

Research Article

The Use of Geographically Weighted Regression for the Relationship among Extreme Climate Indices in China

Chunhong Wang,^{1,2} Jiangshe Zhang,^{1,2} and Xiaodong Yan³

¹ *School of Science, Xi'an Jiaotong University, Xi'an 710049, China*

² *State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China*

³ *Key Laboratory of Regional Climate-Environment in Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Science, Beijing 100029, China*

Correspondence should be addressed to Chunhong Wang, zqlwch@gmail.com

Received 9 May 2011; Accepted 2 September 2011

Academic Editor: Weihai Zhang

Copyright © 2012 Chunhong Wang 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.

The changing frequency of extreme climate events generally has profound impacts on our living environment and decision-makers. Based on the daily temperature and precipitation data collected from 753 stations in China during 1961–2005, the geographically weighted regression (GWR) model is used to investigate the relationship between the index of frequency of extreme precipitation (FEP) and other climate extreme indices including frequency of warm days (FWD), frequency of warm nights (FWN), frequency of cold days (FCD), and frequency of cold nights (FCN). Assisted by some statistical tests, it is found that the regression relationship has significant spatial nonstationarity and the influence of each explanatory variable (namely, FWD, FWN, FCD, and FCN) on FEP also exhibits significant spatial inconsistency. Furthermore, some meaningful regional characteristics for the relationship between the studied extreme climate indices are obtained.

1. Introduction

There is a general agreement that changes in the frequency or intensity of extreme climate events are likely to exert a much greater impact on nature and humanity than shifts in the mean value [1]. Starting from IPCC (1996) [2], many scientists have stressed the importance to study extreme climate events [3–5]. In the research field of extreme temperature and precipitation events, indices that are based on either fixed thresholds [6] or relative thresholds [7] are commonly used. To the best of our knowledge, most of the previous studies of climate

extremes mainly focus on some individual extreme climate index; however, the investigation of the relationship between them is relatively rare.

As for the relationship between some extreme climate indices, researchers generally assume that it is stationary over space and use an ordinary linear regression (OLR) model to analyze it. Nevertheless, it is known that an OLR model can only represent global relationship and it hardly takes into consideration the variations in relationships over space, in other words, the explicit incorporation of space or location has not been that commonly considered. In this context, there has been recently a surge focusing on the inclusion of spatial effects in climate models. A geographically weighted regression (GWR) model, which extends the traditional regression framework by allowing regression coefficients to vary with individual locations (spatial nonstationarity), is an effective method of utilizing spatial information to improve this issue [8–13]. Hence, GWR produces locally linear regression estimates for every point in space. For this purpose, weighted least squares methodology is used, with weights based on the distance between observations i and all the others in the sample. GWR allows the exploration of variation of the parameters as well as the testing of the significance of this variation. It is of great appeal to apply GWR technique to analyze spatial data in a number of areas such as geography econometrics, epidemiology, and environmental science [14–16].

China is strongly influenced by the East Asian monsoon [17]. During the winter half year, the climate is mostly cold and dry. Cold days and strong winds accompanied by dust storms are the major climate features particularly observed in northern China [18]. During the summer period, the rain belt moves gradually from south to north with the hot and humid climate in eastern China [19]. The regional characteristics of extreme climate are particularly prominent in China. The purpose of this paper is to analyze the spatially varying impacts of some temperature extreme indices on one precipitation extreme index in China. In this paper, relative thresholds based on the 1961–1990 base period were firstly used to build some extreme indices, namely, FEP (frequency of extreme precipitation), FWD (frequency of warm days), FWN (frequency of warm nights), FCD (frequency of cold days), and FCN (frequency of cold nights). The spatial distributions of these indices were then analyzed. In order to investigate the relationship among these indices, a GWR model was utilized to study how FEP was affected by the other indices. Moreover, two statistical tests were carried out to confirm some of our guesswork and some promising results were obtained.

The rest of the paper is organized as follows. Section 2 presents the data source, gives the definitions of extreme climate indices used in this paper, and briefly outlines the method of GWR. Results for annual mean extreme climate indices over China are displayed in Section 3. Section 4 provides a conclusion.

2. Data and Method

2.1. Experimental Data

The experimental data sets used in this paper consist of daily maximum and minimum temperatures and daily precipitation observed at 753 meteorological stations in China from January 1, 1961 to December 31, 2005, which were offered by National Meteorological Information Center in China Meteorological Administration. Because the study must rely on reliable data, the missing data in each month should be no more than three days. Therefore, the data collected from the 504 stations (Figure 1) which comply with this requirement

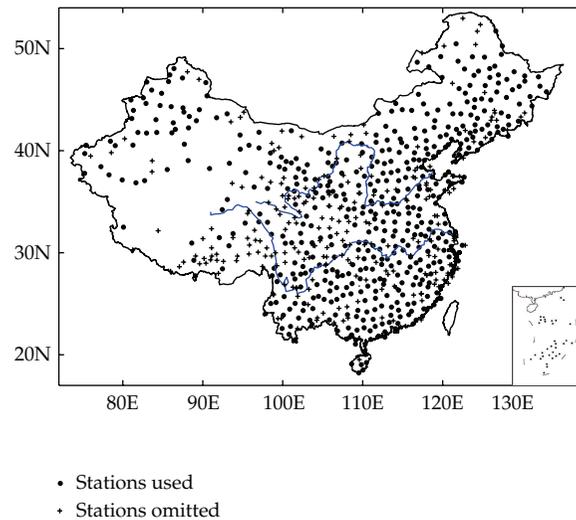


Figure 1: Stations for which data were available in China. (●) Stations used in this paper; (+) stations omitted due to excessive missing data.

were utilized in this work. With respect to the missing values in these 504 stations, a linear interpolation method was adopted to impute them.

2.2. Extreme Climate Indices

Numerous temperature indices have been used in previous studies of climate events. Some indices involved arbitrary thresholds, such as the number of hot days exceeding 35°C and summer days exceeding 25°C . As indicated by Manton et al. [5], these are suitable for regions with little spatial variability in climate, but arbitrary thresholds are inappropriate for regions spanning a broad range of climates. In China, climates vary widely from monsoon region in the eastern part to the westerly region in the northwestern part of the country, so there is no single temperature threshold that would be considered an event in all regions. For this reason, some studies have used weather and climate indices based on statistical quantities such as the 10th (5th) or 90th (95th) percentile [20, 21]; detailed information can be found from the European Climate Assessment & Dataset (ECA&D) Indices List (<http://www.knmi.nl/>). Upper and lower percentiles of temperature indices are used in all regions, but vary in absolute magnitude from site to site. A regional climate study in the Caribbean region using the same indices can also be found in [21].

As this study covers a broad region in China, climate indices chosen are based on the 10th and 90th percentiles. The extreme climate indices studied in this paper include FEP, FWD, FWN, FCD, and FCN whose definitions are described in detail in Table 1. As for the experimental data of these extreme indices based on the 1961–1990 base period, the relative values of them were calculated. For each station, the values for FEP, FWD, FWN, FCD, and FCN are their respective values averaged over the period 1961–2005, which are still denoted as FEP, FWD, FWN, FCD, and FCN in order to facilitate the following discussions.

Table 1: Five extreme climate indices calculated based on daily temperature and precipitation data.

Indicator name	Indicator definition (unit: days)
FEP	Let Tp_{ij} be the daily precipitation on day i of year j , and let Tp_{in90} be the calendar day 90th percentile centered on a 5-day window for the base period 1961–1990. Frequency of extreme precipitation (FEP) in year j is the annual count of days when $Tp_{ij} > Tp_{in90}$.
FWD	Let Tx_{ij} be the daily maximum temperature on day i of year j , and let Tx_{in90} be the calendar day 90th percentile centered on a 5-day window for the base period 1961–1990. Frequency of warm days (FWD) in year j is the annual count of days when $Tx_{ij} > Tx_{in90}$.
FWN	Let Tn_{ij} be the daily minimum temperature on day i of year j , and let Tn_{in90} be the calendar day 90th percentile centered on a 5-day window for the base period 1961–1990. Frequency of warm nights (FWN) in year j is the annual count of days when $Tn_{ij} > Tn_{in90}$.
FCD	Let Tx_{ij} be the daily maximum temperature on day i of year j , and let Tx_{in10} be the calendar day 10th percentile centered on a 5-day window for the base period 1961–1990. Frequency of cold days (FCD) in year j is the annual count of days when $Tx_{ij} < Tx_{in10}$.
FCN	Let Tn_{ij} be the daily minimum temperature on day i of year j , and let Tn_{in10} be the calendar day 10th percentile centered on a 5-day window for the base period 1961–1990. Frequency of cold nights (FCN) in year j is the annual count of days when $Tn_{ij} < Tn_{in10}$.

2.3. Geographically Weighted Regression (GWR)

The technique of linear regression estimates a parameter β that links the explanatory variables to the response variable. However, when this technique is applied to spatial data, some issues concerning the stationarity of these parameters over the space come out. In “normal” regression, it is generally assumed that the modeling relationship holds everywhere in the study area—that is, the regression parameters are “whole-map” statistics. In many situations this is not the case, however, as mapping the residuals (the difference between the observed and predicted data) may reveal. The realization in the statistical and geographical sciences that a relationship between an explanatory variable and a response variable in a linear regression model is not always constant across a study area has led to the development of regression models allowing for spatially varying coefficients. Many different solutions have been proposed for dealing with spatial variation in the relationship. One of them, developed by Brunsdon et al. [8], has been labelled geographically weighted regression (GWR), which provides an elegant and easily grasped means of modeling such relationships by subtly incorporating the spatial characteristics of data via allowing regression coefficients to depend on some covariates such as longitude and latitude of the meteorological stations. Specifically, it is a nonparametric model of spatial drift that relies on a sequence of locally linear regressions to produce estimates for every point in space by using a subsample of data information from nearby observations. That is to say, this technique allows the modeling of relationships that vary over space by introducing distance-based weights to provide estimates β_{ki} for each variable k and each geographical location i . Thus the spatial variation of regression relationship can be effectively analyzed and the inherent disciplines of spatial data by the estimated coefficients over different locations can be better understood.

An ordinary linear regression (OLR) model can be expressed by

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.1)$$

where y_i , $i = 1, 2, \dots, n$, are the observation of the response variable y , β_j ($j = 1, 2, \dots, p$) represents the regression coefficients, x_{ij} is the i th value of the explanatory variable x_j , and ε_i are normally distributed error terms with zero mean and constant variance.

In GWR model, the global regression coefficients are replaced by local parameters

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^p \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (2.2)$$

where (u_i, v_i) denotes the longitude and latitude coordinates of the i th meteorological station, $(y_i; x_{i1}, x_{i2}, \dots, x_{ip})$ represent the observed value of the response Y and explanatory variables X_1, X_2, \dots, X_p at (u_i, v_i) , $\beta_0(u_i, v_i)$ is the intercept, and $\beta_j(u_i, v_i)$ ($j = 1, 2, \dots, p$) are p unknown coefficient functions of spatial locations, which represent the strength and type of relationship that the j th explanatory variable X_j has to the response variable Y . Additionally, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are error terms which are generally assumed to be independent and identically distributed variables with mean 0 and common variance σ^2 . It is worth noticing that the OLR model is actually a special case of the GWR model where $\beta_j(u_i, v_i)$ are constant for all $i = 1, 2, \dots, n$.

The coefficient function vector $\hat{\beta}(u_i, v_i)$ for the i th observation in GWR can be estimated via the locally weighted least square procedure [22] as

$$\begin{aligned} \hat{\beta}(u_i, v_i) &= \left(\hat{\beta}_0(u_i, v_i), \hat{\beta}_1(u_i, v_i), \dots, \hat{\beta}_p(u_i, v_i) \right)^T \\ &= (X^T W(i) X)^{-1} X^T W(i) Y, \quad i = 1, 2, \dots, n, \quad (2.3) \\ \hat{\beta}_j(u, v) &= \left(\hat{\beta}_j(u_1, v_1), \hat{\beta}_j(u_2, v_2), \dots, \hat{\beta}_j(u_n, v_n) \right)^T, \quad j = 0, 1, 2, \dots, p, \end{aligned}$$

where

$$X = \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad (2.4)$$

$$W(i) = \text{diag}[K_h(d_{i1}), K_h(d_{i2}), \dots, K_h(d_{in})] \quad (2.5)$$

is a diagonal weight matrix, ensuring that observations near to the location have greater influence than those far away. Here, d_{ij} denotes the distance between two observed locations (u_i, v_i) and (u_j, v_j) , which can be calculated as

$$d_{ij} = R \arccos(\sin v_i \sin v_j + \cos v_i \cos v_j \cos(u_i - u_j)), \quad (2.6)$$

where R is the earth radius, namely, 6371 kilometers. In (2.5), $K_h(\cdot) = 1/hK(\cdot/h)$ with $K(\cdot)$ being Gaussian kernel function

$$K(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) \quad (2.7)$$

and h being the bandwidth which can be estimated by some data-driven procedures such as the cross-validation (CV) method [23], the generalized cross-validation (GCV) procedure [13], or the corrected Akaike information criterion (AIC_c) [24]. In this paper, the CV method utilized by [23] was employed to select the optimal h which was chosen to minimize

$$CV(h) = \sum_{i=1}^n (y_i - \hat{y}_{(i)}(h))^2, \quad (2.8)$$

where $\hat{y}_{(i)}(h)$ is the fitted value of y_i under bandwidth h with the observation at location (u_i, v_i) omitted from the fitting process.

Although GWR is very appealing in analyzing spatial nonstationarity, from the statistical viewpoint, two critical questions still remain. One is the goodness-of-fit test, that is, a OLR model is compared to a GWR model to see which one provides the best fit. Usually, a GWR model can fit a given data set better than an OLR model. However, the simpler a model, the easier it can be applied and interpreted in practice. If a GWR model does not perform significantly better than an OLR model, it means that there is no significant drift in any of the model parameters. Thus, we will prefer an OLR model in practice. On the other hand, if a GWR model significantly outperforms an OLR model, we will be concerned with the second question, that is, whether each coefficient function estimate $\hat{\beta}_j(u, v)$ ($j = 1, 2, \dots, p$) exhibits significant spatial variation over the studied area [11, 25]. If the answer to this question is positive, the characteristics of the data will be investigated in more details.

To compare the goodness-of-fit of a GWR model and an OLR model, a simplified procedure is summarized as follows.

(1) Formulate the hypothesis

$$\left(\begin{array}{l} H_0 : Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \varepsilon, \\ H_1 : Y = \beta_0(u, v) + \beta_1(u, v) X_1 + \dots + \beta_p(u, v) X_p + \varepsilon \end{array} \right). \quad (2.9)$$

(2) Construct the test statistic

$$F = \frac{Y^T(I - H)Y - Y^T(I - L)^T(I - L)Y}{Y^T(I - L)^T(I - L)Y}. \quad (2.10)$$

Here, $H = X(X^T X)^{-1} X^T$, I is an identity matrix of order n , and

$$L = \begin{pmatrix} x_1^T (X^T W(1) X)^{-1} X^T W(1) \\ x_2^T (X^T W(2) X)^{-1} X^T W(2) \\ \vdots \\ x_n^T (X^T W(n) X)^{-1} X^T W(n) \end{pmatrix} \quad (2.11)$$

is an $n \times n$ matrix. If H_0 is true, the test statistic F is to be

$$F = \frac{\boldsymbol{\varepsilon}^T \left((I - H) - (I - L)^T (I - L) \right) \boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}^T (I - L)^T (I - L) \boldsymbol{\varepsilon}}. \quad (2.12)$$

(3) Test the hypothesis. The p value should be calculated as

$$p_0 = P_{H_0}(F > F_0) = P_{H_0} \left(\boldsymbol{\varepsilon}^T \left((I - H) - (1 + F_0)(I - L)^T (I - L) \right) \boldsymbol{\varepsilon} > 0 \right), \quad (2.13)$$

where F_0 is the observed value of F in (2.12). Since it is difficult to derive the null distribution of F theoretically, the three-moment χ^2 approximation procedure [26, 27] devoted to approximate the distribution of normal variable quadratic form such as $\boldsymbol{\varepsilon}^T \left((I - H) - (1 + F_0)(I - L)^T (I - L) \right) \boldsymbol{\varepsilon}$ was used to compute the p value defined in (2.13). Given a significance level α , if $p_0 < \alpha$, the null hypothesis should be rejected. Otherwise, we may conclude that the GWR model cannot improve the fitness significantly in comparison with the OLR model.

In order to test whether each coefficient function estimate $\hat{\beta}_j(u, v)$ ($j = 1, 2, \dots, p$) exhibits significant variation over the studied area, we employed the method developed by [12] to achieve the goal. The main steps of it are summarized as follows.

(a) For a given k ($k = 0, 1, 2, \dots, p$), formulate the hypothesis

$$\left(\begin{array}{l} H_{0k} : \beta_k(u_1, v_1) = \beta_k(u_2, v_2) = \dots = \beta_k(u_n, v_n) = \beta_k \\ H_{1k} : \text{not all } \beta_k(u_i, v_i) \text{ (} i = 1, 2, \dots, n \text{) are equal} \end{array} \right). \quad (2.14)$$

(b) Construct the test statistic

$$T_k = \frac{Y^T B^T (I - (1/n)11^T) B Y}{Y^T (I - L)^T (I - L) Y}. \quad (2.15)$$

Here, $\mathbf{1}$ is an $n \times 1$ column vector with unity for each element, and

$$B = \begin{pmatrix} e_{k+1}^T (X^T W(1) X)^{-1} X^T W(1) \\ e_{k+1}^T (X^T W(2) X)^{-1} X^T W(2) \\ \vdots \\ e_{k+1}^T (X^T W(n) X)^{-1} X^T W(n) \end{pmatrix}. \quad (2.16)$$

e_{k+1} is an $n \times 1$ column vector which takes value 1 for the $(k + 1)$ th element and zero for the other elements. Under the null hypothesis H_{0k} , the test statistic T_k is simplified as

$$T_k = \frac{\boldsymbol{\varepsilon}^T B^T (I - (1/n)\mathbf{1}\mathbf{1}^T) B \boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}^T (I - L)^T (I - L) \boldsymbol{\varepsilon}}. \quad (2.17)$$

(c) Test the hypothesis. The p value is

$$p_k = P_{H_{0k}}(T_k > T_{0k}) = P_{H_{0k}}\left(\boldsymbol{\varepsilon}^T \left(B^T \left(I - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right) B - T_{0k}(I - L)^T (I - L)\right) \boldsymbol{\varepsilon} > 0\right), \quad (2.18)$$

where T_{0k} is the observed value of T_k in (2.17). Similar to the goodness-of-fit test, the three-moment χ^2 approximation procedure was used to derive the p value defined in (2.18). Given a significance level α , if $p_k < \alpha$, reject H_{0k} ; accept H_{0k} otherwise.

3. Analysis of Results

In this part, we will carry out numerical experiments for the OLR model and GWR model. All programs are written in Matlab.

3.1. Spatial Distributions of Extreme Climate Indices

Based on the values of FWD, FWN, FCD, FCN, and FEP, Figure 2 presents the spatial distributions for each of them over the 504 stations in China.

As shown in Figure 2, FWD, FWN, FCD, FCN, and FEP exhibit some regional features. Generally, there are 16 to 29 times per year for FWD and the larger values for FWD are mainly located in the north as well as the east of China. There are 18–35 times per year for FWN. If using the Yangtze River as the boundary, FWN values in the north are generally larger than those in the south. As for FCD, there are 14 to 26 times per year. Specially, FCD has small values about 14–18 times per year in most parts of northwest China. With regard to FCN, it is about 13–28 times per year and it has small values in southern China. Furthermore, FEP values are between 9 and 33 times per year. In most of the country, its value varies from 25

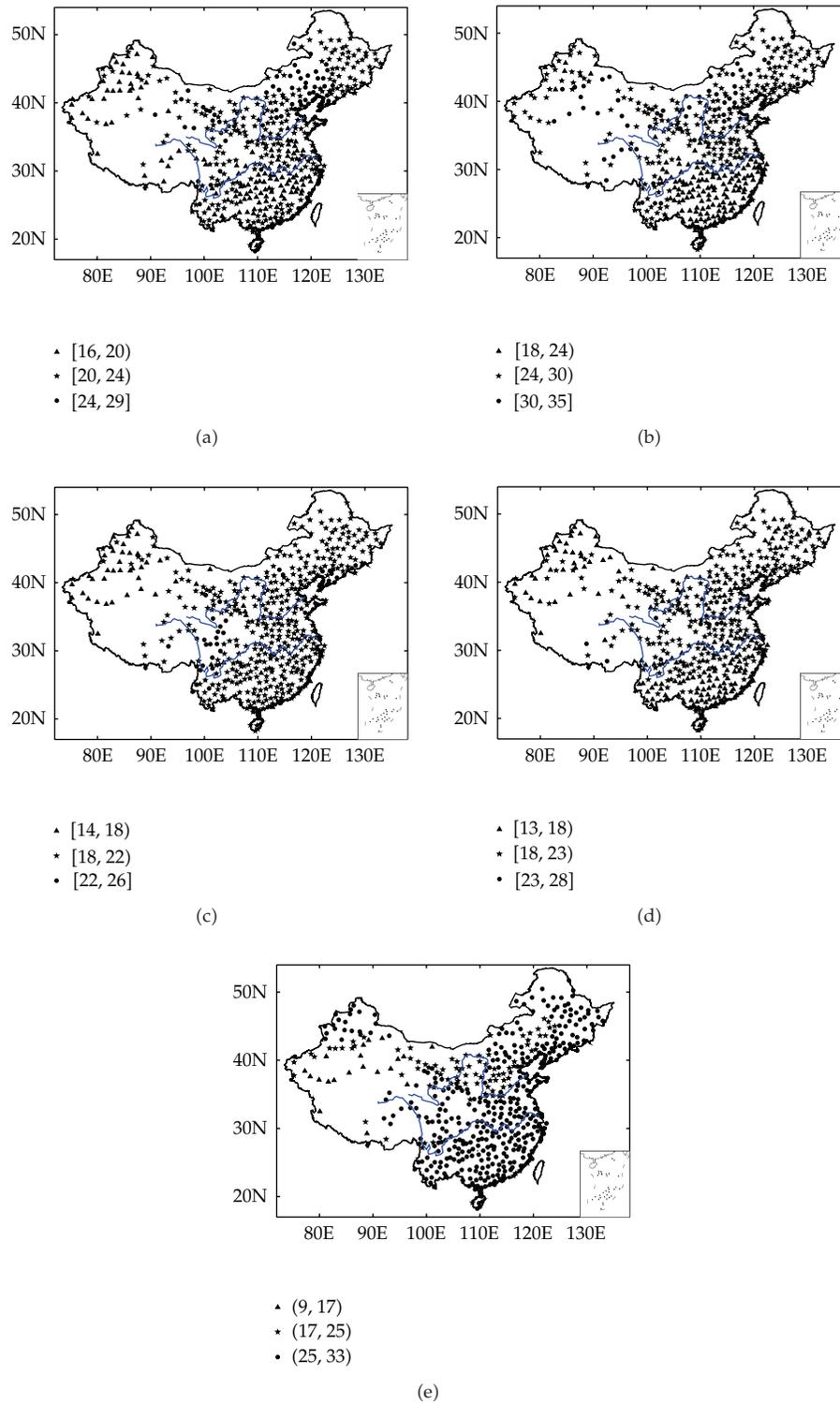


Figure 2: Spatial distributions of the considered extreme climate indices ((a) FWD, (b) FWN, (c) FCD, (d) FCN, and (e) FEP) over the 504 stations in China.

Table 2: Correlation coefficients of the independent variables, that is, FWD, FWN, FCD, and FCN.

	FWD	FWN	FCD	FCN
FWD	1.0000	0.3862	0.3453	0.1836
FWN		1.0000	0.1318	0.1329
FCD			1.0000	0.4174
FCN				1.0000

Table 3: Correlation coefficients of the GWR coefficient estimates, that is, $\hat{\beta}_0(u, v)$, $\hat{\beta}_1(u, v)$, $\hat{\beta}_2(u, v)$, $\hat{\beta}_3(u, v)$, and $\hat{\beta}_4(u, v)$.

	$\hat{\beta}_0(u, v)$	$\hat{\beta}_1(u, v)$	$\hat{\beta}_2(u, v)$	$\hat{\beta}_3(u, v)$	$\hat{\beta}_4(u, v)$
$\hat{\beta}_0(u, v)$	1.0000	-0.4068	-0.1734	-0.5044	0.0828
$\hat{\beta}_1(u, v)$		1.0000	0.0491	-0.3281	-0.0366
$\hat{\beta}_2(u, v)$			1.0000	-0.4295	0.6309
$\hat{\beta}_3(u, v)$				1.0000	-0.6375
$\hat{\beta}_4(u, v)$					1.0000

to 33 times per year, and only in some stations in southern Xinjiang and Tibet, its values lie between 9 and 17 times per year.

3.2. The Fitted Geographically Weighted Regression Model

In order to make clear the relationship among these extreme climate indices in 504 stations in China so that some useful information can be provided to decision-makers to help them to deduce the disaster caused by extreme weather, a GWR model was fitted by considering FEP as the response variable Y and FWD, FWN, FCD, and FCN as the explanatory variables (X_1, X_2, X_3, X_4) , respectively. Letting n be equal to 504 and p equal to 4 and letting $(y_i, x_{i1}, x_{i2}, x_{i3}, x_{i4})$ be the observations of the variables (Y, X_1, X_2, X_3, X_4) at the location (u_i, v_i) , the model (2.2) can be expressed as

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^4 \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, 504, \quad (3.1)$$

based on the data collected from the 504 stations.

When we apply a fixed Gaussian function, the minimum score of (2.8) is obtained when the bandwidth h equals approximately 240 km. Thus, the weighting matrix $W(i)$ is estimated, where $w_{ij} = (1/(240 * \sqrt{2\pi})) \exp(-d_{ij}^2 / (2 * 240^2))$. Based on (2.3), $\hat{\beta}_j(u, v)$ ($j = 0, 1, 2, 3, 4$) are calculated by the locally weighted least square approach. Hence, the strength and type of relationship that FWD (FWN, FCD, FCN) has with FEP over 504 stations in China can be studied.

Because Wheeler [28–30] raised the multicollinearity issues, correlation coefficients of the independent variables as well as that of the GWR coefficient estimates were presented in Tables 2 and 3, respectively.

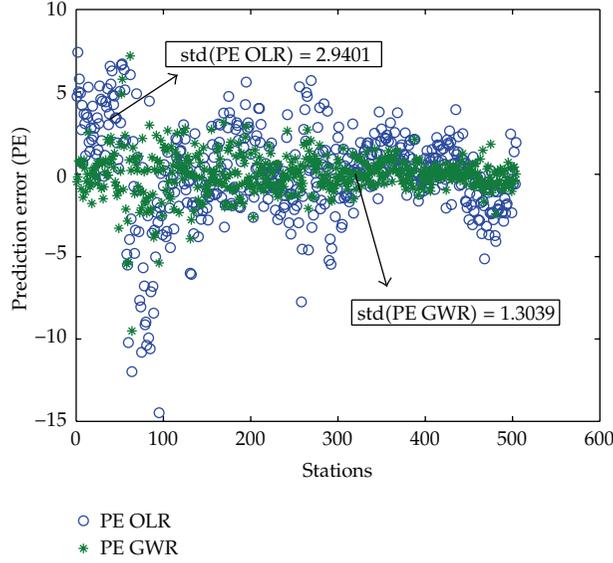


Figure 3: Prediction error (PE) of the responsible variable, FEP, for ordinary linear regression (OLR) and geographically weighted regression (GWR) over the 504 stations in China.

As shown in Tables 2 and 3, correlation coefficients of the independent variables as well as that of the GWR coefficient estimates are all not large, except for that between $\hat{\beta}_2(u, v)$ and $\hat{\beta}_4(u, v)$, as well as $\hat{\beta}_3(u, v)$ and $\hat{\beta}_4(u, v)$, whose absolute values are more than 0.5. It indicates that $\hat{\beta}_4(u, v)$ has a positive correlation with $\hat{\beta}_2(u, v)$, while it has a negative correlation with $\hat{\beta}_3(u, v)$. We ignore the correlation between the independent variables in this paper.

After conducting the goodness-of-fit test, the computed p value is smaller than the significance level 0.05. Thus, the GWR model can describe the regression relationship significantly better than the OLR model and it indicates that the relationship between FEP and FWD, FWN, FCD, and FCN has spatial nonstationarity. Define

$$R^2 = \frac{\sum_{i=1}^{504} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{504} (y_i - \bar{y})^2} \quad (3.2)$$

to measure the goodness of fit of the regression relationship on the given data set. The R^2 values for the OLR and GWR model are 0.3953 and 0.7750, respectively, which indicates that the GWR model can capture a larger amount (77.50%) of variance of FEP based on the climate indices FWD, FWN, FCD, and FCN, than the OLR model. The prediction errors (i.e., residual errors) for the OLR and GWR model are presented in Figure 3, which shows the prediction error of the GWR model and its standard error are both lower than that of the OLR model.

Furthermore, the statistical significance tests for the variations of the coefficient functions are carried out. The obtained results show that all the regression coefficient estimates

Table 4: p value of relevant tests for the GWR model (3.1).

Global stationarity for regression relationship	Significance for $\hat{\beta}_0(u, v)$	Significance for $\hat{\beta}_1(u, v)$	Significance for $\hat{\beta}_2(u, v)$	Significance for $\hat{\beta}_3(u, v)$	Significance for $\hat{\beta}_4(u, v)$
0	$1.4816 * 10^{-6}$	0	0.0018	0	$1.5184 * 10^{-6}$

$\hat{\beta}_j(u, v)$ ($j = 0, 1, 2, 3, 4$) vary significantly with the locations, that is, the influence of each explanatory variable (viz., FWD, FWN, FCD, and FCN) on the response variable FEP has spatial inconsistency. All p values of relevant tests for the GWR model (3.1) are presented in Table 4.

In order to visualize these spatial inconsistencies, Figure 4 shows geographic distributions of the estimated GWR coefficient functions in China. As there is not much meaning of $\hat{\beta}_0(u_i, v_i)$, the plot of it is omitted here.

Figure 4(a) shows that the values of $\hat{\beta}_1(u, v)$ are between -3.5 and 2.6 . Negative values of $\hat{\beta}_1(u, v)$ can be observed in most of mainland China, and the most largest absolute values are located in the northern and western parts of the Xinjiang region. Few stations with positive values of $\hat{\beta}_1(u, v)$ are concentrated in the southern part of Tibet, Gansu, Chongqing, and the eastern part of north China and east China.

As Figure 4(b) manifests, the values of $\hat{\beta}_2(u, v)$ are between -1.3 and 0.28 , and some stations with positive values of $\hat{\beta}_2(u, v)$ are concentrated in Jilin, northern inner Mongolia, eastern coast and Hainan. However, for China as a whole, it is obvious that many areas show negative values, especially in the Xinjiang, Tibet region as well as the middle Yellow River valley and the southern part of Northeast China.

From Figure 4(c), it can be seen that the values of $\hat{\beta}_3(u, v)$ are between -1.3 and 4.6 . Its value is positive in most parts of the country, and it is larger in western China than in eastern China. Scattered stations with negative values can be found in the northern part of inner Mongolia and south China, especially concentrated in Yunnan and Guangdong Province.

As for $\hat{\beta}_4(u_i, v_i)$, it can be found in Figure 4(d) that its values are between -2.3 and 0.57 . Negative values occur in the western China and center China, while in the north of the northeast China, north of north China and south China, positive values can be found.

On the basis of the above analysis, some regional characteristics for the relationship between the studied extreme climate indices can be observed. In western China, FEP increases with the increase of FCD, while it decreases with the increase of FWD, FWN, and FCN. In southern China, FEP increases with the increase of FCN, while it decreases with the increase of FWD, FWN, and FCD. In the northern part of northeast China, FEP increases with the increase of FCD and FCN, while it decreases with the increase of FWD and FWN. The impacts of FCN and FCD on the FEP are roughly the opposite over almost all China.

4. Conclusions

Based on the Chinese daily temperature and precipitation data collected at 753 meteorological stations from 1961 to 2005, the relationship among the numbers of days that experience extreme temperature or precipitation events (i.e., FEP, FWD, FWN, FCD, and FCN) is

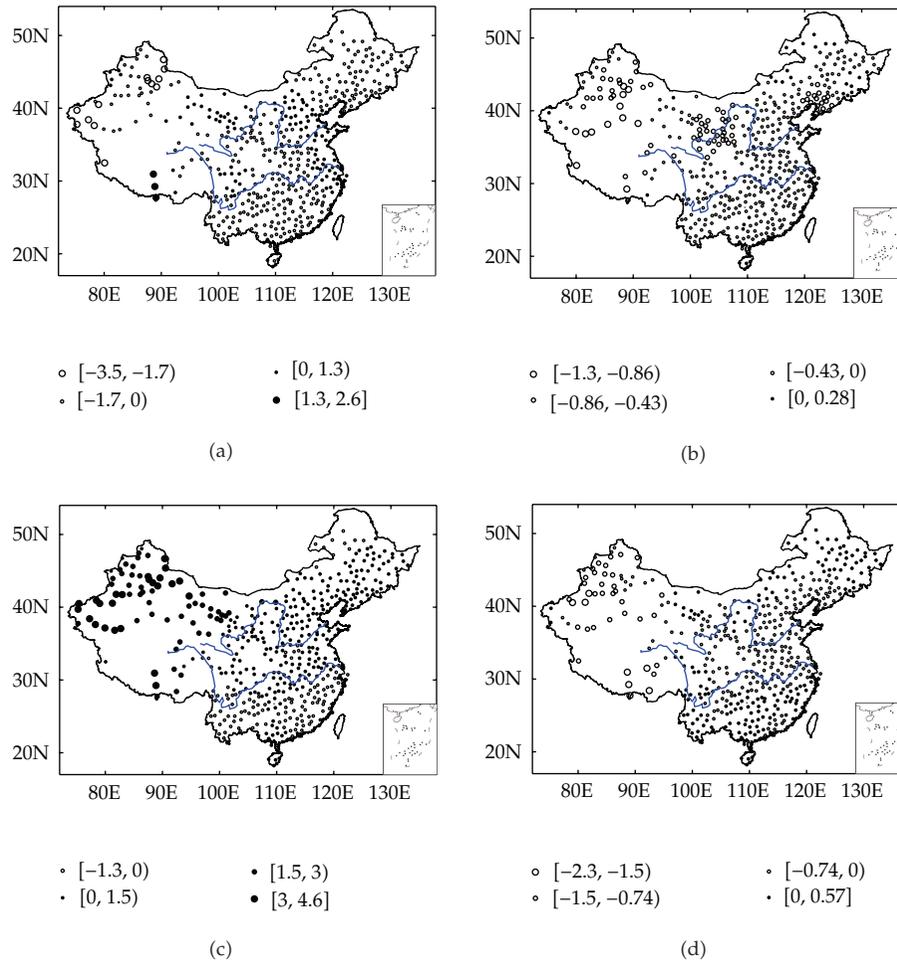


Figure 4: Geographic distributions of the estimated GWR coefficient functions ((a) $\hat{\beta}_1(u, v)$, (b) $\hat{\beta}_2(u, v)$, (c) $\hat{\beta}_3(u, v)$, and (d) $\hat{\beta}_4(u, v)$) over the 504 stations in China.

investigated by a GWR model and their spatial distributions in China. The main conclusions can be summarized as follows.

- (1) FWD, FWN, FCD, FCN and FEP exhibit different spatial variations. There are larger values about 24–29 times per year for FWD mainly in northeast China. In the north of the Yangtze River, FWN has larger values of 24–35 times per year. FCD has larger values about 18–26 times per year in most part of China but northwest China. As for FCN, most of China has larger values about 18–28 times except for the south. Except in some stations in southern Xinjiang and Tibet, FEP has larger values of 17–33 times per year.
- (2) With respect to how FWD, FWN, FCD, and FCN affect FEP, the GWR model is significantly superior to the OLR model at the significance level 0.05. Furthermore,

the statistical tests indicate that the influence of each explanatory variable (viz., FWD, FWN, FCD, and FCN) on FEP has spatial inconsistency.

- (3) Some regional features are detected for the relationship between the studied extreme climate indices. In western China, FCD has a positive effect on FEP, which is contrary to that of FWD, FWN, and FCN. However, it is just the opposite in southern China. The effects of FCD as well as FCN on FEP are positive in the northern part of Northeast China, while those of FWD and FWN are negative. Meanwhile, FCN and FCD have the opposite influence on FEP over most of China.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant nos. 60675013, 10531030) and the National Basic Research Program of China (973 Program) (Grant no. 2007CB311002).

References

- [1] R. W. Katz and B. G. Brown, "Extreme events in a changing climate: variability is more important than averages," *Climatic Change*, vol. 21, no. 3, pp. 289–302, 1992.
- [2] J. T. Houghton, L. G. M. Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, UK, 1996.
- [3] T. R. Karl, N. Nicholls, and A. Ghazi, "CLIVAR/GCOS/WMO Workshop on Indices and Indicators for climate extremes," *Climatic Change*, vol. 42, no. 1, pp. 3–7, 1999.
- [4] C. K. Folland, C. Miller, D. Bader et al., "Workshop on indices and indicators for climate extremes, asheville, NC, USA, 3–6 June 1997, breakout group C: temperature indices for climate extremes," *Climatic Change*, vol. 42, no. 1, pp. 31–41, 1999.
- [5] M. J. Manton, P. M. Della-Marta, M. R. Haylock et al., "Trends in extreme daily rainfall and temperature in southeast Asia and the south Pacific: 1961–1998," *International Journal of Climatology*, vol. 21, no. 3, pp. 269–284, 2001.
- [6] B. D. Su, T. Jiang, and W. B. Jin, "Recent trends in observed temperature and precipitation extremes in the Yangtze River basin, China," *Theoretical and Applied Climatology*, vol. 83, no. 1–4, pp. 139–151, 2006.
- [7] P. Zhai, A. Sun, F. Ren, X. Liu, B. Gao, and Q. Zhang, "Changes of climate extremes in China," *Climatic Change*, vol. 42, no. 1, pp. 203–218, 1999.
- [8] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton, "Geographically weighted regression: a method for exploring spatial nonstationarity," *Geographical Analysis*, vol. 28, no. 4, pp. 281–298, 1996.
- [9] A. S. Fotheringham, M. Charlton, and C. Brunsdon, "The geography of parameter space: an investigation of spatial non-stationarity," *International Journal of Geographical Information Systems*, vol. 10, no. 5, pp. 605–627, 1996.
- [10] C. Brunsdon, S. Fotheringham, and M. Charlton, "Geographically weighted regression: modelling spatial non-stationarity," *Journal of the Royal Statistical Society Series D*, vol. 47, no. 3, pp. 431–443, 1998.
- [11] Y. Leung, C. L. Mei, and W. X. Zhang, "Statistical tests for spatial nonstationarity based on the geographically weighted regression model," *Environment and Planning A*, vol. 32, no. 1, pp. 9–32, 2000.
- [12] Y. Leung, C. L. Mei, and W. X. Zhang, "Testing for spatial autocorrelation among the residuals of the geographically weighted regression," *Environment and Planning A*, vol. 32, no. 5, pp. 871–890, 2000.
- [13] A. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, Wiley, Chichester, UK, 2002.
- [14] V. Huang and Y. Leung, "Analysing regional industrialisation in Jiangsu province using geographically weighted regression," *Journal of Geographical Systems*, vol. 4, no. 2, pp. 233–249, 2002.

- [15] P. A. Longley and C. Tobón, "Spatial dependence and heterogeneity in patterns of hardship: an intra-urban analysis," *Annals of the Association of American Geographers*, vol. 94, no. 3, pp. 503–519, 2004.
- [16] T. Nakaya, "Local spatial interaction modelling based on the geographically weighted regression approach," *GeoJournal*, vol. 53, no. 4, pp. 347–358, 2001.
- [17] W. Qian and Y. Zhu, "Climate change in China from 1880 to 1998 and its impact on the environmental condition," *Climatic Change*, vol. 50, no. 4, pp. 419–444, 2001.
- [18] W. Qian, L. Quan, and S. Shi, "Variations of the dust storm in China and its climatic control," *Journal of Climate*, vol. 15, no. 10, pp. 1216–1229, 2002.
- [19] W. Qian, H. S. Kang, and D. K. Lee, "Distribution of seasonal rainfall in the East Asian monsoon region," *Theoretical and Applied Climatology*, vol. 73, no. 3-4, pp. 151–168, 2002.
- [20] T. C. Peterson, M. A. Taylor, R. Demeritte et al., "Recent changes in climate extremes in the Caribbean region," *Journal of Geophysical Research D*, vol. 107, no. 21, p. 4601, 2002.
- [21] N. Plummer, M. J. Salinger, N. Nicholls et al., "Changes in climate extremes over the Australian region and New Zealand during the twentieth century," *Climatic Change*, vol. 42, no. 1, pp. 183–202, 1999.
- [22] C. Brunsdon, A. S. Fotheringham, and M. Charlton, "Some notes on parametric significance tests for geographically weighted regression," *Journal of Regional Science*, vol. 39, no. 3, pp. 497–524, 1999.
- [23] L. Desmet and I. Gijbels, "Local linear fitting and improved estimation near peaks," *The Canadian Journal of Statistics*, vol. 37, no. 3, pp. 453–475, 2009.
- [24] C. M. Hurvich, J. S. Simonoff, and C.-L. Tsai, "Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion," *Journal of the Royal Statistical Society B*, vol. 60, no. 2, pp. 271–293, 1998.
- [25] C. Brunsdon, M. Aitkin, S. Fotheringham, and M. Charlton, "A comparison of random coefficient modelling and geographically weighted regression for spatially non-stationary regression problems," *Geographical and Environmental Modelling*, vol. 3, no. 1, pp. 47–62, 1999.
- [26] C. L. Mei and W. X. Zhang, "Testing linear regression relationships via locally-weighted-fitting technique," *Journal of Systems Science and Mathematical Sciences*, vol. 22, no. 4, pp. 467–480, 2002.
- [27] C.-h. Wei and X.-z. Wu, "Testing linearity for nonparametric component of partially linear models," *Mathematica Applicata*, vol. 20, no. 1, pp. 183–190, 2007.
- [28] D. Wheeler and M. Tiefelsdorf, "Multicollinearity and correlation among local regression coefficients in geographically weighted regression," *Journal of Geographical Systems*, vol. 7, no. 2, pp. 161–187, 2005.
- [29] D. C. Wheeler, "Diagnostic tools and a remedial method for collinearity in geographically weighted regression," *Environment and Planning A*, vol. 39, no. 10, pp. 2464–2481, 2007.
- [30] D. C. Wheeler, "Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso," *Environment and Planning A*, vol. 41, no. 3, pp. 722–742, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

