finance Department, Ming Chuan University, Taipei, Taiwan

H.-D. Yang
Department of Mathematics, National Changhua University of Education, Changhua, Taiwan

C.-S. Chen (✉)
Institute of Statistics and Information Science, National Changhua University of Education, Changhua, Taiwan
e-mail: cschen@cc.ncue.edu.tw

1 Introduction

Taiwan is an island country and locates in the subtropical zone. Due to its special geographical position, earthquakes and typhoons both are major natural disasters. In this paper, we focus on discussing the associated problems of typhoon. In Taiwan, typhoon usually happens in summer from early May to late August and attacks us about three or four times every year. It generally brings heavy rain and strong wind, resulting in many floods and mudslides and stops the electrical and water supply. At the same time, typhoon also usually causes heavy economic loss and many people die. In recent years, weather satellites have been able to help us to gain more information about typhoon prior to its arrival, so we may go through a series of preparatory tasks to prevent possible damages; even so, typhoon is still the main cause of disaster loss in Taiwan every year. Therefore, in addition to taking necessary precautions against typhoon, how to develop an efficient statistical methodology to accurately assess the path, rainfall, or intensity of typhoon prior to its arrival is essential. As pointed out by Yeh et al. (1999), there are relatively few methods that can be applied for solving the associated problems. In the past, a regression-based model is a popular statistical method. For example, Neumann (1972) and Xu and Neumann (1985) used a regression model to forecast the path of typhoon, while DeMaria and Kaplan (1994) and Fitzpatrick (1997) also applied a regression model to forecast the intensity of typhoon. In addition, Carter et al. (1989)

applied regression prediction equations to an usual weather forecast, and then the Central Weather Bureau (CWB) of Taiwan adopted the same scheme to establish a statistical forecast system to predict the daily high and low temperatures, cloud cover, and rainfalls (Chen et al. 2000). Other issues regarding the rainfall estimation have been widely studied and the readers may consult Unal et al. (2004), Xu and Tung (2009), Lee et al. (2010), and Haddad et al. (2010) for more details. For predicting the rainfall brought by typhoons, Yeh et al. (2001) and Yeh (2002) proposed the forecast models of rainfall based on a linear regression model and the empirical orthogonal function modes, respectively. In addition, Fan and Lee (2007) further combined the Bayesian technique to develop a mixture model for predicting the typhoon rainfall at Taipei. However, none of these studies accommodate the spatial dependence in the data, resulting in the estimation and prediction that may be inaccurate. Hoeting et al. (2006) also indicated that the importance of particular explanatory variables may not be significant when spatial correlation is ignored. Therefore, incorporating spatial dependence into models to assess the spatial risk variations of rainfall is essential.

Spatial statistics arises when the data are points in some Euclidean space and has been widely developed in many fields of study, such as geography, biology, and epidemiology. In the past, researchers mostly focused on modeling the spatial variations in disease risk (e.g., Kelsall and Wakefield 2002; Stern and Cressie 2002). However, how to assess the spatial risk variations for the natural disaster caused by earthquake or typhoon is also an important problem, but it has not been extensively studied in the literatures. Chen and Yang (2011) proposed a joint spatial modeling approach to assess the spatial variations in earthquake risks and obtained much useful information. To our knowledge, there does not have a suitable spatial model that can be satisfactorily applied to the problems of typhoon cumulated rainfall. It is known that most of the disasters caused by typhoons in Taiwan are due to heavy rain, and therefore we focus on assessing the spatial risk variations of typhoon cumulated rainfall at Taipei with respect to typhoon locations in this paper. Motivated from Chen and Yang (2011), we use the spatial hierarchical Bayesian model combined with the spatial conditional autoregressive (CAR) model to account for the spatial dependence among the risks, where a family of stochastic algorithms based on a Markov chain Monte Carlo (MCMC) technique is designed to estimate the model parameters. Based on the framework, some important factors influencing the typhoon cumulated rainfall at Taipei are also discussed.

The rest of this paper is organized as follows. Section 2 introduces the proposed spatial models for assessing the spatial risk variations of typhoon cumulated rainfall at Taipei with respect to typhoon locations and then displays the procedures of estimating the model parameters. Section 3 applies the proposed method to a real data set recorded by Taiwan's CWB from 1961 to 1994. Finally, a brief discussion is given in Sect. 4.

2 Modeling

Assume that $\mathbf{A} \subset \mathbb{R}^2$ is the study region of interest. We partition $\mathbf{A}$ into $N = n_1 \times n_2$ non-overlapping regular grids and denote them as A_i; $i = 1, 2, \ldots, N$. In this paper, we are interested in investigating the impact of typhoon locations with respect to h-hour cumulated rainfall over M (in millimeter) at some rainfall station, where h and M are pre-specified values and a data set of h-hour cumulated rainfall at some rainfall station is required for each typhoon within $\mathbf{A}$. Thus, there may be several events (i.e., h-hour cumulated rainfall over M at some rainfall station) for the same typhoon. Let Y_i be a nonnegative integer random variable which counts the number of events when the locations of candidate typhoons are in grid A_i; $i = 1, 2, \ldots, N$. In addition, we define the corresponding expected number to be $E_M \equiv N^{-1} \sum_{j=1}^{N} Y_j$, where M is a pre-specified threshold value of typhoon cumulated rainfall. In general, the quantity $r_i = Y_i/E_M$ is usually used to estimate the relative risk of h-hour cumulated rainfall over M at some rainfall station when a typhoon passes through grid A_i, but it does not accommodate the spatial dependence in the data. It may result in less accurate estimation and large variation. Instead of using independent r_i, in the next section, we introduce the spatial hierarchical Bayesian model combined with the spatial CAR model to account for the spatial dependence among the risks.

2.1 Spatial statistical models

Suppose that R_i is the relative risk of h-hour cumulated rainfall over M at some rainfall station when typhoon locations are in grid A_i. We consider to model the random variables $\mathbf{Y} \equiv (Y_1, \ldots, Y_N)'$ to be Poisson distributions as follows

$$Y_i | R_i \sim Poi(R_i E_M); \quad i = 1, 2, \ldots, N, \tag{1}$$

where $R_i E_M$ is the intensity rate of the Poisson process. As suggested in Besag et al. (1991), we further model the relative risks as

$$\log(R_i) = \alpha + \mathbf{X}_i \boldsymbol{\beta} + \delta_i; \quad i = 1, 2, \ldots, N, \tag{2}$$

where $\mathbf{X}_i = (x_{i1}, \ldots, x_{ik})$ are grid-level covariates, α and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$ are regression parameters, and

$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)'$ are random effects with spatial correlations that reflect small-scale spatial variabilities. In other words, all the variations of $\log(R_i)$, apart from the small-scale fluctuations of spatial dependence, are absorbed mostly into the mean structure $\alpha + X_i\boldsymbol{\beta}$. For the spatial random component $\boldsymbol{\delta}$, it is often assumed to be a stationary process and follows a multivariate normal distribution as follows

$$\boldsymbol{\delta}|\sigma^2, \phi \sim N(0, \sigma^2 V(\phi)), \qquad (3)$$

where $V(\phi)$ is an $N \times N$ correlation matrix with the unknown parameter ϕ measuring the degree of spatial dependence among $\boldsymbol{\delta}$, and σ^2 is the variance of δ_i; $i = 1, \ldots, N$. Here, we further consider the CAR model for the process $\{\delta_i : i = 1, \ldots, N\}$, resulting in the covariance matrix $\sigma^2 V(\phi)$ of (3) that can be decomposed into $(I - \phi C)^{-1} M$, where I is an identity matrix, $C = (c_{ij})$ is a pre-specified $N \times N$ spatial association matrix with $c_{ii} = 0$ and $c_{ij} = c_{ji}$, and $M = \sigma^2 I$. Note that $(I - \phi C)$ is invertible and $(I - \phi C)^{-1} M$ is symmetric and positive definite when $\phi \in (\phi_{min}, \phi_{max})$, where ϕ_{min} and ϕ_{max} are determined by the reciprocals of the smallest and the largest eigenvalues of C. There are many spatial correlation

depends on the value of ϕ. Applying the Factorization Theorem of Besag (1974) and the properties of a multivariate normal distribution (Cressie 1993, p. 413), the joint distribution of $\delta_i|\sigma^2, \phi, \boldsymbol{\delta}_{-i}$; $i = 1, \ldots, N$ is a multivariate normal distribution of (3). More details regarding the CAR model can refer to Besag (1974) and Cressie (1993).

2.2 Prior and posterior distributions

Under the Bayesian framework, the prior specification and derivation of posterior distribution both are required. First, some mutually independent and non-informative or conjugate priors for σ^2, ϕ, α, and $\boldsymbol{\beta}$ are given as follows: (i) the usual non-informative prior is considered for α and $\boldsymbol{\beta}$; (ii) an inverse gamma prior $IG(a, b)$ is assumed for σ^2, where constants a and b are pre-specified values so that the prior has variance as large as possible; (iii) we assume an uniform prior $U(0, \phi_{max})$ for the spatial correlation parameter ϕ because a negative value for ϕ seems unlikely in most applications. Then, applying the Bayes' theorem, the joint posterior distribution of $(\sigma^2, \phi, \alpha, \boldsymbol{\beta}, \boldsymbol{\delta})$ conditioned on Y satisfies

$$
\begin{aligned}
p(\sigma^2, \phi, \alpha, \boldsymbol{\beta}, \boldsymbol{\delta}|Y) &\propto \prod_{i=1}^{N} p(Y_i|R_i) p(\boldsymbol{\delta}|\sigma^2, \phi) \pi(\sigma^2) \pi(\phi) \pi(\alpha) \pi(\boldsymbol{\beta}) \\
&\propto \exp\left(\sum_{i=1}^{N} Y_i(\alpha + X_i\boldsymbol{\beta} + \delta_i) - E_M \exp(\alpha + X_i\boldsymbol{\beta} + \delta_i) \right) \\
&\quad \times (\det(\sigma^2 V(\phi)))^{-1/2} \exp\left(-\frac{1}{2\sigma^2} \boldsymbol{\delta}' V(\phi)^{-1} \boldsymbol{\delta} \right) \quad \times \sigma^{-2(a+1)} \exp\left(-\frac{1}{b\sigma^2} \right),
\end{aligned}
\qquad (4)
$$

functions that can be considered to construct the spatial association matrix C. However, it would suffer a problem regarding the selection of covariance functions but is beyond the scope of this study. We will briefly discuss the problem in Sect. 4 For simplicity in this study, we defined the (i, j)th entry of C to be $I(i \sim j)$, where $I(\cdot)$ is an indicator function and $i \sim j$ represents that grids A_i and A_j are neighbors. The definition is typically called the rook contiguity and $I(i \sim i) \equiv 0$. Thus, the conditional distribution of δ_i conditioned on $\boldsymbol{\delta}_{-i}$ is given by

$$\delta_i|\sigma^2, \phi, \boldsymbol{\delta}_{-i} \sim N\left(\phi \sum_{j \in N_i} c_{ij}\delta_j, \sigma^2 \right); \quad i = 1, 2, \ldots, N,$$

where $\boldsymbol{\delta}_{-i}$ is a vector with the ith element deleted from $\boldsymbol{\delta}$ and $N_i \equiv \{j : j \sim i\}$ is the neighborhood set of grid A_i. Obviously, the spatial dependence is found based on the information of neighbors and the degree of dependence

where $\pi(\cdot)$ represents a given prior distribution. Because the joint posterior distribution of (4) can not be summarized analytically, we design a family of stochastic algorithms based on MCMC techniques to estimate the model parameters in the next subsection.

2.3 Estimation of parameters

Due to the conjugate prior, we obtain that the conditional posterior distribution of σ^2 conditioned on all other variables follows an inverse gamma distribution as follows

$$\sigma^2|\phi, \boldsymbol{\delta} \sim IG\left(\frac{N}{2} + a, \left(\frac{1}{2} \boldsymbol{\delta}'(I - \phi C)\boldsymbol{\delta} + \frac{1}{b} \right)^{-1} \right). \qquad (5)$$

Therefore, we can apply the Gibbs sampler technique (e.g., Geman and Geman 1984) to generate the posterior samples of σ^2, where the stochastic algorithm successively samples

from the conditional posterior distribution of (5) and results in a Markov chain that converges to the joint posterior distribution given in (4) under mild conditions (Tierney 1994).

For α, $\boldsymbol{\beta}$, and δ_i; $i = 1, \ldots, N$, the corresponding conditional posterior distributions are all non-standard. We summarize them as follows

$$p\left(\alpha|\sigma^2, \phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{Y}\right) \propto \prod_{i=1}^{N} p(Y_i|R_i)\pi(\alpha) \propto \exp\left(\sum_{i=1}^{N} \alpha Y_i - E_M \exp(\alpha + \boldsymbol{X}_i\boldsymbol{\beta} + \delta_i)\right),$$

$$p\left(\boldsymbol{\beta}|\sigma^2, \phi, \alpha, \boldsymbol{\delta}, \boldsymbol{Y}\right) \propto \prod_{i=1}^{N} p(Y_i|R_i)\pi(\boldsymbol{\beta}) \propto \exp\left(\sum_{i=1}^{N} Y_i\boldsymbol{X}_i\boldsymbol{\beta} - E_M \exp(\alpha + \boldsymbol{X}_i\boldsymbol{\beta} + \delta_i)\right),$$

$$p\left(\delta_i|\sigma^2, \phi, \alpha, \boldsymbol{\beta}, \boldsymbol{\delta}_{-i}, \boldsymbol{Y}\right) \propto p(Y_i|R_i)p(\delta_i|\sigma^2, \phi, \boldsymbol{\delta}_{-i}) \propto \exp\left(Y_i\delta_i - E_M \exp(\alpha + \boldsymbol{X}_i\boldsymbol{\beta} + \delta_i) - \frac{1}{2\sigma^2}\left(\delta_i^2 - 2\delta_i\phi\sum_{j\in N_i} c_{ij}\delta_j\right)\right).$$

For each variable, a Metropolis–Hastings algorithm is applied to iteratively generate an ergodic Markov chain that yields the posterior samples (e.g. Metropolis et al. 1953; Hastings 1970; Chib and Greenberg 1995). In each step, an update of the current state of the chain is generated from a proposal distribution and the update is then accepted or rejected according to a certain acceptance probability. In practice, a most commonly used algorithm is Gaussian random-walk Metropolis, where the proposal distribution is Gaussian with mean being equal to the current state.

Finally, the conditional posterior distribution of the spatial dependent parameter ϕ is also non-standard but has the following proportional form

$$p\left(\phi|\sigma^2, \alpha, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{Y}\right) \propto p(\boldsymbol{\delta}|\sigma^2, \phi)\pi(\phi)$$
$$\propto (\det(\boldsymbol{V}(\phi)))^{-1/2}\exp\left(\frac{\phi}{2\sigma^2}\boldsymbol{\delta}'\boldsymbol{C}\boldsymbol{\delta}\right).$$

In this paper, a discrete method is considered to generate ϕ due to its computational simplicity, where matrix $\boldsymbol{V}(\phi)$ can be computed in advance on fine grid points of ϕ. For each step, the probability mass function of ϕ on find grid points can be evaluated and then the posterior sample of ϕ is generated.

3 Analysis on the typhoon cumulated rainfall

In this section, we carry out models (1)–(3) to a real data set of typhoon cumulated rainfall accompanied with some covariates for evaluating the spatial risk variations of typhoon cumulated rainfall in Taipei area with respect to typhoon locations.

3.1 Description of data

A data set of h-hour typhoon cumulated rainfall at Taipei rainfall station with respect to typhoon locations was supplied by Taiwan's CWB from 1961 to 1994, which contains 145 typhoons and 9595 data regarding the typhoon cumulated rainfalls at Taipei rainfall station. Note that the original data were collected by Joint Typhoon Warning Center of USA based on 6-h best tracks data of typhoons, and then Taiwan's CWB adopted the linear interpolation technique to obtain the hourly typhoon center position and relative information. As mentioned in Yeh et al. (1999, 2001), Taipei area is easier to have heavy rains when typhoons pass through $A = [120°\text{E}, 125°\text{E}] \times [21.9°\text{N}, 25.5°\text{N}]$. Thus, in this paper we let A be our study region and we partition it into $N = 15(= 5 \times 3)$ non-overlapping regular grids according to their suggestion. It results in 68 typhoons that appear within A. Figure 1 shows the study region A and 15 non-overlapping regular grids. In this study, we use three threshold values of 24-h typhoon cumulated rainfalls, say, $M = 50$, 100, and 130, to respectively assess the spatial risk variations of 24-h
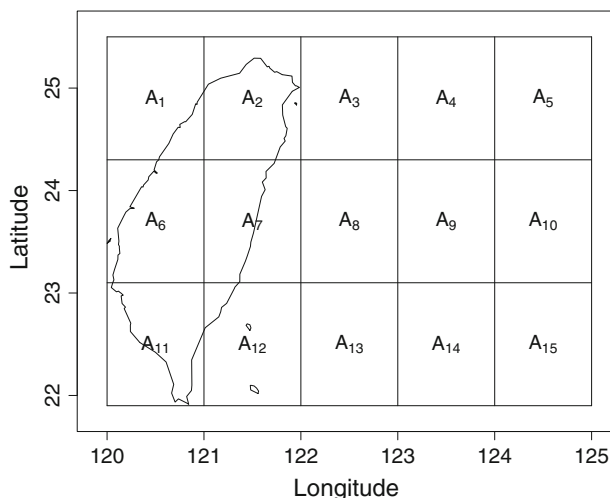

Fig. 1 The study region A and 15 non-overlapping regular grids

Table 1 Three test cases

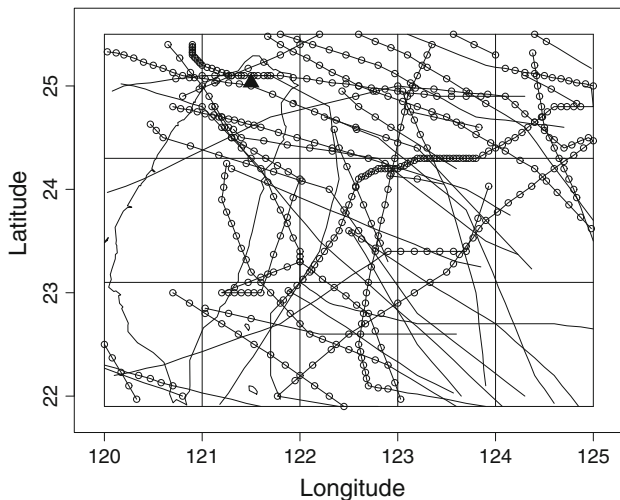| M	Number of typhoons	Events	E_M
50	38	486	32.4
100	17	177	11.8
130	9	87	5.8

Fig. 2 The paths of 38 typhoons and the scatter plots of 486 events for $M = 50$, where the triangle represents Taipei rainfall station

Fig. 3 The paths of 17 typhoons and the scatter plots of 177 events for $M = 100$, where the triangle represents Taipei rainfall station
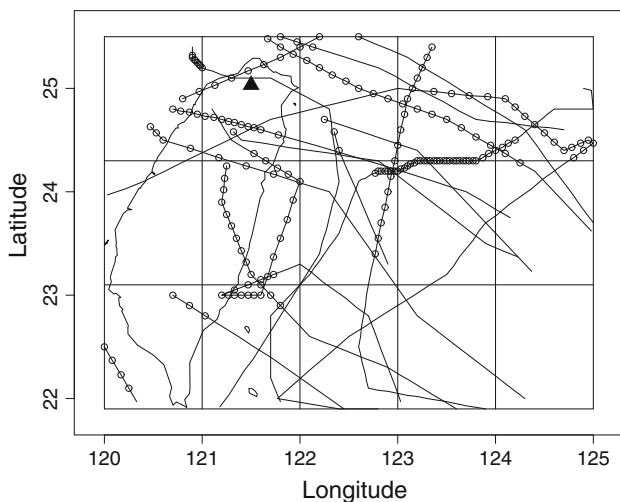

Fig. 4 The paths of 9 typhoons and the scatter plots of 87 events for $M = 130$, where the triangle represents Taipei rainfall station

typhoon cumulated rainfall at Taipei with respect to typhoon locations. According to the definition of Taiwan's CWB, $M = 50$ and $M = 130$ are called heavy rain and extremely heavy rain, respectively. We summary the three cases in Table 1 and Figs. 2, 3, 4 display the paths of typhoons and the scatter plots of events with circle marks (i.e., the typhoon locations where 24-h cumulated rainfalls
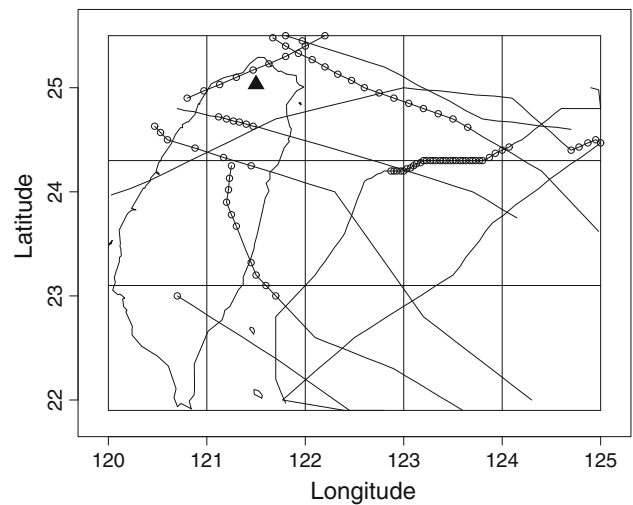
of Taipei rainfall station are over M) for the three cases, where Taipei rainfall station locates in A_2 and is marked as a triangle.

In addition to the typhoon cumulated rainfall at Taipei rainfall station, 7 corresponding explanatory variables regarding the characteristics of the Taipei rainfall station and typhoons were also recorded. Note that these explanatory variables were the same as Yeh et al. (2001) and Fan and Lee (2007) who used the linear regression model and mixture models to predict typhoon cumulated rainfalls at Taipei area, respectively. We summarized 7 explanatory variables as follows:

X_1: Maximum wind speed of typhoon center
X_2: Distance between the Taipei rainfall station (x_s, y_s) and the typhoon center (x, y)

$$X_2 = \sqrt{((x - x_s) \times \cos 20°)^2 + (y - y_s)^2}$$

X_3: Moving direction angle of typhoon

$$X_3 = \frac{360}{2\pi} \times \arctan\left(\frac{y_q - y_p}{x_q - x_p}\right), \text{ where } (x_p, y_p) \text{ and } (x_q, y_q)$$

are locations of a typhoon at time t and time $t + \Delta t$, respectively, and $\arctan\left(\frac{y_q - y_p}{x_q - x_p}\right)$ is the corresponding direction radian.

X_4: Moving speed of typhoon

$$X_4 = \sqrt{((x_q - x_p) \times \cos 20°)^2 + (y_q - y_p)^2} / \Delta t$$

X_5: Surface pressure of Taipei rainfall station
X_6: Wind speed of Taipei rainfall station
X_7: Transformation of the Julian day

$$X_7 = \sin\left((X_7' - 41) \times \frac{\pi}{364.75}\right), \text{ where } X_7' \text{ is the Julian}$$

day of a typhoon at that time.

Note that the transformation formula of X_7 is based on Neumann (1992) and its maximum value occurs at the

Julian date X_7' within July–September. In Taiwan, typhoon usually happens in this period. Therefore, if the variable X_7 is significant with a positive regression coefficient, it indicates that a typhoon happens in the period will bring heavy rainfalls for the Taipei area. In addition to variable X_7, we can divide variables X_1–X_6 into two groups. The first group consists of variables X_1–X_4 which are associated with the characteristics of typhoons and another group includes X_5 and X_6 which show the status of the Taipei rainfall station.

3.2 Computations and results

For $M = 50$, 100, and 130, we apply models (1)–(3) with the spatial association matrix C given in Sect. 2.1 and consider $X_i = (x_{i1}, \ldots, x_{i7}); i = 1, \ldots, 15$, as grid-level covariates in model (2). That is, if there are n_i events within grid $A_i; i = 1, \ldots, 15$, and each event has 7 explanatory variables, then each component of X_i can be obtained by taking average over the corresponding n_i values. In addition, for the settings of priors, we take $\sigma^2 \sim IG(3, 0.001)$, $\phi \sim U(0, \phi_{max})$, and $U(-20, 20)$ for α and $\beta_j; j = 1, \ldots, 7$, where $\phi_{max} = 0.32$ is obtained by taking the reciprocal of the smallest eigenvalue of C.

For each of $M = 50$, 100, and 130, we run 200,000 iterations for the posterior computations and discard the first 100,000 iterations as a burn-in. We retain every 10th set of parameter values to obtain approximately independent joint posterior samples of size 10,000, where convergence is assessed by examination of trace and autocorrelation plots. In addition, some theoretical results indicated that the proposal variances in Metropolis–Hastings steps should be tuned to achieve acceptance rates around 0.23 for random walk updates; see Roberts et al. (1997). In our study, the average acceptance rates for updating spatial random effects δ and regression parameters are near 0.22 and 0.20, respectively. Tables 2, 3, and 4 summarize the posterior inferences of model parameters for $M = 50$, 100, and 130, respectively. Again, parameter ϕ measures the degree of the spatial dependence among the risks. As expected, the posterior mean or median of ϕ for each case is distant from zero, it indicates that the spatial correlation exists among the risks. Therefore, it can not be ignored in analyzing the spatial risk variation of typhoon cumulated rainfalls. Based on the results of Tables 2, 3, and 4, we summarize important explanatory variables for each case in Table 5. In general, we notice that $M = 50$ and $M = 100$ have near important explanatory variables but are very different to the case of $M = 130$. For heavy rains (e.g., $M = 50$ or $M = 100$), the results of Table 5 indicate the following four facts (i)–(iv) correspond to variables X_1, X_2, X_4, and X_7, respectively: (i) if the wind speed of typhoon center is large, Taipei area will have a large risk to suffer the 24-h typhoon cumulated rainfall over M; (ii) when a typhoon approaches the Taipei rainfall station, Taipei area will have a large risk to suffer the 24-h typhoon cumulated rainfall over M; (iii) a typhoon with a slow moving speed will result in a large risk for the Taipei area to suffer 24-h typhoon cumulated rainfall over M; (iv) a typhoon happens within July–September will bring heavy rainfalls for the Taipei area. However, for the case of extremely heavy rain (i.e., $M = 130$), the results indicate that the moving direction angle of typhoons (X_3) and the surface pressure of Taipei rainfall station (X_5) are major reasons to cause the Taipei area to have the 24-h typhoon cumulated rainfall over 130. Importantly, it is very sensible that the moving speed of typhoon (X_4) is always significant to influence the cumulated rainfalls for all cases. In addition, as mentioned in Sect. 2.1, all the variations among the risks, apart from the small-scale fluctuations of spatial dependence, are absorbed mostly into the mean structure. However, the estimate of σ^2 in model (3) is large for each case. It indicates that there may be other important explanatory variables that have not been discovered in analyzing the typhoon cumulated rainfall of Taipei area.

Finally, the estimated relative risks of Taipei area experiencing the 24-h typhoon cumulated rainfall over 50, 100, and 130 when the typhoon locations are within grid $A_i; i = 1, \ldots, 15$, are shown in Figs. 5, 6, and 7, respectively. For all three cases, we notice that there are smaller relative risks for Taipei area when the path of typhoon appears within grids A_6, A_{11} and A_{12}. This is because the counterclockwise rotation of typhoon usually results in its structure being destroyed by the Central Mountain Range of Taiwan. In addition, we also notice that there is a higher relative risk for Taipei area when a typhoon is a west-forward type.

To sum up, the methodology presented here will help us in assessing the spatial risk variations of the typhoon cumulated rainfall at Taipei area with respect to typhoon locations and hence we can take necessary precautions against typhoon.

4 Discussion and conclusion

In this paper, we developed a spatial hierarchical Bayesian model combined with a spatial conditional autoregressive model to assess the spatial risk variations in the typhoon cumulated rainfalls for the Taipei area with respect to typhoon locations. The proposed method can be easily implemented through the MCMC algorithms. In this study, we focus on assessing the spatial risk variations of typhoon cumulated rainfalls for the Taipei area as the results shown

Table 2 Summary of posterior inferences for the model parameters when $M = 50$

Parameter	2.5%	5%	Median	95%	97.5%	Mean	SD
σ^2	71.453	77.877	129.427	233.883	264.537	139.174	49.978
ϕ	0.031	0.031	0.124	0.279	0.279	0.142	0.078
α	−10.714	−10.333	−4.951	−0.430	−0.127	−5.028	3.147
β_1	0.245	0.260	0.426	0.622	0.658	0.428	0.106
β_2	−11.892	−11.652	−9.169	−4.874	−4.215	−8.779	2.245
β_3	−0.105	−0.101	−0.029	0.030	0.034	−0.036	0.042
β_4	−284.752	−271.065	−198.202	−133.901	−127.011	−196.787	38.976
β_5	−0.016	−0.013	0.003	0.018	0.022	0.004	0.009
β_6	−2.049	−1.896	−0.259	1.796	1.896	−0.104	1.311
β_7	30.020	31.882	63.539	105.833	110.919	67.749	26.377

Table 3 Summary of posterior inferences for the model parameters when $M = 100$

Parameter	2.5%	5%	Median	95%	97.5%	Mean	SD
σ^2	99.771	109.014	190.161	358.224	408.770	206.640	81.696
ϕ	0.031	0.031	0.155	0.279	0.310	0.161	0.082
α	−22.428	−21.645	−18.212	−12.317	−10.608	−17.733	2.827
β_1	0.102	0.119	0.319	0.599	0.647	0.337	0.151
β_2	−23.270	−21.402	−15.592	−11.885	−11.388	−15.833	2.703
β_3	−0.106	−0.098	0.034	0.071	0.079	0.015	0.051
β_4	−261.037	−238.434	−144.798	−18.859	−11.375	−141.287	72.263
β_5	−0.008	−0.005	0.021	0.054	0.055	0.027	0.019
β_6	−5.307	−5.193	−1.939	−0.761	−0.546	−2.532	1.508
β_7	28.386	33.469	116.531	128.968	130.109	98.049	32.404

Table 4 Summary of posterior inferences for the model parameters when $M = 130$

Parameter	2.5%	5%	Median	95%	97.5%	Mean	SD
σ^2	96.317	107.390	195.945	393.531	453.455	215.850	93.539
ϕ	0.031	0.031	0.155	0.310	0.310	0.162	0.084
α	−21.799	−18.909	−13.155	−7.103	−6.188	−13.140	3.730
β_1	−0.223	−0.107	0.197	0.606	0.631	0.218	0.235
β_2	−8.011	−7.081	0.075	8.514	9.396	0.588	4.650
β_3	0.110	0.120	0.205	0.292	0.301	0.202	0.052
β_4	−213.495	−206.087	−136.767	−68.775	−54.257	−136.380	43.512
β_5	0.014	0.016	0.036	0.063	0.065	0.038	0.015
β_6	−4.869	−4.673	−2.771	0.291	0.696	−2.493	1.506
β_7	−9.608	−7.892	11.797	32.098	34.346	11.413	12.349

Table 5 Significant explanatory variables for the three test cases, where "+" and "−" represent positive correlation and negative correlation, respectively

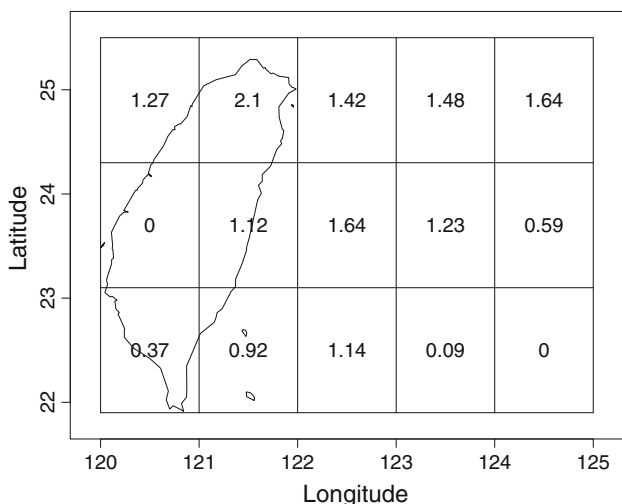| M	X_1	X_2	X_3	X_4	X_5	X_6	X_7
50	+	−		−			+
100	+	−		−		−	+
130			+	−	+		

Fig. 5 Posterior means of $R_i = \exp(\alpha + X_i\beta + \delta_i)$; $i = 1, \ldots, 15$, for $M = 50$
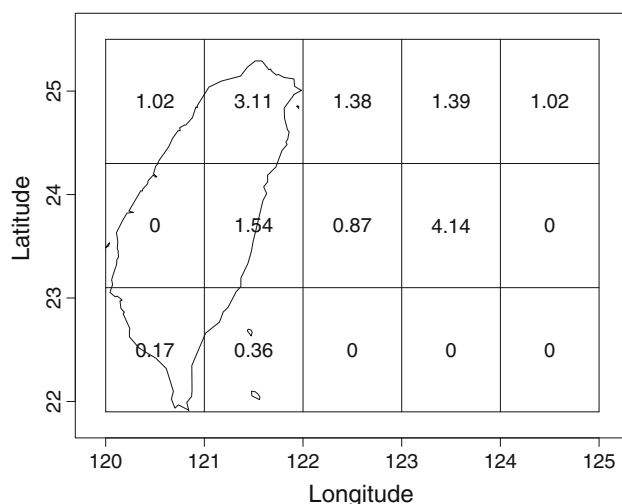

Fig. 7 Posterior means of $R_i = \exp(\alpha + X_i\beta + \delta_i)$; $i = 1, \ldots, 15$, for $M = 130$
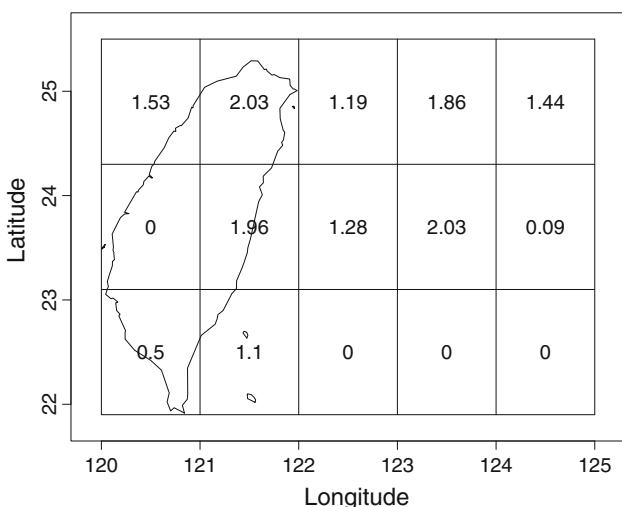

Fig. 6 Posterior means of $R_i = \exp(\alpha + X_i\beta + \delta_i)$; $i = 1, \ldots, 15$, for $M = 100$

in Figs. 5, 6, and 7. The risk trends of the three cases are basically the same, but we notice that there are different important variables that influence the relative risks of Taipei area experiencing heavy rains when different threshold values M are considered. These results would help the Taiwan's CWB to forecast the typhoon cumulated rainfall of Taipei area within 24-h based on the typhoon locations and the changes of important explanatory variables. Therefore, we can take necessary precautions against typhoon.

In our study, we recommend the spatial random effects δ in the model to reflect the spatial correlations among the risks, in which parameter ϕ measures the degree of spatial dependence of δ. As demonstrated in Tables 2, 3, and 4, the estimate of ϕ is distant from zero for each case. It

indicates that the spatial correlation exists among the risks and hence can not be ignored. In addition, Hoeting et al. (2006) also pointed out that the important explanatory variables may not be significant when spatial correlation is ignored in the modeling procedure. Therefore, incorporating spatial dependence into models to assess the spatial risk variations in typhoon cumulated rainfall is essential. For the spatial association matrix C, we use a simple structure, called the rook contiguity, into our model. Although it may not be the best one, the results of data analysis indicate some useful and significant information for us. Undoubtedly, other spatial correlation functions could also be considered, but it would suffer a problem regarding the selection of covariance functions. For the problem, a fair selection criterion of covariance functions should be proposed, but it is beyond the scope of this study.

Furthermore, we could apply the same technique to analyze spatial risk variations of typhoon cumulated rainfalls for other rainfall stations of Taiwan with respect to typhoon locations. The results would provide some useful information for the prevention of possible damages from typhoons.

References

Besag J (1974) Spatial interaction and the statistical analysis of lattice systems. J R Stat Soc B 36:192–225

Besag J, York J, Mollié A (1991) Bayesian image restoration, with two applications in spatial statistics (with discussion). Ann Inst Stat Math 43:1–59

Carter GM, Dallavale JP, Glahn HR (1989) Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. Weather Forecast 4:401–412

Chen CS, Yang HD (2011) A joint modeling approach for spatial earthquake risk variations. J Appl Stat 38:1733–1741

Chen CG, Luo CW, Wang HM, He JG (2000) The development of the statistical forecast system in the Central Weather Bureau. Meteorol Bull 43:18–33

Chib S, Greenberg E (1995) Understanding the Metropolis Hastings algorithm. Am Stat J 49:327–335

Cressie N (1993) Statistics for spatial data (revised edition). Wiley, New York

DeMaria M, Kaplan J (1994) A statistical hurricane intensity prediction scheme for the atlantic basin. Weather Forecast 9:209–220

Fan TH, Lee YH (2007) A bayesian mixture model with application to typhoon rainfall predictions in Taipei. Int J Contemp Math Sci 2:639–648

Fitzpatrick PJ (1997) Understanding and forecasting tropical cyclone intensity change with the tropical intensity prediction scheme. Weather Forecast 12:826–846

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741

Haddad K, Rahman A, Green J (2010) Design rainfall estimation in Australia: a case study using L moments and generalized least squares regression. Stoch Environ Res Risk Assess. doi:10.1007/s00477-010-0443-7

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57:97–109

Hoeting JA, Davis RA, Merton AA, Thompson SE (2006) Model selection for geostatistical models. Ecol Appl 16:87–98

Kelsalll J, Wakefield J (2002) Modeling spatial variation in disease risk: a geostatistical approach. J Am Stat Assoc 97:692–701

Lee CH, Kim T, Chung G, Choi M, Yoo C (2010) Application of bivariate frequency analysis to the derivation of rainfall-frequency curves. Stoch Environ Res Risk Assess 24:389–397

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. J Chem Phys 21:1087–1092

Neumann CJ (1972) An alternate to the HURRAN tropical cyclone forecast system. NOAA Tech Mem NWS, SR-62, p 25

Neumann CJ (1992) A revised climatology and persistence model (WPCLPR) for the prediction of Western North Pacific tropical cyclone motion. SAIC/NORAL Contract Report N00014-90-C-6042 (PART 1), p 40

Roberts GO, Gelman A, Gilks WR (1997) Weak convergence and optimal scaling of random walk Metropolis algorithms. Ann Appl Probab 7:110–120

Stern HS, Cressie N (2002) Posterior predictive model checks for disease mapping models. Stat Med 19:2377–2397

Tierney L (1994) Markov chains for exploring posterior distributions (with discussion). Ann Stat 22:1701–1762

Unal NE, Aksoy H, Akar T (2004) Annual and monthly rainfall data generation schemes. Stoch Environ Res Risk Assess 18:245–257

Xu Y, Neumann CJ (1985) A statistical model for the prediction of Western North Pacific tropical cyclone motion. NOAA Tech Mem NWS NHC-28, p 30

Xu YP, Tung YK (2009) Constrained scaling approach for design rainfall estimation. Stoch Environ Res Risk Assess 23:697–705

Yeh TC (2002) Typhoon rainfall over Taiwan area: the empirical orthogonal function modes and their applications on the rainfall forecasting. Terr Atmos Ocean Sci 13:449–468

Yeh TC, Wu SC, Shieh SL (1999) A study of typhoon rainfall statistics forecast over Taiwan area part I: spatial distribution of the forecasts. Atmos Sci 27:395–412

Yeh TC, Fan TH, Lee YH (2001) Typhoon rainfall regression predictions over Taiwan area (I). The linear regression model for predicting typhoon rainfalls at Taipei. Atmos Sci 29:77–96