Hindawi
Advances in Meteorology
Volume 2018, Article ID 5479576, 15 pages
https://doi.org/10.1155/2018/5479576

Research Article

Spatiotemporal Exploration and Hazard Mapping of Tropical Cyclones along the Coastline of China

Shaobo Zhong, Chaolin Wang, Zhichen Yu, Yongsheng Yang, and Quanyi Huang

Department of Engineering Physics, Institute of Public Safety Research, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Shaobo Zhong; zhongshaobo@tsinghua.edu.cn

Received 24 October 2017; Accepted 15 April 2018; Published 6 June 2018

Academic Editor: Anthony R. Lupo

Copyright © 2018 Shaobo Zhong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Spatiotemporal patterns are one of the greatest interests and provide valuable insights into chronological events occurring in space. A tropical cyclone (TC) track is defined as a sequence of successive points, and several different types of analyses are performed to explore the temporal and spatial patterns of the TCs in the Northwest Pacific and along the coastline of China during 1949–2014. Results show that (1) the number of TCs is getting more frequent from April to August and less frequent from August to October with the peak occurring in August almost every year; (2) the mean of the sizes of the annual temporal clusters during 1949–2014 is 52.5 (days), the standard deviation is 17.0 (days), and the average starting point is the 210.5th day; (3) the spatial clusters are located in two areas: the boundary of Guangxi and Guangdong provinces and the boundary of Fujian and Zhejiang provinces; and (4) the within-strata variance is less than the between-strata variance, which implies the locational and seasonal factors are the potential determinants of the heterogeneity of the TCs. Furthermore, several maps representing the hazards of TCs are produced. According to the resultant maps, 12 coastal prefectures (Zhanjiang, Maoming, Fuzhou, Huizhou, Yangjiang, Qinzhou, Ningde, Quanzhou, Jiangmen, Nanning, Zhangzhou, and Hangzhou) have return periods of less than two years, and the two island provinces of Hainan and Taiwan are visited by TCs the most. Guangdong, Guangxi, Fujian, and Zhejiang provinces in particular suffered severely from the destructive TCs.

1. Introduction

A tropical cyclone (TC) is a rapidly rotating storm system characterized by a low-pressure center, strong winds, and a spiral arrangement of thunderstorms that produce heavy rain. Depending on the strength and the location where the storm occurs, tropical cyclones have some synonyms such as tropical storm, hurricane, typhoon, tropical depression, cyclonic storm, and simply a cyclone [1]. An intense tropical cyclone, which brings gales, heavy rain, and storm surge in the course of its landfall, is among the most devastating of all natural hazards and causes huge losses every year all over the world, especially in some coastal areas [2]. China is located in the West Pacific and has a length of continental coastline more than 18,000 km in the east and south parts of the country. There are numerous cities (including Tianjin, Shanghai, Guangzhou, Shenzhen, etc., metropolises) distributed in 14 coastal provinces (autonomous regions, municipalities, and special administrative regions) (12 continental ones and 2 island ones). Due to suitable natural and climate conditions, more than 600 million people live in these coastal provinces, which amount to 43% of the total population of the country. In terms of the national statistics yearbook of 2014, the accumulated GDP (Gross Domestic Product) of the 10 continental coastal provinces (data are unavailable for Hongkong and Macau) are 36,952 billion RMB (Ren Min Bi) Yuan which is even more than 54% of the total GDP of Mainland China. The nation standard of China "Grade of tropical cyclones" that was put in practice since June 15, 2006 divides tropical cyclones into six scales: TD (tropical depression), TS (tropical storm), STS (severe tropical storm), TY (typhoon), STY (severe typhoon), and SuperTY (super typhoon). Although the fatality of tropical cyclones has been significantly reduced by a highly successful program of warnings and advanced building construction over the last decades, economic losses are escalating rapidly owing to the accelerated construction in TC-prone areas. The main affected areas in China by tropical cyclones are concentrated in coastal cities where population density and

economic level are universally higher than other areas. These areas suffer huge losses from typhoons every year. 16 (including 10 coastal and 6 inland) out of 28 provinces in China are affected by TC-induced disasters [3]. In recent decades, the number of casualties caused by hazardous TCs in China has slightly decreased, but the property loss has significantly increased. Since the year of 2000, direct and indirect economic loss from hazardous TCs exceeds tens of billion RMB every year. Densely populated and well-developed coastal areas have suffered high casualty and property loss.

Analysis of spatiotemporal characteristics for TCs is useful to recognize distributional patterns, identify clusters, and predict spatial variables based on historical data, which is able to provide useful support for prevention and response of hazardous TCs, including comprehensive risk assessment, evacuation planning, resource allocation, among others. For example, Wang et al. [4] show the interdecadal variation in TC tracks over the Northwest Pacific taking landfalling locations of TCs in Xiamen as study objects. The results from Gu et al. [5] suggest that since the early 1960s, there has been an overall decreasing trend in the frequency of occurrence, intensity, peak intensity, length of movement, and lifetime of TCs. Chan and Liu [6] explored interannual variations of typhoon activity and concluded these variations appear to be largely constrained by the large-scale atmospheric factors that are closely related to the El Niño—Southern Oscillation (ENSO) phenomenon. Tang et al. [7] analyzed the annual and monthly variations of TC from 1951 to 2006 statistically, including frequencies, intensities, and wind intensity indices. By calculating the TC frequency and wind intensity indices in each $1^\circ \times 1^\circ$ longitude-latitude grid, Tang et al. [7] also analyzed the spatial distribution and the influence extent of TCs. In summary, in spatiotemporal characteristics of TCs, currently existing literature are mainly focused on summary statistics on different scales. Though some temporal and spatial patterns are presented in these papers, it is necessary to get deep insight into spatiotemporal characteristics of TCs with advanced techniques of spatial pattern analysis. Temporal analysis helps ones effectively get insights into chronological phenomena through establishing references of activity and discover periodic patterns in which data are presented by different scales (e.g., hour, day, month, and year). Temporal analyses are often shown in simple linear and circular charts to illustrate chronological order and cycle [8]. Different from the charts, some statistical methods can be used to detect clusters of events along a timeline and assert the confidence level of the results. Whether the occurrences of chronological events have significant anomalies along a timeline is essentially a problem how to identify the clusters of points along a line. There are some existing methods able to solve this problem [9–11]. These kinds of one-dimensional scan statistic methods are also widely applied in cluster detection of temporal data such as disease, drought, and forest fire. There are two kinds of commonly used techniques for exploring spatial patterns: global clustering and local clustering. Global clustering is to judge whether there exists significant spatial clustering trends in spatial phenomena (represented by observed events), which is generally determined by global test statistics [12, 13]. However, global clustering cannot obtain the specific locations of clusters. Local clustering is capable of identifying the centers of clusters and even their sizes and shapes [14–17]. Stratified heterogeneity, which refers to the within-strata variance less than the between-strata variance, is ubiquitous in spatial phenomena. Stratified heterogeneity may imply the determinants of events [18].

To visualize the hazard of disasters, an alternative approach is to evaluate the spatial density of disasters. While point density estimation is widely used to quantify hazard of disasters with concentric impact, a line-based density estimation is imperative to map the hazard of disasters with linear distribution. Density estimation techniques have been applied to examine spatiotemporal dynamics in some domains [19–24]. These types of techniques generally use a density kernel to spread the values of the samples out over a surface. The magnitude at each sample is distributed throughout the study area, and a density value is estimated for each cell in the output surface. Steiniger and Hunter [22] extend the point-based kernel density estimation (KDE) approach to work with sequential GPS-point tracks, the outcome of which is a line-based KDE. Demšar et al. [23] present an alternative geovisualization method for spatiotemporal aggregation of trajectories of tagged animals: stacked space-time densities. Given tracks of tropical cyclones are represented as lines, and in this study, we propose a track density algorithm to map the density of tropical cyclones.

Some GIS tools of mapping and analysis are widely used in manipulation of georeferenced data. Through adding a few geographic elements into the tornado definition and then characterizing tornado density as a density field in GIS, Deng et al. [25] examine how GIS factors function in the process of tornado density mapping. Liu et al. [26] resort to the extension of ArcGIS (a proprietary GIS software package developed by Esri Corporation): Geostatistical Analyst to carry out interpolation analysis of flood risk factors. Also, in those widely used proprietary GIS software packages like ArcGIS, some commonly used methods of spatial analysis and visualization have been devised. Moreover, almost all products also provide script languages to facilitate end-users to glue all kinds of functions (including built-in and customized ones) for domain-oriented problem solving. As will be seen, in our study, we will make full use of these kinds of capability devised by the proprietary GIS.

For purpose of discovering temporal and spatial characteristics of TCs along the coastline of China in the Northwest Pacific, following the exploratory data analysis, this paper mainly explores temporal and spatial clusters with advanced spatial techniques of spatiotemporal cluster analysis and density mapping coupled with GIS based on the TCs record data between 1949 and 2014. In the remainder of this paper, we first explore the temporal clusters of the TCs in the Northwest Pacific during 1949–2014 and the spatial clusters of the landfalling TCs along the coast of China to identify the clustering durations in these years and the clustering areas along the coastal prefectures, respectively. Following these explorations, the degree of temporal and spatial stratified heterogeneity is measured and its significance is tested. Then, the exceedance probabilities (EP) and

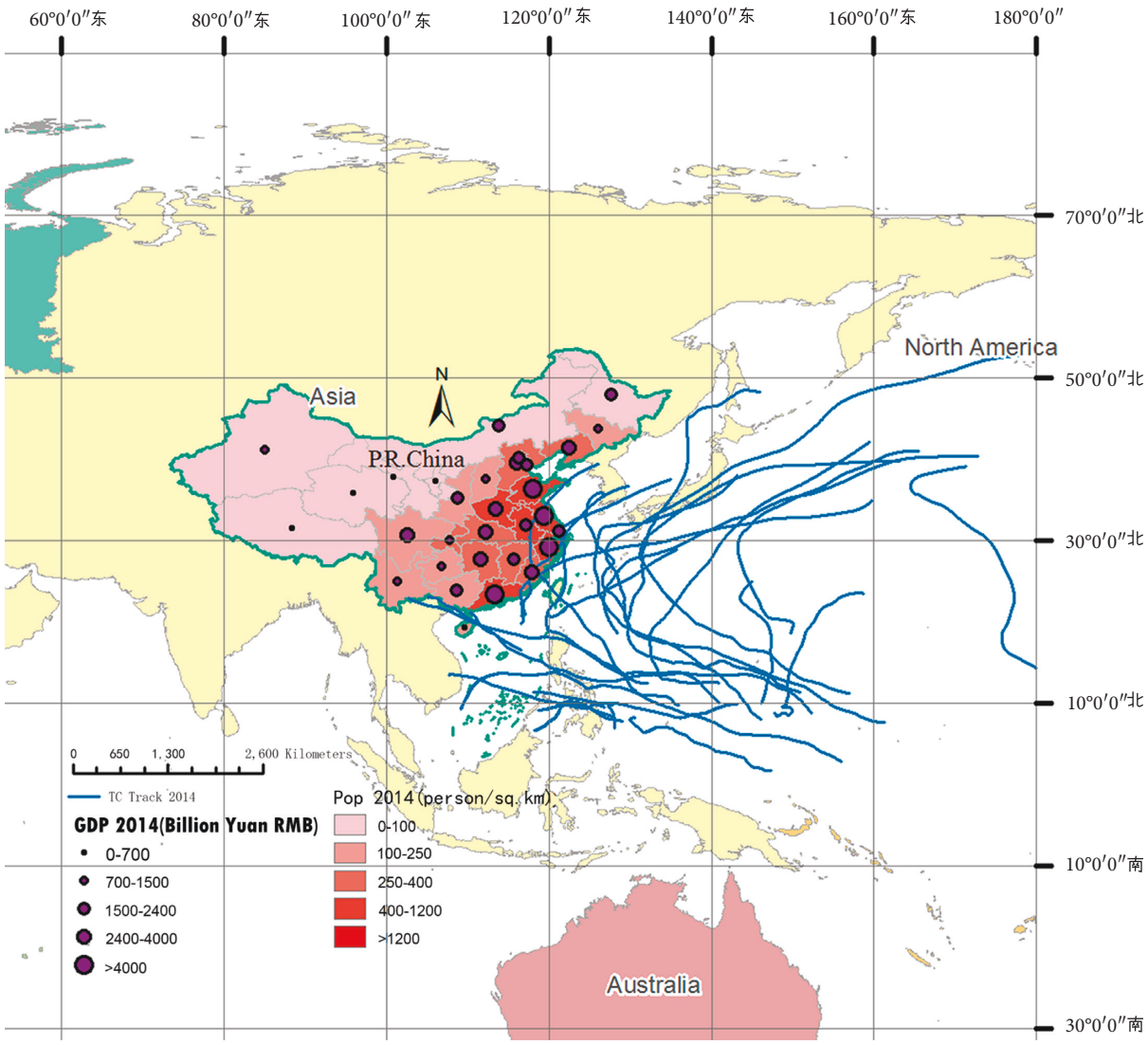


Figure 1: The administration map of China and its surrounding areas (the population density and GDP of Mainland China in 2014 are plotted) and tropical cyclones in the Northwest Pacific (here only show the tracks in 2014).

the return periods of the landfalling TCs in the coastal prefectures are rendered. Next, we map the hazards of the TCs in the Northwest Pacific. In the end, the paper summaries the main conclusions, proposes some problems to be solved, and suggests further research aspects.

2. Study Area and Data

2.1. Study Area. The Northwest Pacific is one of the international TC zones with an annually average number of 35 TCs, which amounts to 31% of all the hazardous TCs in the world, two times more than any other zone [27]. Moreover, about 80% of TCs in the Northwest Pacific develop into more severe ones (tropical storm or above). Figure 1 shows the administration map of China at the provincial level and its surrounding areas. For an intuitional recognization of the hazard of TCs, the population and GDP of 2014 are symbolized with graduated colors and graduated symbols, respectively, in Figure 1.

2.2. TC Track Data. The TC track data used in this study are obtained from the China Typhoon Web (http://www.typhoon.gov.cn). These data are published by Shanghai Typhoon Institute (STI) which is entrusted by the China Meteorological Administration (CMA) to compile the Best Track Dataset for Tropical Cyclones over the Northwest Pacific (CMA-STI Best Track Dataset for Tropical Cyclones over the Northwest Pacific) (http://www.typhoon.org.cn). The basin is to the north of the equator and to the west of 180°E, including the South China Sea, and the time span is from 1949 to 2014. In the data set, the data are organized track by track. Each TC track starts with a header and records a sequence of points (the location of the TC head at that time) every 6 hours. Each file in the data set contains the annual data, which is saved as a text format (with a .txt extension). For detailed information, please refer to the data set and its document.

In order to import and analyze the data in ArcGIS (a proprietary GIS software) [28], we wrote a program to parse

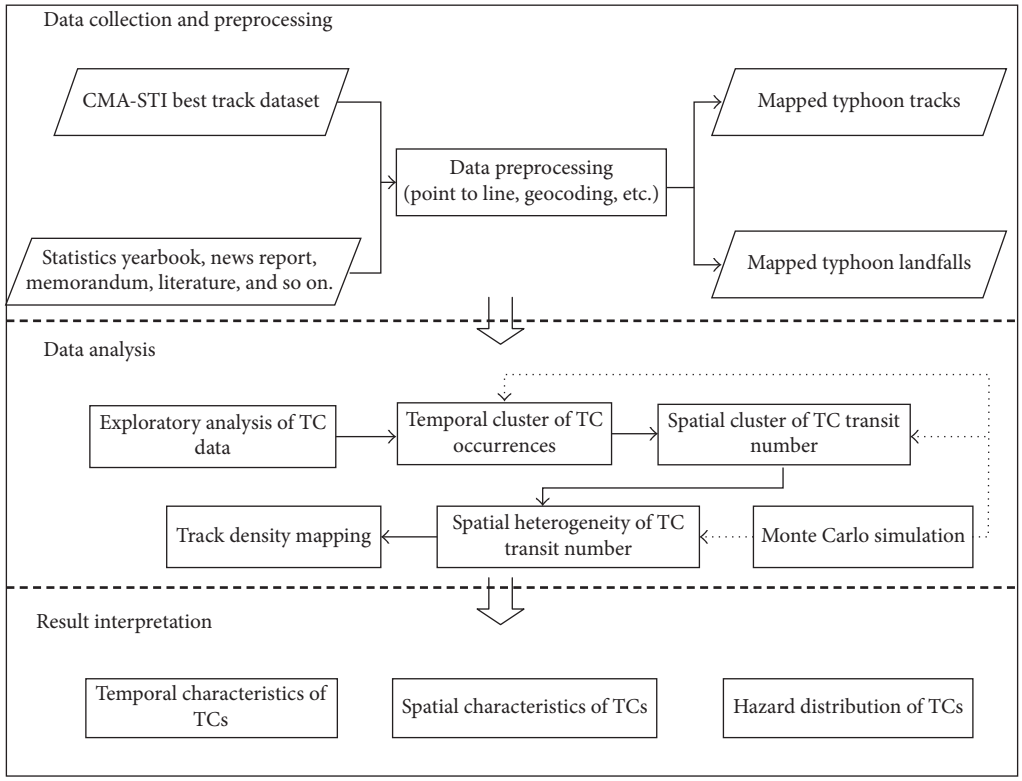


Figure 2: Flowchart of the proposed methodology.

the original text files into XY Data files (required by the Add XY Data operation of ArcGIS as input), and then these resultant files are imported into ArcGIS and transferred into point features. Finally, the Points To Line tool is executed to obtain the TC tracks of line features.

2.3. TC Landfall Data. As mentioned above, the time interval of the adjacent track points recorded is 6 hours. Thus, it is difficult to determine the accurate landfall locations of the TCs. To obtain the most accurate landfall locations, some materials are collected from various sources (such as statistics yearbooks, news reports, memorandum, and the literature). In these materials, generally the landfall locations of TCs are recorded as addresses (e.g., Putuo, Zhoushan, Zhejiang Province). The geocoding technique is applied to interpret the addresses into coordinates of longitude and latitude. Given the abundant gazetteer possessed by some online maps like the Baidu map, we created a Baidu map API-based program to automatically geocode the landfall addresses of TCs of interest in batch.

3. Methods

3.1. Workflow. The study of temporal and spatial characteristics of tropical cyclones is performed with the following steps:

(1) *Data collection and preprocessing*: The methods and procedure are explained in the above sections.

(2) *Data analysis*: Exploratory analysis, temporal pattern analysis, and spatial pattern analysis are done with some selected methods. A Monte Carlo simulation is employed to support the hypothesis test of the temporal and spatial cluster methods. To visualize the hazard of TCs, the track density is mapped through a track density estimation algorithm (line-based KDE).

(3) *Result interpretation*: From the results of data analysis, the temporal and spatial characteristics of TCs are summarized.

A brief overview of the methodology is shown in Figure 2.

The main data analysis techniques will be explained in Sections 3.2–3.5.

3.2. Temporal Cluster Detection of TC Occurrences. Let $t_1 \le t_2 \le \cdots \le t_N$ be N TC occurrences on a time period (e.g., one year) ordered according to occurrence dates. A scan statistic with a variable window (interval) is as follows [11]:

$$\Lambda = \sup_{0<d<n/N, n\ge n_0} \left(\frac{n}{N}\right)^n \left(\frac{N-n}{N}\right)^{N-n} \left(\frac{1}{d}\right)^n \left(\frac{1}{1-d}\right)^{N-n}, \quad (1)$$

where d is the scan window width, n is the number of TC occurrences contained in the scan window, and n_0 is the minimum number of occurrences contained in the clusters to be detected.

This test is the generalized likelihood ratio test for a uniform Null distribution against an alternative of non-random clustering which will detect the most likely cluster (interval) whose observed value of the likelihood ratio is

maximum. This method allows for clusters of variable width. For more details, please refer to Appendix A.

3.3. Spatial Clustering Analysis of TC Crossing Number. Spatial clusters are those regions (prefectures in this study) whose values are significantly higher than their neighbors. Identifying spatial clusters allows one to map "hot" areas and "cold" areas of TCs. The local Moran's I statistic is an index for identifying clusters and outliers and computed according to the following formula:

$$I_i = \frac{x_i - \overline{X}}{S_i^2} \sum_{j=1, j\neq i}^{n} w_{i,j}\left(x_j - \overline{X}\right), \tag{2}$$

where x_i is an attribute of region i, $\overline{X}$ is the mean of the corresponding attribute, $w_{i,j}$ is the spatial weight between region i and j, and

$$S_i^2 = \frac{\sum_{j=1, j\neq i}^{n}\left(x_j - \overline{X}\right)^2}{n-1} - \overline{X}^2, \tag{3}$$

with n equating to the total number of regions in a study area.

Generally, the distribution of I_i does not meet normality. Though the mean and standard deviation are roughly consistent with those for a standard normal distribution, the kurtosis and the skewness are not [14]. In this study, it is more unlikely that the distribution of the Null hypothesis is normal since we are going to explore the spatial clusters of TCs along the coastal regions that is a zonary area. A Monte Carlo simulation is employed to construct a Null hypothesis (see Section 3.5 for the explanation of the simulation process), and further by a classic hypothesis test procedure, the P value for each region is obtained. The P values represent the statistical significance of the computed index values. A Moran's scatterplot can be used to explore the clusters of attributes of regions in a study area in which the X axis denotes the normalized observed attributes and the Y axis is the spatial lags [14].

3.4. Stratified Heterogeneity of TC Crossing Number. Wang et al. [29] define a q-statistic to measure stratified heterogeneity. The q-statistic distribution is defined as

$$P(q < x) = P\left(F < \frac{N-L}{L-1}\frac{x}{1-x}\right) = 1 - \alpha, \tag{4}$$

where N is the number of regions in a study area, L is the number of stratum, F follows a noncentral F-distribution with the 1st df $L-1$, the 2nd df $N-L$, and noncentrality λ, and α is the probability of $q \geq x$. For more information, please refer to Appendix B.

According to the q-statistic, a hypotheses test is carried out by defining the Null hypotheses and alternative hypotheses, respectively. The procedure can be implemented by a software: GeoDetector, which is freely downloadable from http://www.geodetector.org.

3.5. Test Statistical Significance. To assess the significance of a cluster (including temporal and spatial), a Monte Carlo method of inference is used to implement the hypothesis tests. First, the test statistic value is computed according to the observed data. Secondly, the same statistic is calculated with a given quantity of realizations drawn independently from the Null hypothesis (generally a uniform distribution). Next, a distribution can be fitted by some methods (e.g., histogram and kernel density), which provides an estimate of the distribution of the test statistic under the Null hypothesis. The proportion of test statistic observed for the actual data set provides a Monte Carlo estimate of the upper-tail P value for a one-sided hypothesis test [30]. Specifically, suppose that T_{obs} denotes the test statistic value for the sample data and $T(1) \geq T(2) \geq \cdots \geq T(N_{\text{sim}})$ denote the test statistic values for the simulated data of realizations. If $T(1) \geq \cdots \geq T(l) \geq T_{\text{obs}} \geq T(l+1)$, then the estimated P value is

$$P\left(T \geq T_{\text{obs}} \middle| H_0\right) = \frac{l}{N_{\text{sim}} + 1}, \tag{5}$$

where one is added to the denominator since the estimate is based on $N_{\text{sim}} + 1$ values $\left(\{T(1), \ldots, T(N_{\text{sim}}), T_{\text{obs}}\}\right)$. The lower-tail P values can be calculated in an analogous manner.

3.6. Track Density Estimation. To visualize the hazard of TCs, an alternative approach is to evaluate the spatial density of the TC tracks. The algorithm for track density can be considered an extension of the widely used 1D kernel density on point data [31] into 2D kernel density on polyline data (tracks). The contribution of each track to the estimated density is calculated separately, which represents the influence of the particular track on its neighborhood in space. Then, those contributions from all tracks are added up to obtain the total density. The contribution of each track (polyline) to the density is defined within a buffer whose centerline is the projection of the track on the plane space. A value is assigned to each point in the buffer according to the distance of the point to the centerline (the perpendicular distance). This value needs to be normalized using the user-specified kernel size, so that the point on the centerline is assigned a maximum value, and the density value decreases with distance from the centerline and reaches a minimum value at the limit of the kernel. Furthermore, in order to be consistent with the probability density function (pdf), the integral of the kernel function along the perpendicular segment must be equal to 1. For the computation detail, please refer to Appendix C.

4. Results and Discussion

4.1. Temporal Analysis of the TC Occurrences. There are a total of 2233 recorded TCs in the Northwest Pacific and 816 landfalling TCs along the continental coastline during 1949–2014. The landfalls are more than 12 in an average year. The landfalls amount to 25 in 1961 with 17 as tropical storms or above. The annual number of the TC occurrences and their landfalls are shown in Figure 3.

Several radar plots (also called sample circular plots or spider plots in some literature) are presented in Figures 4 and 5. Figure 4 shows the monthly pattern of the recorded

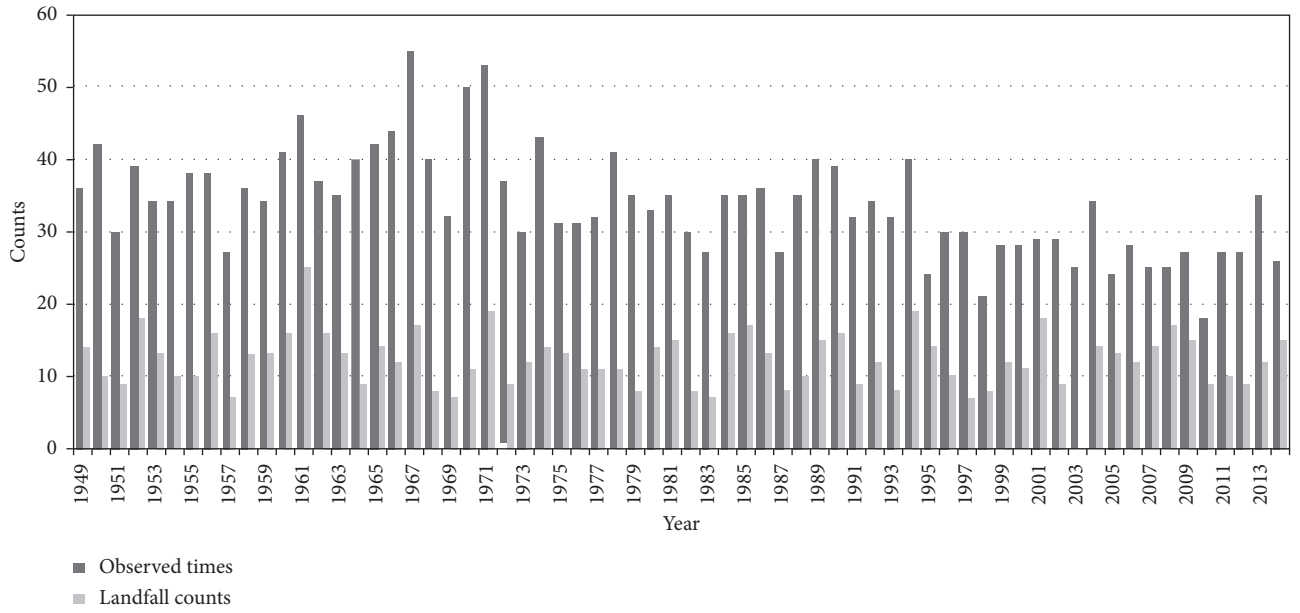


Figure 3: The annual number of the observed TC occurrences in the Northwest Pacific and their landfall counts approaching China.

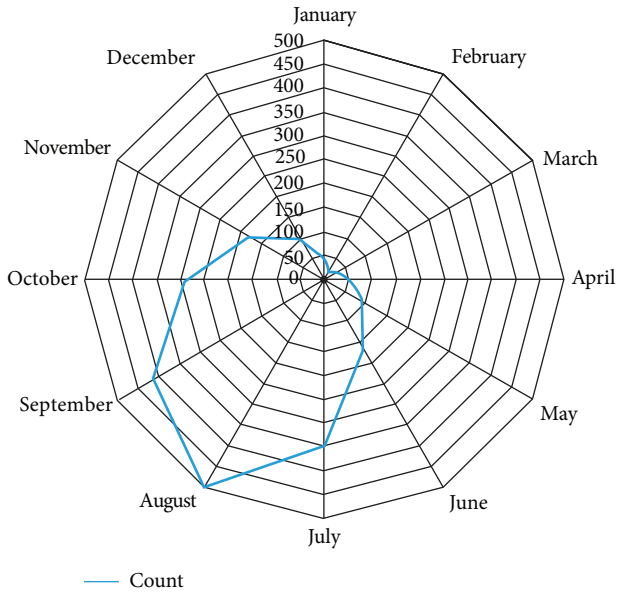


Figure 4: Monthly distribution of the observed TC occurrences in the Northwest Pacific during 1949–2014.

TC occurrences in the Northwest Pacific during 1949–2014. Figure 5 shows the monthly pattern of the landfalling TCs along the coastline of China during 1949–2014. These two figures show that the number of TCs increased from April to August and decreased from August to October with the peak occurring in August. According to Figure 5, fortunately, STY and SuperTY, the most destructive TCs, occupy a relatively small proportion in all TCs. Nonetheless, the tremendous consequences caused by STY and SuperTY indicate that the need for prevention and response to them cannot be overemphasized.

Next, we try to detect the most likely clusters of the TC occurrences during 1949–2014 in the Northwest Pacific, which is able to be achieved through a scan over the period with a variable window. For each year, the most likely cluster of points (the times of the TC occurrences) along the whole year (from January 1 to December 31) is detected. The cluster detection of points along line is implemented to look for the most likely clusters. The algorithm presented in (1) and Appendix A is coded with Matlab to find the most likely cluster. Meanwhile, the algorithm of the Monte Carlo simulation is coded to test the significance of the most likely cluster. In our analysis, n_0 is set to 5, the max length of the clusters is limited to 178 (days), and the simulation times are set to 999. After finishing the computation, the starting position, the number of points, and P value are output.

Figure 6 shows the detection result. From the visual examination, these clusters are obviously concentrated in a certain range of time. The mean of the durations of all 66 clusters from 1949 to 2014 is 52.5 (days) and the corresponding standard deviation is 17.0 (days). The average starting point is the 210.5th day. The mean of the centers of these clusters is the 237th day which is consistent with Figure 4.

4.2. Spatial Clusters of Landfalling TCs along the Coastline. The local Moran's I statistic aforementioned is used to analyze clusters of TCs traveling across those prefectures along the continental coastline of China. The landfall times of TCs for every city are summarized. This operation is accomplished through the Spatial Join tool of ArcGIS, which can join the tracks to the prefectures according to the spatial topological relationship of Interaction (meaning that the TC passes through the prefectures). ArcGIS also has been devised with several spatial cluster analysis tools including the Local Moran's I and Getis-Ord G_i^* [15, 32]. The tools produce a map in terms of the calculated results of the statistic itself and P value. Here, we define the spatial weight w_{ij} with Rook contiguity [33]. Figure 7(a) shows the clusters of the times of TCs traveling across the continental coastal prefectures of China, and the corresponding P values are

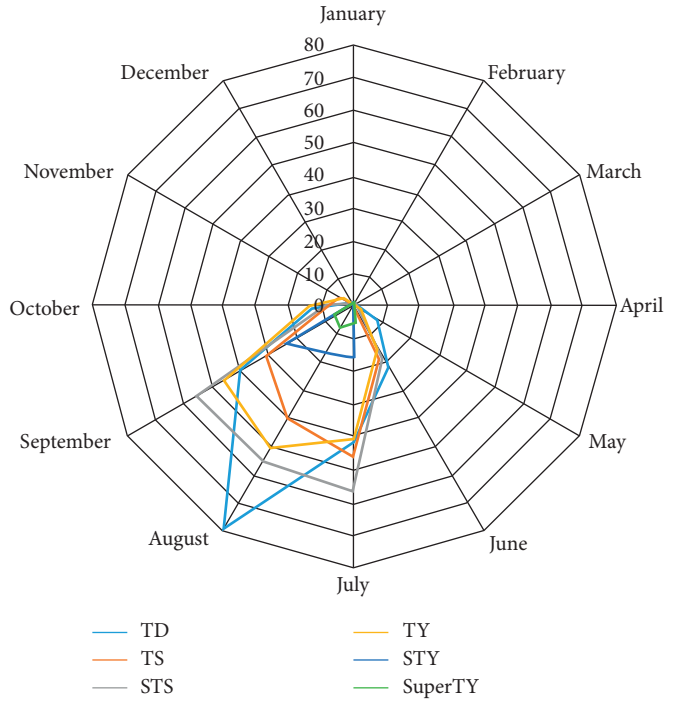


Figure 5: Monthly distribution of the landfalling TCs along the coastline of China during 1949–2014.

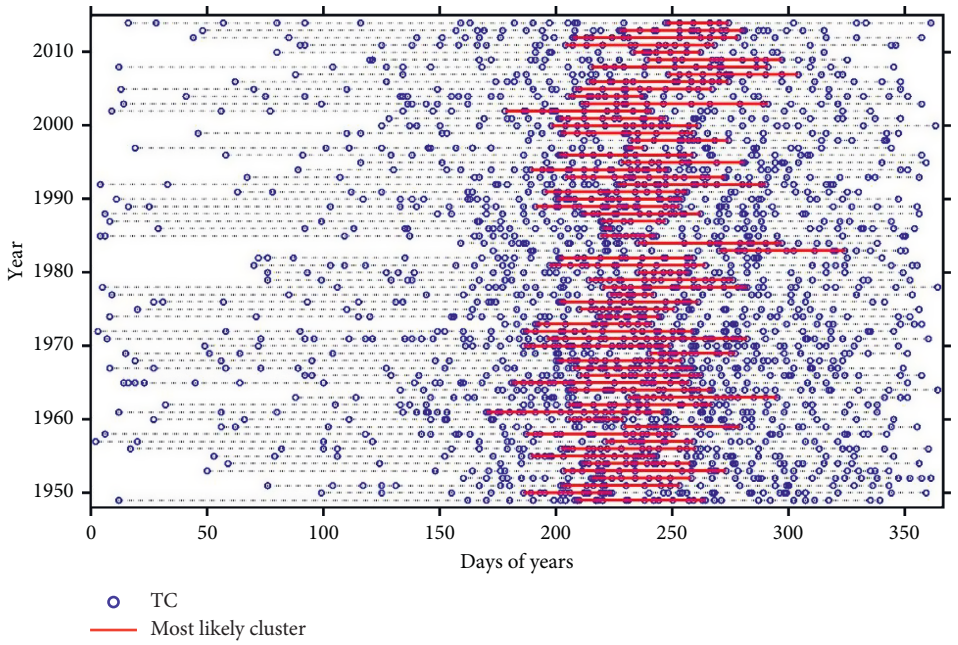
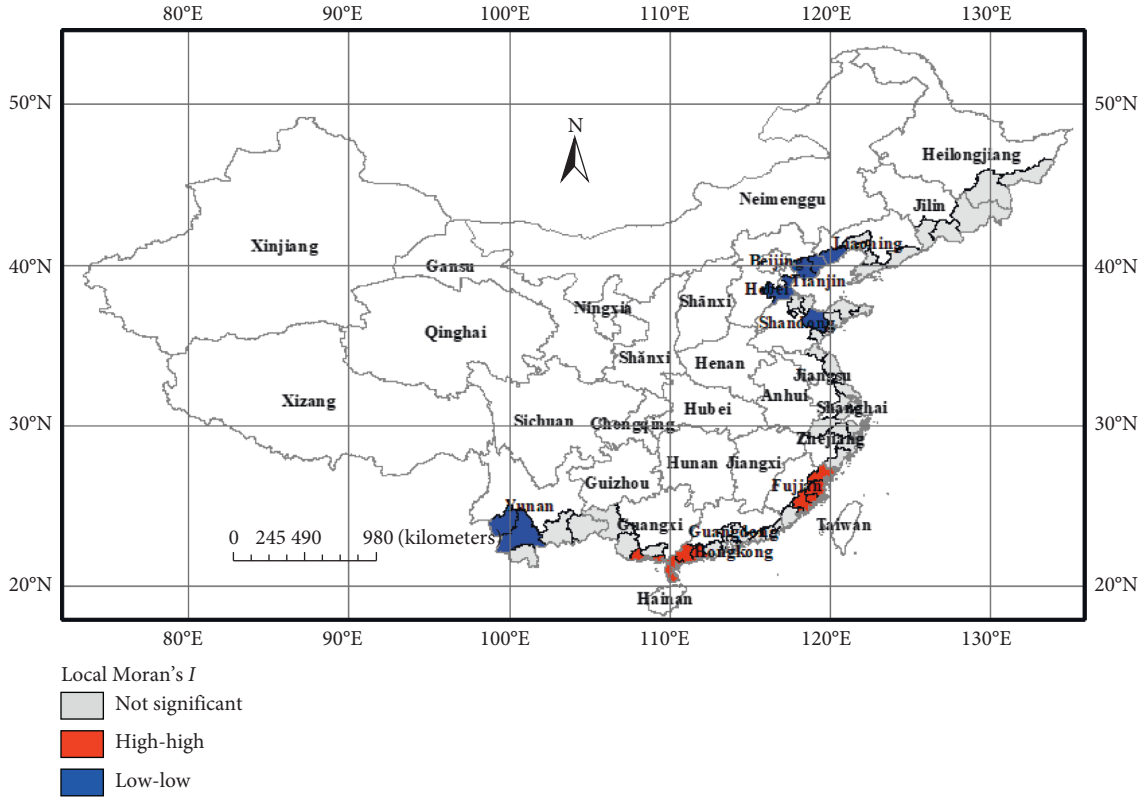
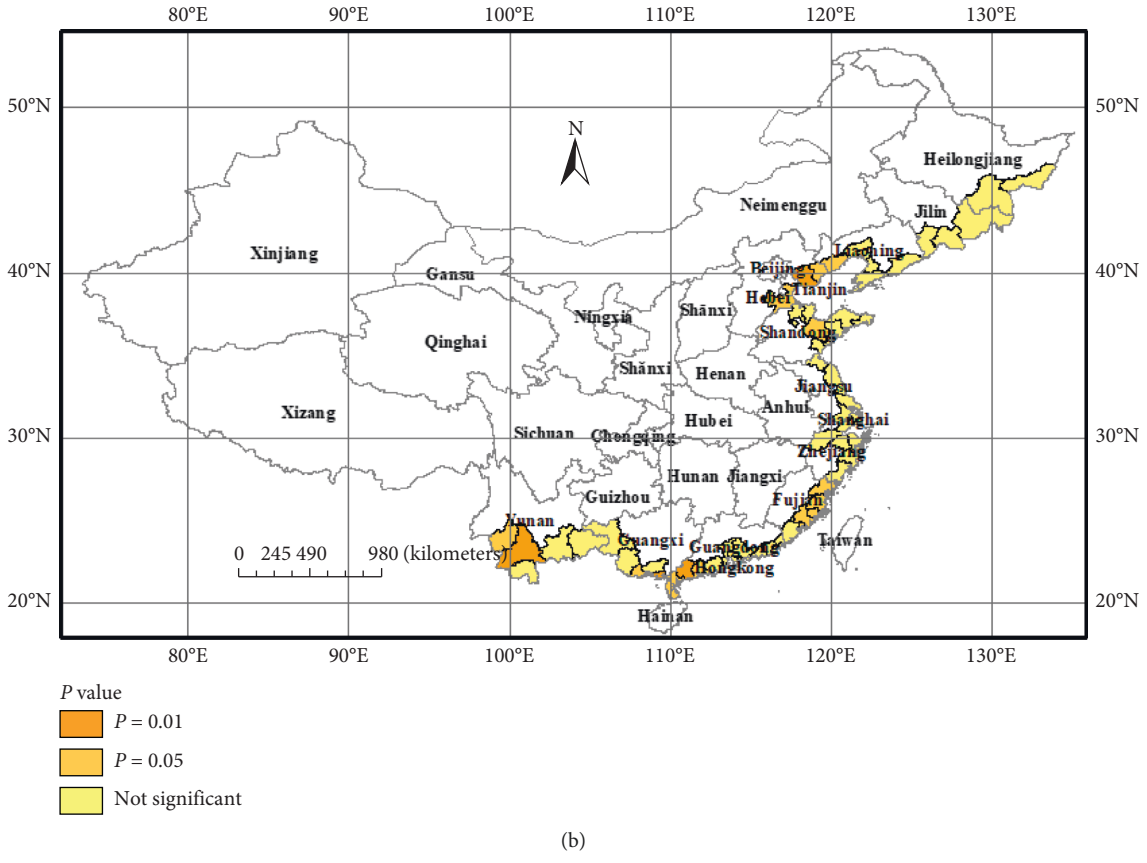


Figure 6: The most likely clusters detected for each year during 1949–2014.

shown in Figure 7(b). As seen, the clusters are located in two areas (filled with red in Figure 7(a)): the boundary of Guangxi and Guangdong provinces and the boundary of Fujian and Zhejiang provinces. These two areas are the most frequently stroke ones that can be easily confirmed by comparison with the historical TC records. Other prefectures do not have statistically significant clusters according to the analysis results. A Moran's scatterplot is drawn from the results of local Moran's I (Figure 8). In a Moran's scatterplot, a point is drawn in one of the four phases according to its X value and Y value. The first phase of the plot contains the HH (High-High) values of the observed data, that is, the attribute of a region is high and the attributes of its neighboring region are also high, which means these regions are a cluster. Similarly, the second phase contains LH (Low-High) values, the third phase contains LL (Low-Low) values, and the fourth phase contains HL (High-Low) values. Here, our interest obviously is on the first phase in order to learn the spatial clusters of the visiting



(a)



(b)

Figure 7: A resultant map of cluster analysis of the landfall times of the TCs along the continental coastal prefectures of China during 1949–2014 according to the observed z-scores. (a) Local Moran's I cluster map; (b) P value map.

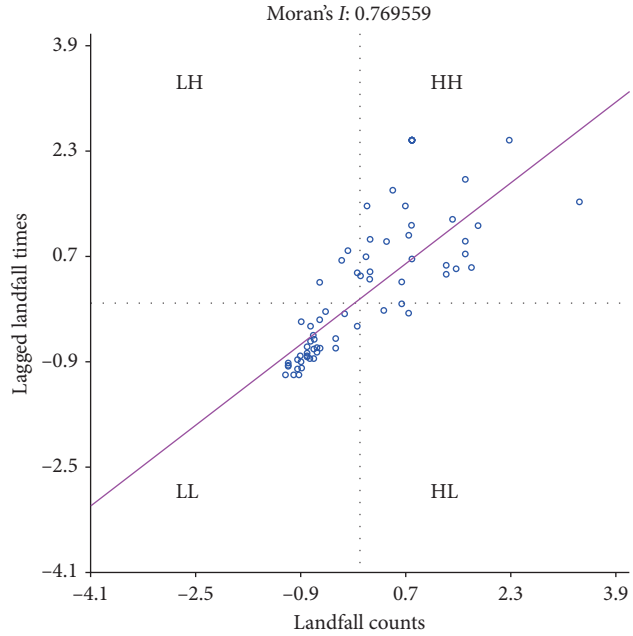


Figure 8: The Moran's scatterplot of the landfalling TCs along the continental coastal prefectures of China during 1949–2014 (the first phase means clusters that the landfalling TCs in those prefectures are high and whose neighbors also have high landfall counts).

Table 1: Outputs of q-statistic and significance tests for spatial stratified heterogeneity.

N	$\overline{Y}$	L	N_h	$\mu_h = \overline{Y}_h$	σ^2	σ_h^2	λ	F_α	F	q-statistic	P value
66	23.71	3	48	21.02	431.47	259.38	3.91	10.96	24.52	0.44	5.430×10^{-6}
			10	53		389.33					
			8	3.25		10.78					

Table 2: Outputs of q-statistic and significance tests for temporal stratified heterogeneity.

N	$\overline{Y}$	L	N_h	$\mu_h = \overline{Y}_h$	σ^2	σ_h^2	λ	F_α	F	q-statistic	P value
36	62.03	4	9	11	2913.8	22	3.25	8.06	54.77	0.87	1.993×10^{-10}
			9	34.67		390.5					
			9	140.33		823.5					
			9	62.11		842.11					

frequencies of TCs. As expected, the dots for detected clusters are drawn in the first phase and have greater observation values and expected mean neighbor values.

4.3. Stratified Heterogeneity of TCs. Two tests of stratified heterogeneity of TCs are carried out. One is for spatial stratified heterogeneity of landfall counts of TCs in the coastal prefectures. In this case, we divide the coastal prefectures into three strata: Low-Low, High-High, and Not Significant according to the analysis results of local Moran's I. In the other, we test the temporal stratified heterogeneity of the TC occurrences in different seasons, that is, we divide the TC occurrences into four strata: January–March, April–June, July–September, and October–December. We assessed these assumptions by the q-statistic to the landfall counts and TC occurrences of the spatial and temporal strata. We also tested the significance of the stratified heterogeneity. The results are shown in Tables 1 and 2. In the spatial case, because $F = 24.52 > F_{0.01}\ (2, 63, 3.9082) = 10.96$ and $P = 5.43 \times 10^{-6} \ll 0.01$, we conclude that the spatial stratified heterogeneity is significant. In the temporal case, because $F = 54.77 > F_{0.01}\ (3, 32, 3.2549) = 8.06$ and $P = 1.99 \times 10^{-10} \ll 0.01$, the temporal stratified heterogeneity is significant. We can also identify the most likely boundary of strata by q-statistic to test various stratifications to reach the one with the maximum q value, which should be performed in a future work [29].

4.4. Track Mapping of TCs. The landfall counts of TCs in the coastal prefectures are by year. Assume the landfall counts n_i for a certain prefecture i obey Poisson distribution with intensity λ_i, that is,

$$\Pr(n_i = k) = \frac{(\lambda_i t)^k e^{-\lambda_i t}}{k!}. \tag{6}$$

The exceedance probability (EP) can be calculated as follows:

$$\Pr(n_i \geq k) = 1 - \sum_{j=0}^{j=k-1} \Pr(n_i = j). \quad (7)$$

According to the exceedance probability, the return period of TCs in prefecture i can be obtained by (8)

$$r_i = \frac{1}{\Pr(n_i \geq 1)}. \quad (8)$$

The annual intensities for every prefecture are estimated according the landfall counts of TCs from 1949 to 2014.

$$\widetilde{\lambda}_i = \frac{\sum_{j=1949}^{2014} f_{ij}}{2014 - 1949 + 1}, \quad (9)$$

where f_{ij} is the observed landfall times of TCs for prefecture i in year j.

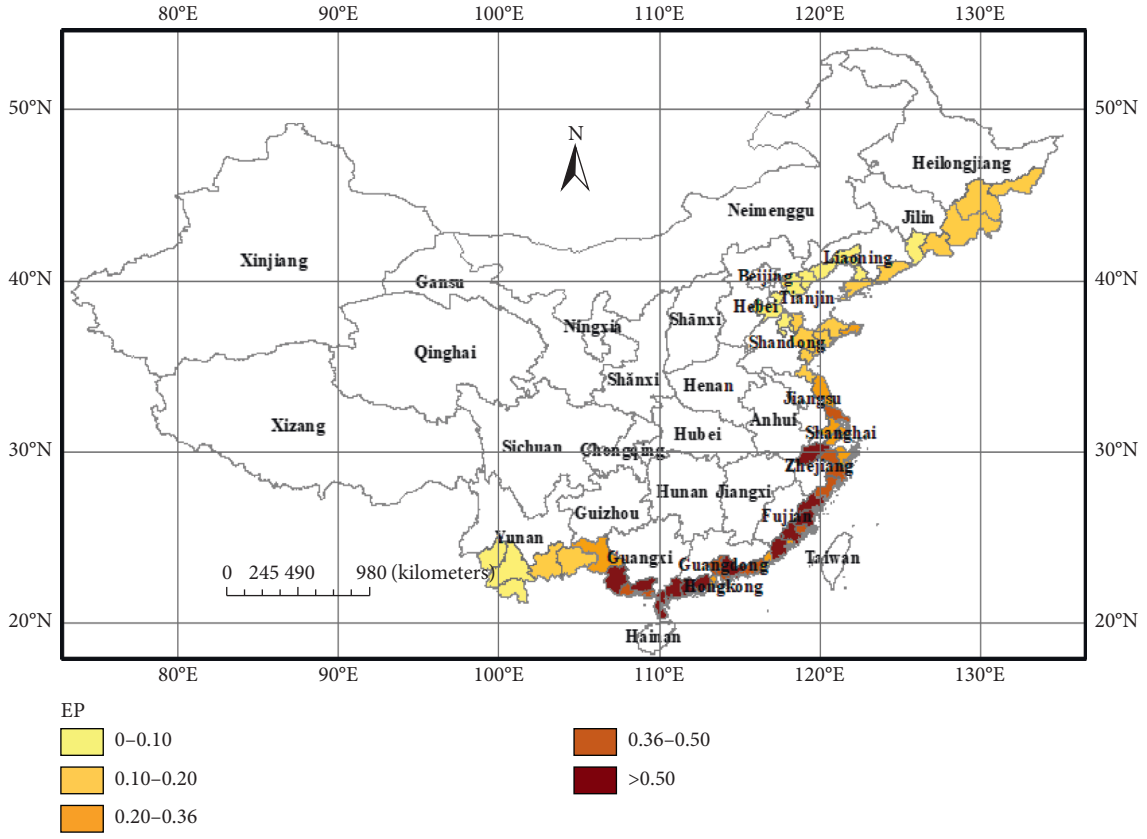
Figure 9 depicts the exceedance probabilities of the 66 coastal prefectures of China for landfall counts of TCs. According to Figure 9, 12 coastal prefectures (Zhanjiang, Maoming, Fuzhou, Huizhou, Yangjiang, Qinzhou, Ningde, Quanzhou, Jiangmen, Nanning, Zhangzhou, and Hangzhou) have return periods of less than two years. We also observe that: (1) Zhanjiang, Maoming, and Fuzhou rank highest in modeled levels of EP; (2) Huizhou, Yangjiang and Qinzhou are at lesser, but still relative to the remaining prefectures, high EP; and (3) of all 66 selected coastal prefectures, Lincang has no landfalling TCs according to historical records.

Furthermore, a track density map is produced to visualize the continuous impact of TCs around the coastline of China (Figure 10). A Gaussian kernel function is selected here. The formation of the kernel function is defined as $k(\|x - x_c\|) = e^{-\|x - x_c^2\|/(2\sigma)^2}$, where x_c is the center of the kernel, and σ is window width (a value of 5° is specified). Through simply a visual investigation, we can find out that there are two peaks in the track density map. One peak is over the South China Sea, the other is located to the southeast of China in the Western Pacific. Furthermore, there is a dominant northwest gradient of the density. We inferred that large-scale atmospheric circulation causes this result. Further statistical and mechanism analyses are required to ascertain the underlying fact. The analysis results show that the two island provinces of Hainan and Taiwan are located near the peaks. Guangdong, Guangxi, Fujian, and Zhejiang provinces suffered severely from destructive TCs. In comparison, several northern coastal provinces have significantly less TC landfalls. These results show consistency with existing literature [34–36]. The track density map provides a continuous presentation of TC hazards of "places" across China, which will be helpful for preparedness and mitigation of TCs such as hazard hotspot identification and risk assessment.

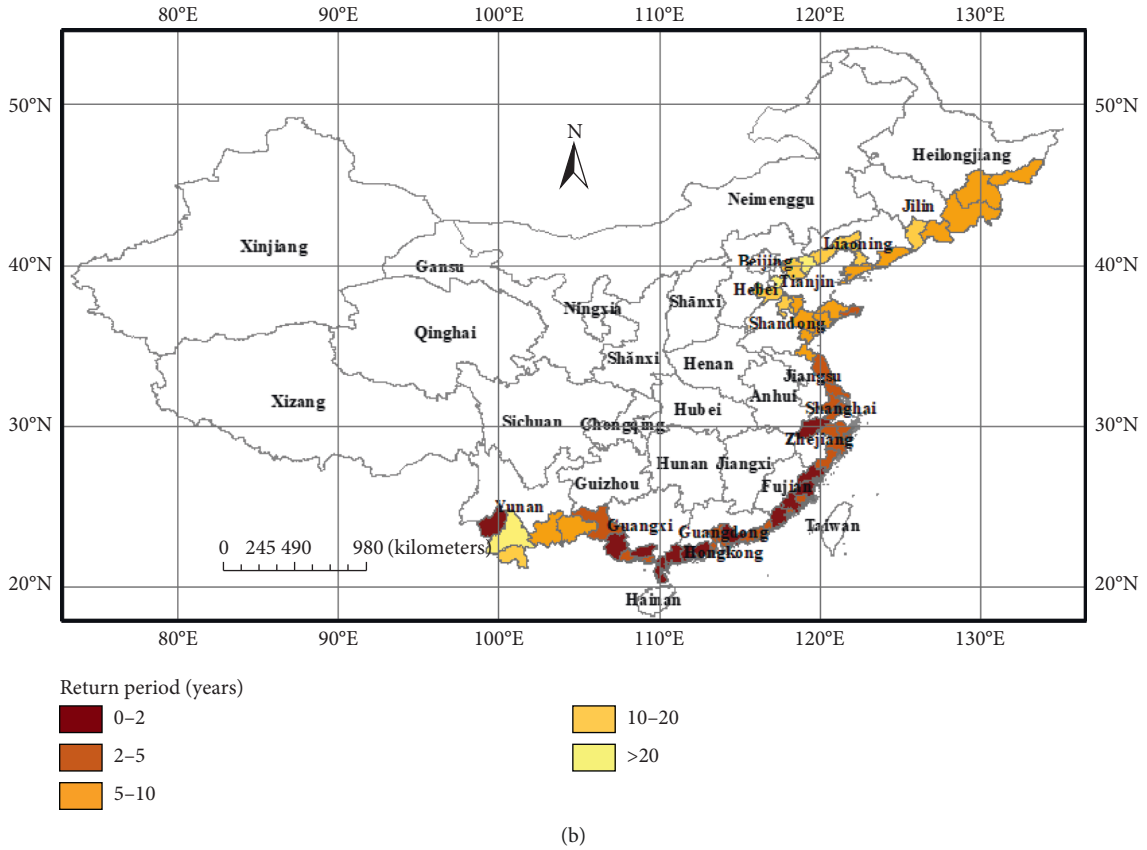
5. Conclusions

Spatiotemporal patterns of TCs are one of the greatest interests in prevention and control of the destructive disasters. While the accurate behavior of a certain TC is difficult to determine thanks to our limited cognition for the complex dynamic mechanism [37, 38], data-oriented analysis provides a kind of feasible approach to mining hidden patterns implied in the massive tracks. Time stamps and geographical locations are two types of substantial attributes for TC track data. As stated in the first law of geography, in a spatial context, spatial (spatiotemporal) correlation must be taken into consideration, which is an essential difference between spatial and nonspatial analysis. Since some proprietary GIS software packages are devised with powerful spatiotemporal analysis tools, some traditional methods of identifying TC patterns in time and space are complemented and enhanced when coupled with GIS. Visualization-based approaches provide an intuitive survey for observations and are widely put in practice, while they usually fail to be accepted in some occasions due to their imprecisions and inherent weakness. In order to discover real potential patterns, some robust methods are used in spatial data analysis. When faced with a question like what methods are appropriate for our analysis, the representation of spatial data is also of great concern. As a typical kind of linear features composed of a sequence of points, linear feature specific methods are usually preferred for track analysis of TCs and the like.

The typical periodicity and centralization of TCs encourages us to explore the potential patterns of TCs. Through plotting the circular map on different time scales (month, season), an obvious monthly and seasonal difference is presented. Point cluster analysis along line (temporal interval) is carried out to confirm the exploratory results by a rigorous approach. The cluster analysis demonstrates not only the consistency, but also the detailed analysis of statistical characteristics of these clusters through a Monte Carlo simulation, which is impossible to be perceived by visualization-based analysis methods. The detected temporal clusters are consistent with empirical knowledge. Since the temporal dimension can only describe the marginal features of TCs, several spatial aspects are explored. At this time, our focus is on the landfall locations of TCs. First, a local Moran's I index is used to detect spatial clusters of TCs. The analytical results are consistent with the intuitive realization of TC clustering along the coastline of China. Additionally, statistical significance is calculated according to the spatial Monte Carlo algorithm. This extended analysis helps refine the visual estimation of clusters and provides a rigorous approach to result assessment. Next, to evaluate the stratified heterogeneity of TCs, both spatial and temporal stratified heterogeneity is carried out according to a q-statistic. The tests show that stratified heterogeneity is significant, which means the within-strata variance is less than the between-strata variance. The results imply the locational and seasonal factors are the potential determinants of the heterogeneity of TCs. Sequentially, the exceedance probabilities and the return periods for the coastal prefectures are rendered. Finally, a line-based kernel density estimation is performed on the TC tracks generated in the Northwest Pacific. Coupled with some referable data such as population and GDP, the resultant map is intended to quantitatively evaluate the hazard of TCs and further provide substantial reference for preparedness and mitigation of future TCs along the coastline of China. Furthermore, compared to



(a)



(b)

Figure 9: Exceedance probabilities and return periods of the landfall counts of TCs in the coastal prefectures. (a) Exceedance probability; (b) return period.

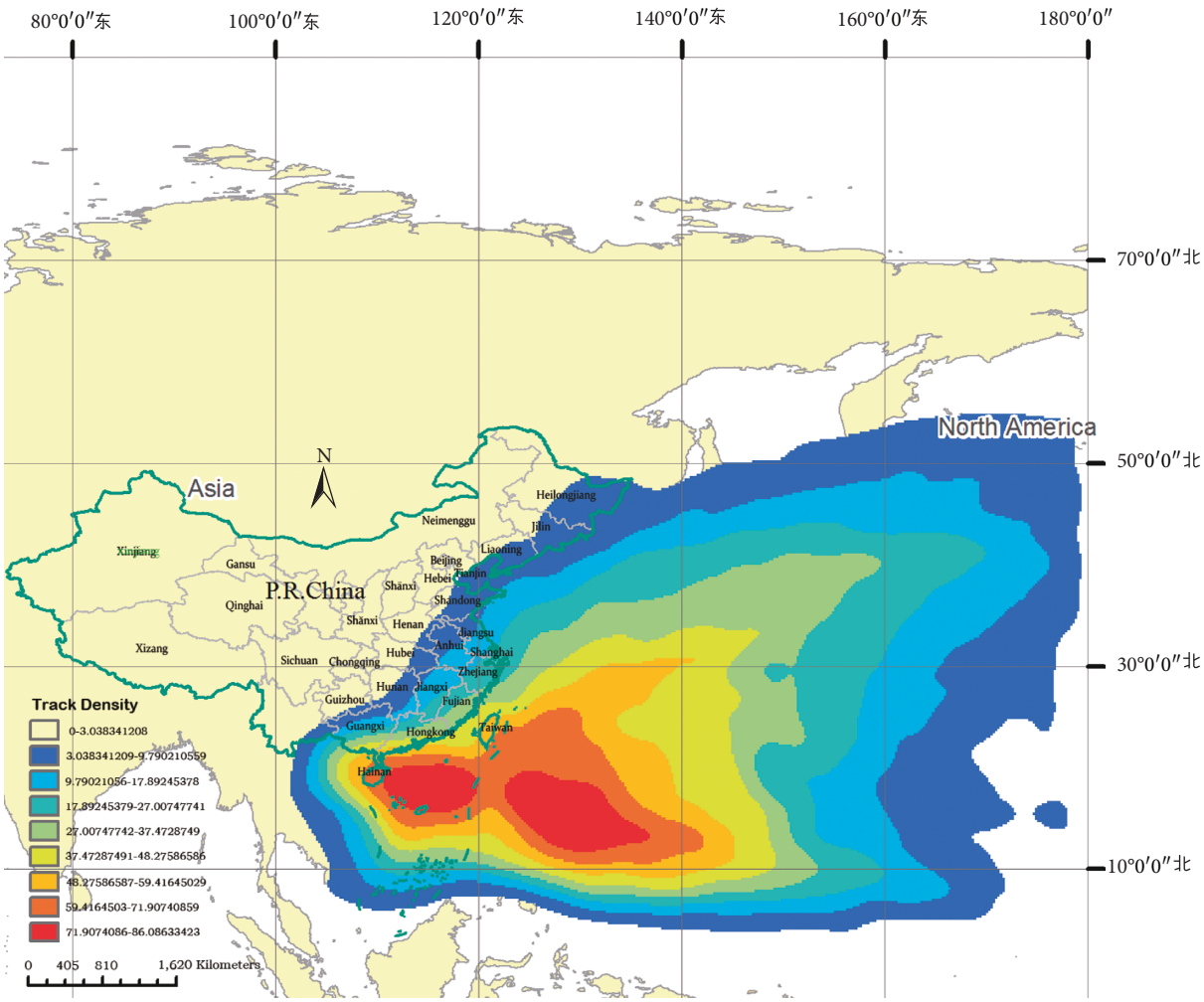


Figure 10: Track density map of TCs in the Northwest Pacific.

some findings in existing literature related to the TC mechanisms, these spatial and temporal patterns are consistent with several universally accepted mechanisms that have been deeply studied by some researchers. For example, the TC swarm occurring in space or time is closely related to atmospheric circulation conditions of the tropical Northwest Pacific [39]. The status of the warm pool in the West Pacific largely affects the interannual change of the moving tracks of TCs [40]. Therefore, those comprehensive effects lead to the patterns.

Overall, this paper systematically explored temporal and spatial patterns and mapped the hazards of TCs, through the integrated application of advanced spatial techniques. This work can help obtain insight into temporal and spatial characteristics of TCs compared to previous efforts. These kinds of findings can assist with proactive destructive prevention activities, thereby enhancing the preparedness and mitigation capacity in a predisaster phase. While the present findings offer useful insights into the spatiotemporal patterns of TCs along the coastline of China, further work is clearly required both to reinforce the results and to extend our understanding of the problems. For example, satellite remote sensing techniques enable retrieval of meteorological parameters from remotely sensed data since 1970s, which makes it necessary to compare between data from meteorological stations and satellites. In prevention and response, a continued effort by researchers, the authorities, and the public is imperative in order to enhance prevention ability and reduce the loss of lives and property damage caused by TCs.

Appendixes

A. Cluster Detection of Points along a Line

Let $x_1 \le x_2 \le \cdots \le x_N$ be N points independently drawn from the uniform distribution on a line segment (or timeline) ordered according to size. The joint p.d.f. at $x_1, x_2, \ldots, x_N$ is $N!$ and $1 - \Pr(n|N; p)$, the probability that no n points are contained in a length p, is the multiple integral of $N!dx_1, \ldots, dx_N$ over $x_1, \ldots, x_N$ in the region: $a \le x_1 \le x_2 \le \cdots \le x_N \le b$, and $x_{n+i-1} - x_i > p$, for all $i = 1, 2, \ldots, N - n + 1$. Naus derived $\Pr(n|N; p)$, the c.d.f. of p, and the size of the smallest interval that contains n out of the N points. This calculation gave rise to a test statistic for detecting clusters, namely, the scan statistic.

Assume $x_1, x_2, \ldots, x_N$ be a random sample of N points from the density:

$$f(x) = \begin{cases} a, & \text{for } b \le x \le b + d \\ \dfrac{1 - ad}{1 - d}, & \text{for } 0 \le x \le b \text{ or } b + d \le x \le 1 \end{cases} \tag{A.1}$$

For the hypothesis test problem,

$$H_0: a = 1, \qquad H_1: \frac{1 < a \le 1}{d}, \tag{A.2}$$

where d is the scan window width and b is the start location of the scan window. Obviously, if $a = 1$, then the sample is drawn from a uniform distribution in [0, 1]. Otherwise, when $a > 1$, the sample has higher density in $[b,\ b + d]$ than in $[0,\ b] \cup [b + d,\ 1]$. This is the basic principle of the hypothesis test.

According to the above notation, the scan statistic is formulated as follows:

$$\Lambda = \sup_{0<d<n/N, n \ge n_0} \left(\frac{n}{N}\right)^n \left(\frac{N-n}{N}\right)^{N-n} \left(\frac{1}{d}\right)^n \left(\frac{1}{1-d}\right)^{N-n}, \tag{A.3}$$

where n_0 is the minimum number of points of the clusters to be detected.

B. Test for Stratified Heterogeneity

Assume that a study area is composed of N regions and is stratified into $h = 1, 2, \ldots, L$ stratum; stratum h is composed of N_h regions; and Y_i and Y_{hi} denote the value of region i in the population and in stratum h, respectively.

The q-statistic is defined as follows:

$$q = 1 - \frac{\sum_{h=1}^{L} \sum_{i=1}^{N_h} \left(Y_{hi} - \overline{Y}_h\right)^2}{\sum_{i=1}^{N} \left(Y_i - \overline{Y}\right)^2} = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2} = 1 - \frac{\mathrm{SSW}}{\mathrm{SST}}, \tag{B.1}$$

where the total sum of squares

$$\mathrm{SST} = \sum_{i=1}^{N} \left(Y_i - \overline{Y}\right)^2 = N\sigma^2 \tag{B.2}$$

and the within sum of squares

$$\mathrm{SSW} = \sum_{h=1}^{L} \sum_{i=1}^{N_h} \left(Y_{hi} - \overline{Y}_h\right)^2 = \sum_{h=1}^{L} N_h \sigma_h^2. \tag{B.3}$$

The value of the statistic is within [0, 1] (0 if there is no stratified heterogeneity and 1 if the population is fully stratified) and increase monotonously with the increase of stratified heterogeneity.

Let

$$F \overset{\text{def}}{=} \frac{\mathrm{SSB}/(L-1)}{\mathrm{SSW}/(N-L)}, \tag{B.4}$$

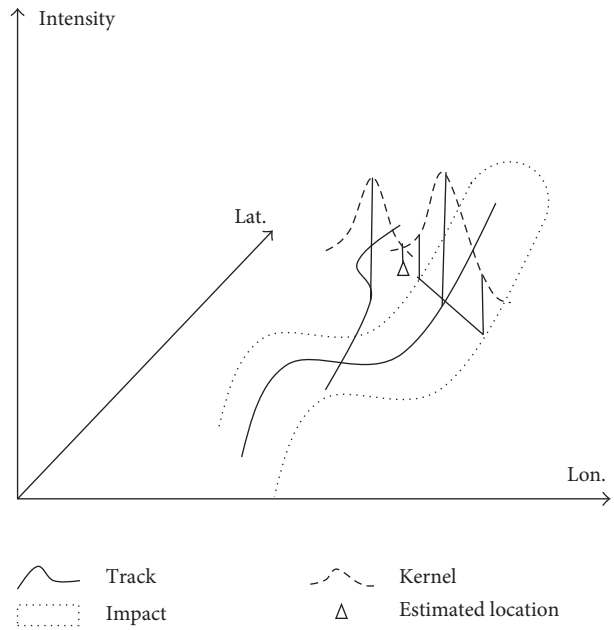


Figure 11: A schematic of the calculation of track density.

where $\mathrm{SSB} = \sum_{h=1}^{L} N_h (\overline{Y}_h - \overline{Y})^2$ is the between-strata sum of squares, then F follows a noncentral F-distribution with the 1st df $L-1$, the 2nd df $N-L$, and noncentrality λ as follows:

$$F \sim F(L-1, N-L; \lambda). \tag{B.5}$$

C. Calculation of Track Density

Let $l_1, l_2, \ldots, l_N$ are the N tracks and for a certain location p, the distances (perpendicular distances) from p to these tracks are $d_{p1}, d_{p2}, \ldots, d_{pN}$, then the track density of p can be calculated as follows:

$$\widetilde{f}(p) = \sum_{i=1}^{N} f\left(d_{pi}\right), \tag{C.1}$$

where $f(\cdot)$ is a distance decay function (kernel function). Figure 11 illustrates how to calculate the track density at a certain location (two tracks are shown in Figure 11).

Kernel estimation methods originally are used to estimate a probability density function. With respect to the track data set, a density function defines the probability of observing a track at a location p. In this sense, only a function conforming to some certain conditions can be adopted as a kernel [30].

The pseudocode in Table 3 presents a more detailed description of the calculation procedure of the track density. For each cell in a tessellation of the area to be kernelled, first calculate the normalized distances (*TrackDensity*) to its neighboring tracks, second, add the densities for each separate track up to a summarized density, and finally, the summarized density (*SumDensity*) is normalized with the number of tracks. Once all cells are processed according to the steps, a raster map of track density will be produced.

Table 3: Pseudocode for calculating the track density.

```
SumDensity = 0
For each Track
{
TrackDensity = 0
   calculate NeighboringArea around the track
   for each Cell in the NeighboringArea
   {
      calculate DistanceToTrack
      TrackDensity = normalized DistanceToTrack
   }
   SumDensity = SumDensity + TrackDensity
}
normalize SumDensity with number of tracks
```

In the algorithm, a track is represented as a list of successive points in 2D space, that is, a 2D polyline consisting of points $p_1, p_2, p_n, p_{n+1}, \ldots, p_m$, where $p_n = (x_n, y_n)$, x_n, and y_n are geographic coordinates. Though a track in the real world is always a continuous line, this discretized representation of the track is a kind of widely used data structure in GIS, which enables storage and analysis of real-world objects in computers. We adopt this standard assumption in our study: a series of straight-line segments between sample points are able to represent an original continuous track approximately.

Conflicts of Interest

No potential conflicts of interest were reported by the authors.

Acknowledgments

The authors acknowledge the support from the National Natural Science Foundation of China (Grant nos. 70901047 and 91224004) and Project in the National Science and Technology Pillar Program during the Twelfth Five-Year Plan Period (Grant nos. 2015BAK10B01 and 2015BAK12B03). The authors also appreciate support for this paper from the Collaborative Innovation Center of Public Safety.

References

[1] NOAA, *What is the Difference Between a Hurricane, a Cyclone, and a Typhoon?*, National Oceanic and Atmospheric Administration, Silver Spring, MD, USA, July 2016, http://oceanservice.noaa.gov/facts/cyclone.html.

[2] P. J. Webster, G. J. Holland, J. A. Curry, and H. R. Chang, "Changes in tropical cyclone number, duration, and intensity in a warming environment," *Science*, vol. 309, no. 5742, pp. 1844–1846, 2005.

[3] D. F. Liu, L. Pang, and B. T Xie, "Typhoon disaster in China: prediction, prevention, and mitigation," *Natural Hazards*, vol. 49, no. 3, pp. 421–436, 2009.

[4] L. Wang, G. H. Chen, and R. H. Huang, "Spatiotemporal distributive characteristics of tropical cyclone activities over the Northwest Pacific in 1979-2006," *Journal of Nanjing Institute of Meteorology*, vol. 32, no. 2, pp. 182–188, 2009.

[5] C. L. Gu, J. C. Kang, G. D. Yan et al., "Spatial and temporal variability of northwest pacific tropical cyclone activity in a global warming scenario," *Journal of Tropical Meteorology*, vol. 22, no. S1, pp. 15–23, 2016.

[6] J. C. L. Chan and K. S. Liu, "Global warming Western North Pacific typhoon activity from an observational perspective," *Journal of Climate*, vol. 17, no. 23, pp. 4590–4602, 2010.

[7] L. Tang, H. U. Deyong, and L. I. Xiaojuan, "Spatiotemporal characteristics of tropical cyclone activities in Northwestern Pacific from 1951 to 2006," *Journal of Natural Disasters*, vol. 21, no. 1, pp. 31–38, 2012.

[8] R. B. Santos, *Crime Analysis with Crime Mapping*, Sage Publications, Thousand Oaks, CA, USA, 2016.

[9] J. I. Naus, "The distribution of the size of the maximum cluster of points on a line," *Journal of the American Statistical Association*, vol. 60, no. 310, pp. 532–538, 1965.

[10] R. J. Huntington and J. I. Naus, "A simpler expression for kth nearest neighbor coincidence probabilities," *Annals of Probability*, vol. 3, no. 5, pp. 894–896, 1975.

[11] N. Nagarwalla, "A scan statistic with a variable window," *Statistics in Medicine*, vol. 15, no. 7–9, pp. 845–850, 1996.

[12] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1-2, pp. 17–23, 1950.

[13] A. D. Cliff and J. K. Ord, *Spatial Processes: Models and Applications*, Pion, London, UK, 1981.

[14] L. Anselin, "Local indicators of spatial association—LISA," *Geographical Analysis*, vol. 27, no. 2, pp. 93–115, 1995.

[15] J. K. Ord and A. Getis, "Local spatial autocorrelation statistics: distributional issues and an application," *Geographical Analysis*, vol. 27, no. 4, pp. 286–306, 1995.

[16] M. Kulldorff, "A spatial scan statistic," *Communications in Statistics-Theory and Methods*, vol. 26, no. 6, pp. 1481–1496, 1997.

[17] M. A. Costa and M. Kulldorff, "Maximum linkage space-time permutation scan statistics for disease outbreak detection," *International Journal of Health Geographics*, vol. 13, no. 1, p. 20, 2014.

[18] J. F. Wang, X. H. Li, G. Christakos et al., "Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China," *International Journal of Geographical Information Science*, vol. 24, no. 1, pp. 107–127, 2010.

[19] A. Asgary, A. Ghaffari, and J. Levy, "Spatial and temporal analyses of structural fire incidents and their causes: a case of Toronto, Canada," *Fire Safety Journal*, vol. 45, no. 1, pp. 44–57, 2010.

[20] U. Demšar and K. Virrantaus, "Space–time density of trajectories: exploring spatio-temporal patterns in movement data," *International Journal of Geographical Information Science*, vol. 24, no. 10, pp. 1527–1542, 2010.

[21] P. Timothée, L. B. Nicolas, S. Emanuele et al., "A network based kernel density estimator applied to Barcelona economic activities," in *Proceedings of the Computational Science and Its Applications (ICCSA 2010)*, pp. 32–45, Fukuoka, Japan, March 2010.

[22] S. Steiniger and A. J. S. Hunter, "A scaled line-based kernel density estimator for the retrieval of utilization distributions and home ranges from GPS movement tracks," *Ecological informatics*, vol. 13, pp. 1–8, 2013.

[23] U. Demšar, K. Buchin, E. E. van Loon, and J. Shamoun-Baranes, "Stacked space-time densities: a geovisualisation approach to explore dynamics of space use over time," *GeoInformatica*, vol. 19, no. 1, pp. 85–115, 2015.

[24] L. Tang, Z. Kan, X. Zhang, F. Sun, X. Yang, and Q. Li, "A network kernel density estimation for linear features in space–time analysis of big trace data," *International Journal of Geographical Information Science*, vol. 30, no. 9, pp. 1717–1737, 2015.

[25] Y. Deng, B. Wallace, D. Maassen, and J. Werner, "A few GIS clarifications on tornado density mapping," *Journal of Applied Meteorology and Climatology*, vol. 55, no. 2, pp. 283–296, 2016.

[26] Y. L. Liu, G. R. Feng, Y. Xue, H. M. Zhang, and R. G. Wang, "Small-scale natural disaster risk scenario analysis: a case study from the town of Shuitou, Pingyang County, Wenzhou, China," *Natural Hazards*, vol. 75, no. 3, pp. 2167–2183, 2015.

[27] A. J. Colbert, B. J. Soden, and B. P. Kirtman, "The impact of natural and anthropogenic climate change on western North Pacific tropical cyclone tracks," *Journal of Climate*, vol. 28, no. 5, pp. 1806–1823, 2015.

[28] Esri, *What is ArcGIS?*, Esri, Redlands, CA, USA, June 2016, https://www.arcgis.com/features/index.html.

[29] J. F. Wang, T. L. Zhang, and B. J. Fu, "A measure of spatial stratified heterogeneity," *Ecological Indicators*, vol. 67, pp. 250–256, 2016.

[30] L. A. Waller and C. A. Gotway, *Applied Spatial Statistics for Public Health Data*, Wiley, Hoboken, NJ, USA, 2004.

[31] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, New York, NY, USA, 1986.

[32] A. Getis and J. K. Ord, "The analysis of spatial association by use of distance statistics," *Geographical Analysis*, vol. 24, no. 3, pp. 189–206, 1992.

[33] B. J. L. Berry and D. F. Marble, *Spatial Analysis: A Reader in Statistical Geography*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1968.

[34] F. C. Guan and Q. H. Xie, "The statistical characteristics of typhoon in the south China sea," *Marine Science Bulletin*, vol. 3, no. 4, pp. 21–29, 1984.

[35] N. Shi and J. D. Zhou, "A statistical analysis of typhoon activities over south China sea and ENSO," *Meteorological Monthly*, vol. 15, no. 4, pp. 9–14, 1989, in Chinese.

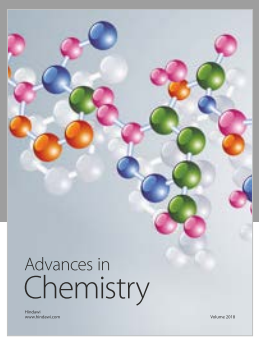
[36] D. S. Wu, X. Zhao, W. Z. Feng et al., "The statistical analysis to the local harmful typhoon of South China Sea," *Journal of Tropical Meteorology*, vol. 21, no. 3, pp. 309–314, 2005, in Chinese.

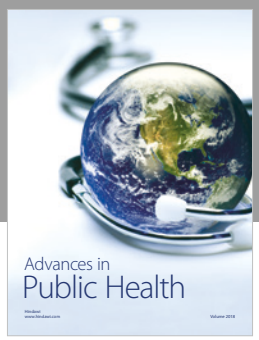
[37] P. J. Vickery and L. A. Twisdale, "Prediction of hurricane wind speeds in the United States," *Journal of Structural Engineering*, vol. 121, no. 11, pp. 1691–1699, 1995.

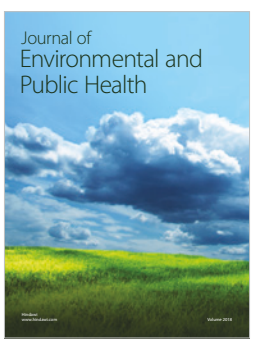
[38] K. Emanuel, S. Ravela, E. Vivant, and C. Risi, "A statistical deterministic approach to hurricane risk assessment," *Bulletin of the American Meteorological Society*, vol. 87, no. 3, pp. 299–314, 2006.

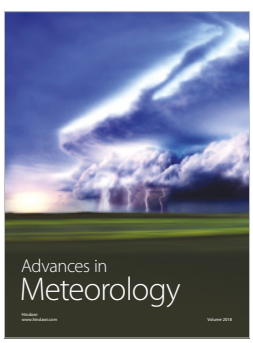
[39] H. Wang, Y. H. Ding, and J. H. He, "Influence of western north Pacific summer monsoon changes on typhoon genesis," *Acta Meteorologica Sinica*, vol. 64, no. 3, pp. 345–356, 2006.

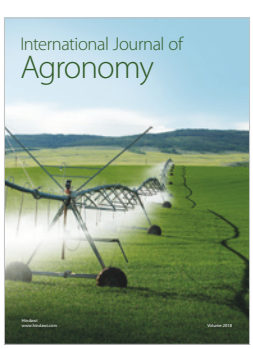
[40] R. H. Huang and G. H. Chen, "Research on interannual variations of tracks of tropical cyclones over Northwest Pacific and their physical mechanism," *Acta Meteorologica Sinica*, vol. 65, no. 5, pp. 683–694, 2007.

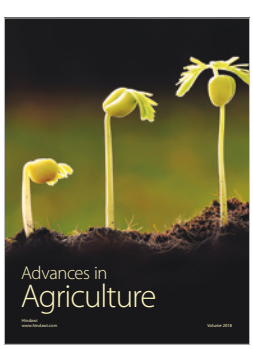


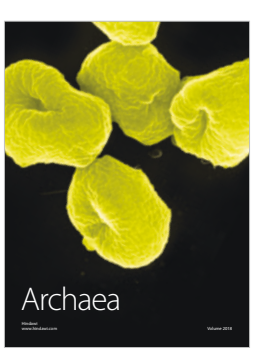


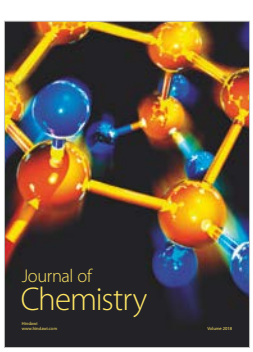













Hindawi

Submit your manuscripts at
www.hindawi.com