



Article

# Spatial Characteristics of Life Expectancy and Geographical Detection of Its Influencing Factors in China

Yafei Wu <sup>1,2,3,†</sup>, Ke Hu <sup>1,2,†</sup>, Yaofeng Han <sup>1,2</sup>, Qilin Sheng <sup>1,2</sup> and Ya Fang <sup>1,2,3,\*</sup>

<sup>1</sup> State Key Laboratory of Molecular Vaccine and Molecular Diagnostics, School of Public Health, Xiamen University, Xiamen 361102, China; wyfyyahcx@163.com (Y.W.); huke2296991140@163.com (K.H.); yaofenghan@163.com (Y.H.); shengqilin5@163.com (Q.S.)

<sup>2</sup> Key Laboratory of Health Technology Assessment of Fujian Province, School of Public Health, Xiamen University, Xiamen 361102, China

<sup>3</sup> National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen 361102, China

\* Correspondence: fangya@xmu.edu.cn

† These authors contributed equally to this work.

Received: 30 December 2019; Accepted: 29 January 2020; Published: 1 February 2020

**Abstract:** Life expectancy (LE) is a comprehensive and important index for measuring population health. Research on LE and its influencing factors is helpful for health improvement. Previous studies have neither considered the spatial stratified heterogeneity of LE nor explored the interactions between its influencing factors. Our study was based on the latest available LE and social and environmental factors data of 31 provinces in 2010 in China. Descriptive and spatial autocorrelation analyses were performed to explore the spatial characteristics of LE. Furthermore, the Geographical Detector (GeoDetector) technique was used to reveal the impact of social and environmental factors and their interactions on LE as well as their optimal range for the maximum LE level. The results show that there existed obvious spatial stratified heterogeneity of LE, and LE mainly presented two clustering types (high–high and low–low) with positive autocorrelation. The results of GeoDetector showed that the number of college students per 100,000 persons (NOCS) could mainly explained the spatial stratified heterogeneity of LE (Power of Determinant ( $PD$ ) = 0.89,  $p < 0.001$ ). With the discretization of social and environmental factors, we found that LE reached the highest level with birth rate, total dependency ratio, number of residents per household and water resource per capita at their minimum range; conversely, LE reached the highest level with consumption level, GDP per capita, number of college students per 100,000 persons, medical care expenditure and urbanization rate at their maximum range. In addition, the interaction of any two factors on LE was stronger than the effect of a single factor. Our study suggests that there existed obvious spatial stratified heterogeneity of LE in China, which could mainly be explained by NOCS.

**Keywords:** life expectancy; spatial characteristics; social and environmental factors; spatial stratified heterogeneity; Geographical Detector

## 1. Introduction

Life expectancy (LE) is a comprehensive and important index for measuring population health, which is vital for policy development and health improvement [1]. In recent years, the LE has continued to rise with the development of economy and living standard in China. From 2000 to 2017, it increased from 71.40 to 76.47 years old [2]. However, LE showed obvious spatial differences in China. According to a research, the difference of LE was larger than 10 years between the east and west in China in 2010 [3]. Therefore, it is an urgent task to explore the spatial characteristics and influencing factors of LE in order to deal with the regional inequality of LE in China.

There have been many studies concerning the various influencing factors of LE. These factors can mainly be divided into two categories: biological and social and environmental factors [4]. Biological factors involve the individual's genetic factors, living habits, etc. Differences in LE between men and women are mainly due to differences in gene expression intensity. In addition, a study also found that people who smoked, drank and exercised less were vulnerable to cancer, cardiovascular and cerebrovascular diseases, which caused a reduction of LE [5]. However, social and environmental factors were the most frequently considered and important influencing factors in LE research [6]. Many studies found that economy and demographic composition were key influencing factors of LE [7–10]. Studies have found that health care and services had a positive effect on disease prevention and treatment, so the LE level was often higher in areas with good medical conditions [11]. People with a higher education level tend to have better health awareness, so the LE of these individuals is higher. The emergence of education differences would enlarge the gap of LE among different populations [12]. Moreover, researchers also found that other social and environmental factors, such as environmental resources, had some impact on LE [13]. However, these studies were mainly based on few social and environmental factors, and they rarely explored the effects of multi-dimensional factors and their interactions from a spatial perspective.

Multiple linear regression (MLR) is the most widely used method for traditional studies of LE [14], while MLR has not yet considered the spatial information contained in data and can not effectively deal with multiple collinearity and interactions between variables. Furthermore, global and local spatial regression models, such as Spatial Error Model (SEM), Spatial Lag Model (SLM) and Geographical Weighted Regression (GWR), have been applied to LE research [15]. Although the above spatial methods can deal with the spatial autocorrelation and heterogeneity of LE and independent variables, the spatial stratified heterogeneity of LE have not been explored yet. As a result, the Geographical Detector technique has been put forward and developed rapidly in recent years. The geographical detector is a specific tool developed by Wang for spatial stratified heterogeneity study without any assumptions on the distribution of dependent and independent variables [16,17]. It can detect the spatial stratified heterogeneity of dependent variables and reveal the driving force behind it with its great flexibility [18]. It has been widely used in various fields, such as natural and social sciences [19–21], environmental pollution [22,23] and disease risk detection [16,24], etc., but it is rarely used in the field of LE research.

Therefore, we analyzed the spatial distribution characteristics of LE by descriptive methods and spatial autocorrelation analysis. Then we used the Geographical Detector technique to reveal the impact of social and environmental factors and their interactions on LE as well as their optimal range for the maximum LE level, providing reference for the research on LE, economic and educational development, utilization of medical resources, environmental protection as well as population management policy-making in China.

## 2. Materials and Methods

### 2.1. Data

In order to explore the impact of social and environmental factors on LE in China, we selected 11 representative social and environmental factors (Table 1) classified to 7 categories. We performed an analysis at the provincial level, including 31 provinces in mainland China. The LE data in 2010 came from China Statistical Yearbooks [25]. The data of TDR (Total dependency ratio) in 2010 was obtained from China Population & Employment Statistics Yearbooks [26], and the data of MCEOUR (Medical care expenditure of urban residents) in 2010 came from China Health Statistics Yearbooks [27]. The data of other social and environmental factors in 2010 were all from China Statistics Yearbooks [28,29].

**Table 1.** Social and environmental factors in the study.

Categories 1	Categories 2	Variables	Abbreviation
Population structure	—	Birth rate	BR
Family living standard	Raising burden	Total dependency ratio	TDR
	Family size	Number of residents per household	NORPH
Environmental resources	—	Water resource per capita	WRPC
Economic development	Residents' living	Consumption level of urban residents	CLOUR
	Economic level	GDP per capita	GPC
Education level	—	Number of college students per 100,000 persons	NOCS
Medical level	Medical expenditure	Medical care expenditure of urban residents	MCEOUR
	Medical resources	Number of hospital beds per 1000 persons	NOHB
		Number of doctors per 1000 persons	NOD
Other	—	Urbanization rate	UR

## 2.2. Statistical Methods

### 2.2.1. Descriptive Methods

In order to describe the spatial distribution of LE and social and environmental factors, the LE and main social and environmental factors of each province were mapped (the value was represented by the color depth) based on the provincial scale map of mainland China.

### 2.2.2. Spatial Autocorrelation Analysis

In order to analyze the spatial clustering characteristics of LE in China, we conducted the global and local spatial autocorrelation by using Moran's index (Moran's I).

#### Global Spatial Autocorrelation Analysis

Global Moran's I is used to judge whether there exists spatial autocorrelation in LE in the whole area. Moran's I ranges from −1 to 1. If it is greater than 0, spatial autocorrelation exists in LE; if it equals to 0, no spatial autocorrelation exists in LE, if it is less than 0, spatial discreteness exists in LE. Global Moran's I can be expressed as follows [30]:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

where  $n$  represents the number of spatial units (provincial administrative unit in the study);  $x_i$  and  $x_j$  is the LE of the  $i$  and  $j$  province, respectively;  $\bar{x}$  is the average of the LE, and  $W_{ij}$  is the spatial weight matrix based on the inverse distance as follows:

$$W_{ij} = d_{ij}^{-\alpha} \quad (2)$$

where  $d_{ij}$  refers to the distance between two provinces,  $\alpha$  is the appropriate constant (such as 1 or 2).

#### Local Spatial Autocorrelation Analysis

We used local Moran's I (*LISA*) to detect the spatial cluster types of LE further. And there are 4 types: high–high (high value is surrounded by high value), low–low (low value is surrounded by low

value), high–low (high value is surrounded by low value) and low–high (low value is surrounded by high value) [31]. The local Moran's I is calculated as follows [32]:

$$I_i = \frac{n (x_i - \bar{x}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})}{\sum_{j=1}^n (x_j - \bar{x})^2} \quad (3)$$

Each index in the above formula is the same as Equation (1). In addition, we also showed the spatial cluster type of LE by creating a Lisa cluster map.

### 2.2.3. Geographical Detector

Based on the spatial variance analysis, Geographic detector is a new tool to detect environmental factors of health risk. By comparing between strata-variance with the total variance in the whole area of dependent variable, the Geographical Detector can detect whether the factor causes the spatial stratified heterogeneity of dependent variable or not [24]. The Geographical Detector consists of 4 parts: factor detector, ecological detector, risk detector and interaction detector. Each of its components is described in detail as follows:

#### Factor Detector

A factor detector can be used to detect the importance of certain factors on LE, and is commonly measured by the Power of Determinant (*PD*) as follows [33]:

$$PD = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} \quad (4)$$

where  $h = 1, \dots, L$  refers to a certain stratum of each factor ( $L$  is the number of strata of the factor);  $\sigma^2$  and  $\sigma_h^2$  represents the variance of LE in the whole and stratum  $h$ , respectively,  $N$  and  $N_h$  represent the sample for them, respectively. *PD* ranges from 0 to 1. If *PD* is closer to 1, the effect of this factor on LE is greater. And the *p*-value of *PD* was also given by the factor detector.

#### Risk Detector

The risk detector is used to judge whether there is a significant difference between the average LE of different strata of each factor. T-test is used for hypothesis tests[16]:

$$t_{\bar{y}_{h=1} - \bar{y}_{h=2}} = \frac{\bar{y}_{h=1} - \bar{y}_{h=2}}{\left[ \frac{\text{Var}(\bar{y}_{h=1})}{n_{h=1}} + \frac{\text{Var}(\bar{y}_{h=2})}{n_{h=2}} \right]^{1/2}} \quad (5)$$

The degree of freedom was:

$$df = \frac{\frac{\text{Var}(\bar{y}_{h=1})}{n_{h=1}} + \frac{\text{Var}(\bar{y}_{h=2})}{n_{h=2}}}{\frac{1}{n_{h=1} - 1} \left[ \frac{\text{Var}(\bar{y}_{h=1})}{n_{h=1}} \right]^2 + \frac{1}{n_{h=2} - 1} \left[ \frac{\text{Var}(\bar{y}_{h=2})}{n_{h=2}} \right]^2} \quad (6)$$

where  $\bar{Y}_h$  represents the average LE of the stratum  $h$ ;  $n_h$  indicates samples for stratum  $h$ ;  $Var$  represents sample variance,  $t$  follows the Student's- $t$ -test distribution. The null hypothesis refers to the equivalent of LE in two areas. If null hypothesis is rejected at the significance level  $\alpha$ , it is considered that there exists a significant difference in LE between the two areas.

### Ecological Detector

The ecological detector is able to detect whether there is a significant difference between the effects of two factors on LE. And it is measured by the following formula:

$$F = \frac{n_{x1}(n_{x2} - 1)SSW_{x1}}{n_{x2}(n_{x1} - 1)SSW_{x2}} \quad (7)$$

$$SSW_{x1} = \sum_{h=1}^{L1} N_h \sigma_h^2 \quad SSW_{x2} = \sum_{h=1}^{L2} N_h \sigma_h^2$$

where  $n_{x1}$  and  $n_{x2}$  represent the samples;  $SSW_{x1}$  and  $SSW_{x2}$  represent the sum of within-strata variance of strata divided by  $x_1$  and  $x_2$  (any two social and environmental factors in the study) respectively;  $L_1$  and  $L_2$  represent the number of strata of variables  $x_1$  and  $x_2$  respectively; The null hypothesis represent  $SSW_{x1}$  and  $SSW_{x2}$  are equal. If null hypothesis is rejected at the significance level of  $\alpha$ , this indicates that there is a significant difference between the effects of two factors on LE.

### Interaction Detector

The interaction detector is used to identify the interaction between different factors. It can evaluate whether the interaction of factors  $x_1$  and  $x_2$  will increase or decrease the influence on LE. Specifically, it calculates the  $PD$  value of  $x_1$  and  $x_2$  after interaction ( $X_1 \cap X_2$ ) and then compares it with  $PD(x_1)$  and  $PD(x_2)$  to judge the interaction type. As shown in line 3 of Table 2, the  $PD$  value of  $x_1$  and  $x_2$  after interaction is greater than the maximum of the original  $PD$  values of the two factors, suggesting that the interaction type is bivariate-enhanced.

**Table 2.** Types of interaction between two covariates [34].

Description	Interaction
$PD(X_1 \cap X_2) < \text{Min}(PD(X_1), PD(X_2))$	Weaken, nonlinear
$\text{Min}(PD(X_1), PD(X_2)) < PD(X_1 \cap X_2) < \text{Max}(PD(X_1), PD(X_2))$	Weaken, univariate
$PD(X_1 \cap X_2) > \text{Max}(PD(X_1), PD(X_2))$	Enhanced, bivariate
$PD(X_1 \cap X_2) = PD(X_1) + PD(X_2)$	Independent
$PD(X_1 \cap X_2) > PD(X_1) + PD(X_2)$	Enhance, nonlinear

### 2.3. Data Preprocessing

The Geographical Detector analysis generally requires the independent variable to be categorical variable, so it is necessary to discretize the 11 social and environmental factors before the analysis. Different discretization schemes will have different impact on performance. Generally, the discretization scheme with the largest  $PD$  value is preferred in a geographical detector analysis [35]. After comparison of various discretization schemes, we used the quantile method to classify the above 11 social and environmental variables and the classification interval is shown in Table 3. As for the dependent variable, it can be either continuous or discrete, so we did not make any preprocessing.

**Table 3.** Descriptive analysis and discretization results of social and environmental factors.

Variables	Max	Min	Classification Interval
BR (‰)	15.99	6.68	8
TDR (%)	51.45	20.94	6
NORPH	4.23	2.45	6
WRPC (m <sup>3</sup> )	153681.9	72.8	4
CLOUR(RMB)	34588	10523	3
GPC (RMB)	76074	13119	4
NOCS	6196.36	1109.34	8
MCEOUR (RMB)	1389.5	352.3	3
NOHB	7.44	2.51	6
NOD	5.24	1.04	7
UR (%)	89.3	22.7	6

#### 2.4. Software

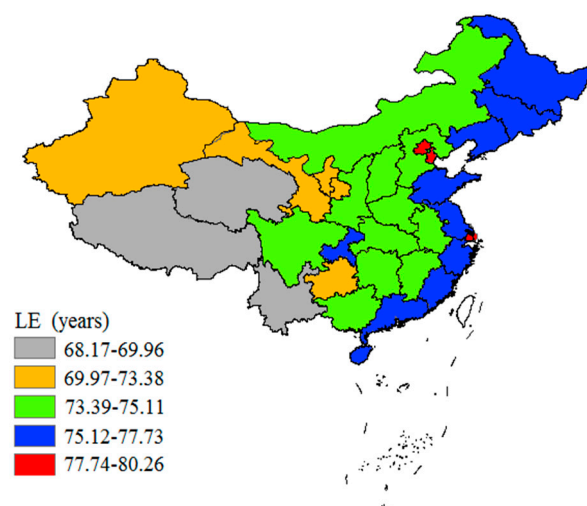
The discretization of social and environmental factors and the spatial autocorrelation analysis were implemented by ArcGIS 10.2, and the analysis of spatial stratified heterogeneity of LE was completed by the GeoDetector.

### 3. Results

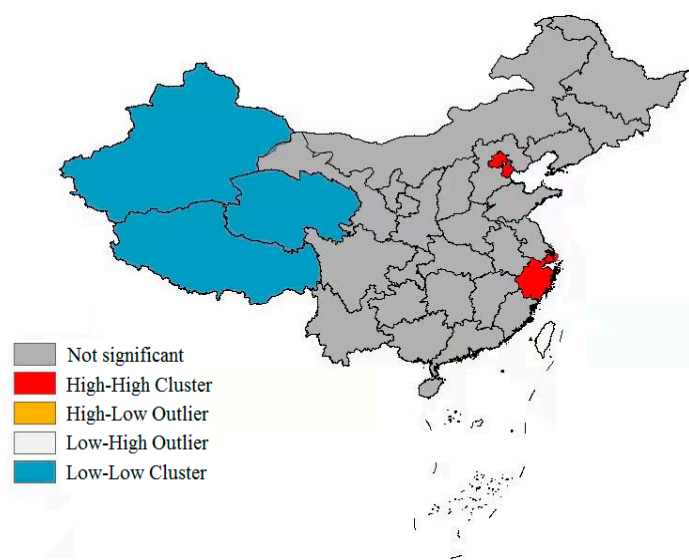
#### 3.1. Spatial Distribution Characteristics of LE in China

In 2010, the average LE was 74.83 years old in China, and the average LE of women was higher than that of men (77.37 vs 72.38). The results show that LE showed a clear downward trend from the east to the west in 2010 in China, which reveals that there existed obvious spatial stratified heterogeneity of LE in China (Figure 1). The average LE of the eastern areas was over 75 years old, especially in Shanghai and Beijing. The average LE of the central areas was a bit lower than the eastern areas which was between 73.39 and 75.11 years, while the average LE in the western areas was the lowest in China. Tibet had the lowest LE of 68.17 years in all the 31 provinces.

From the results of the global autocorrelation analysis, the global Moran's I was 0.266 ( $p < 0.001$ ), indicating that there existed obvious spatial clustering for LE in China. In order to further explore the cluster type of LE, we conducted a local autocorrelation analysis and obtained a Lisa cluster map (Figure 2). It can be seen from Figure 2 that LE mainly presented two cluster types: high–high and low–low, among which high–high types were mainly distributed in the eastern coastal areas, including Beijing, Tianjin, Shanghai and Zhejiang provinces. The low–low type was mainly located in Western inland regions, such as Xinjiang, Tibet and Qinghai provinces.



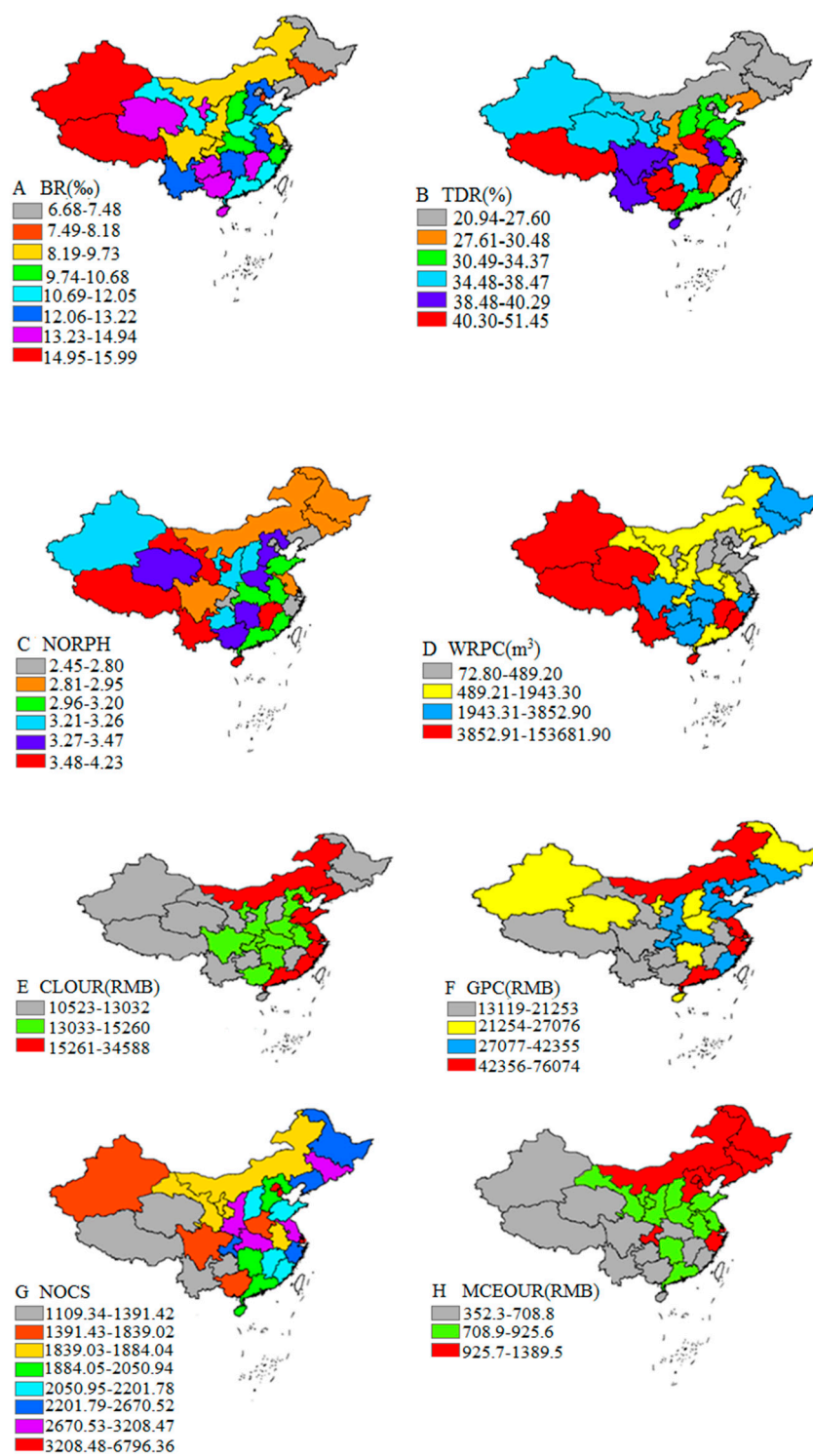
**Figure 1.** Distribution of LE (Life expectancy) in China in 2010.



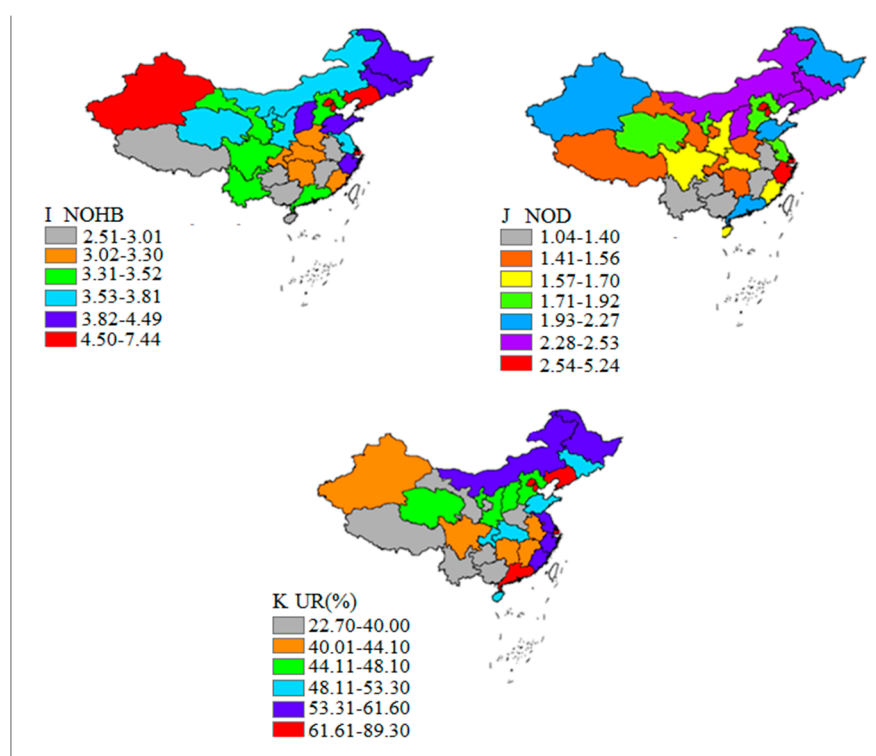
**Figure 2.** Lisa cluster map of LE in China in 2010.

### 3.2. Spatial Distribution of Main Social and Environmental Factors

We also performed spatial maps of social and environmental factors to reveal their spatial characteristics (Figure 3). We can see that there exist certain spatial differences in BR, NORPH and TDR. And the three indicators show an upward trend from the east to the west, which were different from the spatial distribution of LE (Figure 3A–C). There were great spatial differences in the WRPC. For instance, the WRPC in Tibet reached the highest level (153681.9 m<sup>3</sup>), while in Tianjin, it was only 72.8 m<sup>3</sup>. In addition, the WRPC increased from the north to the south in China (Figure 3D). NOHB and NOD showed a decline trend from the north to the south (Figure 3I–J). There existed big differences in the CLOUR, GPC, NOCS, MCEOUR and UR among all provinces in China, showing a downward trend from the east to the west (Figure 3 E–H,K)







**Figure 3.** Spatial distribution map of social and environmental factors of LE in China in 2010. (A) Distribution of BR (Birth rate); (B) Distribution of TDR (Total dependency rate); (C) Distribution of NORPH (Number of residents per household); (D) Distribution of WRPC (Water resource per capita); (E) Distribution of CLOUR (Consumption level of urban residents); (F) Distribution of GPC (GDP per capita); (G) Distribution of NOCS (Number of college students per 100,000 persons); (H) Distribution of MCEOUR (Medical care expenditure of urban residents); (I) Distribution of NOHB (Number of hospital beds per 1000 persons); (J) Distribution of NOD (Number of doctors per 1000 persons); (K) Distribution of UR (Urbanization rate).

### 3.3. The Influence of Social and Environmental Factors on LE Based on Factor Detector

Based on the factor detector, we analyzed the importance of social and environmental factors on LE. Table 4 lists the *PD* value and its *p* value of each factor. The *PD* value of medical resources (NOHB, NOD) was not statistically significant ( $p > 0.05$ ), which indicates that medical resources had no effect on LE at the provincial scale. However, the *PD* of most variables ranged from 0.45 to 0.65. In particular, the *PD* value of NOCS was 0.89, which indicates that NOCS can mainly explain the spatial stratified heterogeneity of LE in our study. Moreover, the *PD* values of WRPC and MCEOUR were less than 0.4, revealing that they had less influence on LE than other factors.

**Table 4.** The result of factor detector about social and environmental factors of LE.

Variables	<i>PD</i>	<i>p</i>
BR	0.61	0.017
TDR	0.49	0.024
NORPH	0.57	0.004
WRPC	0.35	0.015
CLOUR	0.47	0.014
GPC	0.50	0.046
NOCS	0.89	0.000
MCEOUR	0.396	0.040
NOHB	0.33	0.621
NOD	0.52	0.647
UR	0.64	0.049

LE: Life expectancy; PD: Power of Determinant.

### 3.4. The Optimal Range of Factors for the Maximum LE Based on Risk Detector

The risk detector showed the average LE of each stratum of all factors and analyzed whether there was significant difference between each stratum. Taking GPC as an example: the relationship between GPC and LE is shown in Figure 4. With the increase of GPC, LE also increased. When the GPC ranged from 13119 to 21253 RMB, the average LE reached 72.54 years old. However, when the GPC ranged from 42355 to 76074 RMB, the average LE rose to 77.80 years old. The significance of the average LE differences between each stratum of GPC is shown in Table 5 (Y: significant, N: not significant). The fourth layer corresponds to the maximum range in Figure 4 (42355 to 76074 RMB), while the first layer corresponds to the minimum range in Figure 4 (13119 to 21253 RMB). We can see that the difference of average LE between the fourth layer and other layers of GPC was statistically significant in Table 5. Furthermore, the risk detector was able to analyze the quantitative relationship between social and environmental factors and LE. As could be seen in Table 6, the results show that the highest ranges of CLOUR, GPC, NOCS, MCEOUR and UR related to the highest level of LE, while the lowest ranges of BR, TDR, NORPH and WRPC corresponded to the highest level of LE. The optimal range of factors corresponds to the maximum value of the LE [36]. Therefore, we displayed the optimal range of factors in Table 6. It is worth noting that the areas with the highest LE value could be identified as the main influencing area for each significant social and environmental factor [34]. We show the results in Figure 5 with visualization technology.

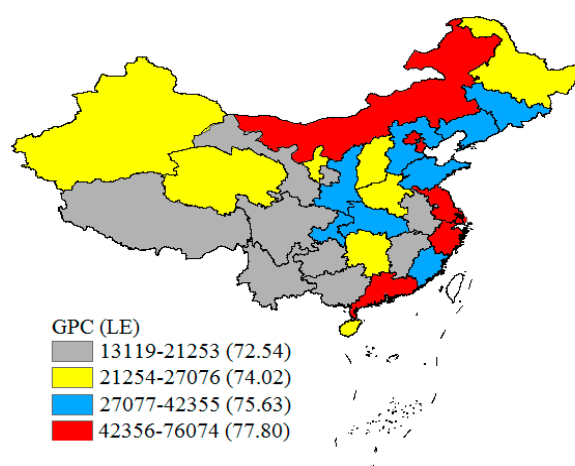


Figure 4. The subregions of the GPC and their LE.

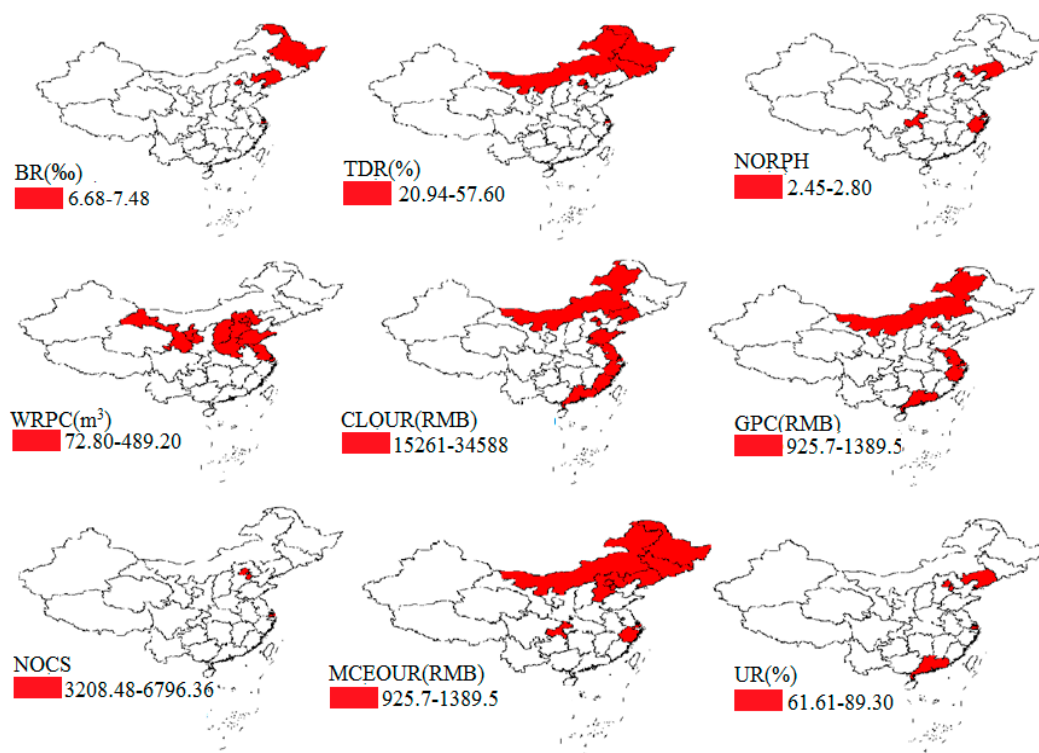
Table 5. Significance of average LE difference between different layers of GPC (GDP per capita).

Stratum	1	2	3	4
1				
2	N			
3	Y	N		
4	Y	Y	Y	

Table 6. The optimal range of social and environmental factors and maximum average LE.

Variables	Optimal range	Maximum average LE (years)
BR(‰)	6.68–7.48	78.20
TDR (%)	20.94–27.60	77.66
NORPH	2.45–2.80	78.19
WRPC (m <sup>3</sup> )	72.80–489.20	76.96

CLOUR(RMB)	15261–34588	77.32
GPC (RMB)	42356–76074	77.80
NOCS	3208–6796	79.78
MCEOUR (RMB)	925.70–1389.50	77.07
UR (%)	61.61–89.30	78.44



**Figure 5.** Distribution of main influencing area of each influencing factors.

### 3.5. Differences between the Impact of Different Factors on LE based on the Ecological Detector

The significance of differences in *PD* values between two factors on LE was compared by the ecological detector (Table 7). The results show that the differences in *PD* values among most factors were not significant (Y: significant, N: not significant). The differences in *PD* values of NOCS and other factors were statistically significant, suggesting that NOCS had a great impact on LE. The *PD* value of WRPC was significantly different from that of BR and UR, respectively. Similarly, the difference in *PD* values of MCEOUR and UR was statistically significant. Combined with the results of the factor detector, this shows that the impact of WRPC and MCEOUR on LE was weak.

**Table 7.** Statistical significance of the differences in *PD* values among different factors.

Impact Factors	BR	TDR	NORPH	WRPC	CLOUR	GPC	NOCS	MCEOUR	UR
BR									
TDR	N								
NORPH	N	N							
WRPC	Y	N	N						
CLOUR	N	N	N	N					
GPC	N	N	N	N	N				
NOCS	Y	Y	Y	Y	Y	Y			
MCEOUR	N	N	N	N	N	N	Y		
UR	N	N	N	Y	N	N	Y	Y	

### 3.6. Interaction Between Different Factors based on the Interaction Detector

We used the interaction detector to reveal the interaction effect and types among the factors. As shown in Table 8, the *PD* value was  $\geq 0.9$  after NOCS interacted with other factors. Notably, the *PD* value of the interaction between NOCS and NORPH reached 0.98, which was closer to 1. The *PD* value of UR interacted with other factors was also high; for example, the *PD* value of UR that interacted with WRPC was 0.92. However, the *PD* value of the interaction between WRPC and MCEOUR was only about 0.5. We found that the interaction effect of any two factors was greater than the individual effect of a certain factor on LE. Even for the factors with a lower *PD* value, their *PD* value increased after the interaction. Moreover, the results show that all interaction types were bivariate-enhanced.

Table 8. *PD* values for interactions between factors on the LE.

Impact Factors	BR	TDR	NORPH	WRPC	CLOUR	GPC	NOCS	MCEOUR	UR
BR	0.61								
TDR	0.86	0.49							
NORPH	0.85	0.75	0.57						
WRPC	0.85	0.73	0.79	0.35					
CLOUR	0.78	0.69	0.71	0.64	0.47				
GPC	0.84	0.63	0.79	0.69	0.65	0.50			
NOCS	0.97	0.96	0.98	0.95	0.96	0.94	0.89		
MCEOUR	0.68	0.60	0.65	0.52	0.66	0.59	0.90	0.40	
UR	0.89	0.81	0.85	0.92	0.80	0.74	0.95	0.78	0.64

## 4. Discussion

LE is an important indicator for measuring health status [37]. To our knowledge, this is the first time that the relationship between LE and social and environmental factors is explored in China from spatial perspective using the Geographic Detector technique.

Many previous studies have shown that LE was mainly affected by economic factors [36,38,39]. For example, one study found that the main reason for the spatial distribution pattern of LE in China was the economy [40]. In contrast, we found that NOCS could mainly explain the spatial stratified heterogeneity of LE at the provincial scale combined with the results of factor detector and ecological detector, that is, the effect of NOCS on LE was significantly greater than that of the other factors. There is some possible explanation. Firstly, compared with other factors, people with higher education were related to good health awareness and more timely access to health care [15,41]; secondly, a higher education population could resist the adverse effects of negative aspects with better psychological quality [42]. At the same time, our study found that WRPC had little effect on LE. Some previous studies also found that this effect showed an upward trend from 2000 to 2010 [15]. Therefore, our government, especially in the eastern developed areas, should pay attention to the protection of the ecological environment while improving social economy. In addition, the impact of medical resources (NOHB, NOD) on LE was not statistically significant. Even though there were abundant medical resources in some areas, their actual efficiency might be very low due to poor infrastructure and low economic level. Therefore, they might not fully play a role, even had no effect on LE. However, the impact of MCEOUR on LE was also quite small, which further reveals the importance of improving the utilization of medical resources [43].

The results of the risk detector show that when the economic factors (GPC and CLOUR) and UR reached the maximum range, the average LE was also closer to the highest level. Because economic status played a role through its effects on people's daily life, such as education, medical care, etc., the average LE would reach the maximum level with the GPC at the maximum value. The consumption level (CLOUR) was closely related to economic situation. When the consumption level reached the maximum value, people would purchase enough food to improve their health [44,45]. People who lived in the areas with highest UR would have the longest average LE because the high UR

corresponded to the high economic conditions, medical and educational opportunities [10]; meanwhile, residents in rural areas also reported much higher rates of disability, injury and high blood pressure compared with urban residents, due to inequalities in education, health care and poverty [46]. Therefore, the Chinese government, especially in the central and western areas, should focus on the alleviation of poverty and urbanization to improve local LE. At the same time, when the BR and family living standard (TDR, NORPH) were in the minimum range, the average LE reached its maximum level. In general, the areas with the lowest BR were usually economically developed, such as Shanghai and Beijing. The social welfare in these areas was also higher than that of other areas [39]. Moreover, areas with a lowest BR tended to have the fewest number of families members (NORPH) and the lowest total dependency ratio (TDR). Therefore, they had the longest average LE.

Based on the interaction detector, we found that the *PD* value of any social environment factor interacted with NOCS was  $\geq 0.9$ , indicating that education combined with other factors could significantly improve LE level. Therefore, the government, especially in Western China, should focus on improving the education as well as economic and medical conditions.

Our research shows the impact of social and environmental factors on life expectancy and their interaction [47]. We display the optimal range of factors for maximum LE and the main influencing area, which was meaningful for health policy development. Moreover, the selected variables covered multiple dimensions and the data on them were authoritative, which came from the national bureau of statistics. However, there are still some limitations in our study. Firstly, this study only focused on social and environmental factors, so there might be some factors influencing LE that were not included, such as air pollutants PM<sub>2.5</sub>, PM<sub>10</sub>, etc. However, we could not obtain the data of air pollution factors in each province in 2010. In addition, the data of LE and social and environmental factors used in this study were all from 2010, so they were insufficient in inferring a causal relationship. Moreover, the relationship between the dependent variable and independent variable was statistical and was not causality but the geographical detectors could filter out highly potential factors of LE for further confirmation, such as longitudinal studies[16]. In addition, this study was performed at the provincial level, which needs to be studied at a more precise scale in the future. Finally, the Geographical Detector could only explore the interaction effect between two factors and failed to further reveal the impact of multiple interactions on LE, which was also a key problem to be solved in the future.

## 5. Conclusions

In conclusion, there exist obvious spatial stratified heterogeneity of LE in China. Among the many social and environmental factors, NOCS could mainly explain the spatial stratified heterogeneity of LE. BR, TDR, NORPH, CLOUR, GPC and UR had less influence on LE, while WRPC and MCEOUR had the lowest influence on LE. Further study is needed to discover the actual causality between LE and these factors. When BR, TDR, NORPH and WRPC were at the minimum range, LE reached the highest level; conversely, LE reached the highest level with CLOUR, GPC, NOCS, MCEOUR and UR at the maximum range. In addition, the interaction of any two social and environmental factors on LE was stronger than the effect of a single factor. Our results provide political basis for the government to formulate economic and educational development, utilization of medical resources, environmental protection and population management policies to solve the regional inequality of LE in China.

**Author Contributions:** Y.W., K.H., Y.H., Q.S. and Y.F. worked together. Specifically, Y.W. conceived the main idea of this study and drafted the manuscript; K.H. analyzed the data and drafted the manuscript; Y.H. gave constructive suggestions for the modeling; Q.S. was mainly responsible for mapping and organizing the manuscript; Y.F. revised the manuscript. All of the authors read and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the National Natural Science Foundation of China (No. 81973144).

**Acknowledgments:** Thank National Bureau of statistics of the People's Republic of China and National Health and Family Planning Commission of PRC for providing us with data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Gou, X. Quantitative analysis of life expectancy differences and influencing factors in countries around the world. *J. Nanjing Coll. Popul. Programme Manage.* **2013**, *29*, 31–36. (In Chinese)
- The World Bank Database. Available online: <https://data.worldbank.org.cn> (accessed on 11 August 2019).
- Guo, Y., Analysis of spatio-temporal changes in life expectancy and its influencing factors in China. *Chin. J. Health Policy* **2018**, *11*, 44–49. (In Chinese)
- Yu, T. Analysis on the Influencing Factor and Prediction of Life Expectancy. Master Thesis, Capital University of Economics and Business, Beijing, China, June 2017. (In Chinese)
- Manuel, D.G.; Perez, R.; Sanmartin, C.; Taljaard, M.; Hennessy, D.; Wilson, K.; Tanuseputro, P.; Manson, H.; Bennett, C.; Tuna, M.; et al. Measuring Burden of Unhealthy Behaviours Using a Multivariable Predictive Approach: Life Expectancy Lost in Canada Attributable to Smoking, Alcohol, Physical Inactivity, and Diet. *PLoS Med.* **2016**, *13*, e1002082.
- Song, X.; Chen, G.; Zheng, X. Chinese Life Expectancy and Policy Implications. *Procedia—Soc. Behav. Sci.* **2010**, *2*, 7550–7555.
- Gilligan, A.M.; Skrepnek, G.H. Determinants of life expectancy in the Eastern Mediterranean Region. *Health Policy Plan* **2015**, *30*, 624–637.
- Wilkinson, R.G. *Unhealthy Societies: The Afflictions of Inequality*; Routledge: London, UK, 2002.
- Rodgers, G.B., Income and inequality as determinants of mortality: An international cross-section analysis. *Popul. Stud.* **1979**, *33*, 343–351.
- Lu, X.; Chen, X., Factors on mean lifespan of each province in China. *Yunan Geogr. Environ. Res.* **2007**, *3*, 72–78. (In Chinese)
- Yang, Z.; Liu, H.; Wang, X., Spatio-temporal variations of population health distribution in China and its influencing factors. *World Reg. Stud.* **2017**, *26*, 161–168. (In Chinese)
- Meara, E. R.; Richards, S.; Cutler, D. M. The gap gets bigger: Changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs* **2008**, *27*, 350–360.
- Hao, Y., A stepwise regression analysis of comprehensive environmental factors affecting life expectancy in Shanghai. *Economist* **2003**, *11*, 33–35. (In Chinese)
- Dong, L.; Li, C., Relationship between life expectancy, GNP and education level. *Stat. Decis.* **2009**, *23*, 95–98. (In Chinese)
- Jiang, J.; Luo, L.; Xu, P.; Wang, P., How does social development influence life expectancy? A geographically weighted regression analysis in China. *Public Health* **2018**, *163*, 95–104.
- Wang, J.F.; Li, X.H.; Christakos, G.; Liao, Y.L.; Zhang, T.; Gu, X.; Zheng, X.Y., Geographical Detectors-Based Health Risk Assessment and its Application in the Neural Tube Defects Study of the Heshun Region, China. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 107–127.
- Zhou, C.; Chen, J.; Wang, S., Examining the effects of socioeconomic development on fine particulate matter (PM<sub>2.5</sub>) in China's cities using spatial regression and the geographical detector technique. *Sci. Total Environ* **2018**, *619–620*, 436–445.
- Wang, J.F.; Zhang, T.L.; Fu, B.J., A measure of spatial stratified heterogeneity. *Ecol. Indic.* **2016**, *67*, 250–256.
- Ju, H.; Zhang, Z.; Zuo, L.; Wang, J.; Zhang, S.; Wang, X.; Zhao, X., Driving forces and their interactions of built-up land expansion based on the geographical detector – a case study of Beijing, China. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 2188–2207.
- Ren, Y.; Deng, L.; Zuo, S.; Luo, Y.; Shao, G.; Wei, X.; Hua, L.; Yang, Y., Geographical modeling of spatial interaction between human activity and forest connectivity in an urban landscape of southeast China. *Landscape Ecol.* **2014**, *29*, 1741–1758.
- Tan, J.; Zhang, P.; Lo, K.; Li, J.; Liu, S., The Urban Transition Performance of Resource-Based Cities in Northeast China. *Sustainability* **2016**, *8*, 1022.
- Lou, C.R.; Liu, H.Y.; Li, Y.F.; Li, Y.L., Socioeconomic Drivers of PM<sub>2.5</sub> in the Accumulation Phase of Air Pollution Episodes in the Yangtze River Delta of China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 928.
- Todorova, Y.; Lincheva, S.; Yotinov, I.; Topalova, Y., Contamination and Ecological Risk Assessment of Long-Term Polluted Sediments with Heavy Metals in Small Hydropower Cascade. *Water Resour. Manage.* **2016**, *30*, 4171–4184.
- Hu, Y.; Wang, J.; Li, X.; Ren, D.; Zhu, J., Geographical detector-based risk assessment of the under-five

- mortality in the 2008 Wenchuan earthquake, China. *PLoS ONE* **2011**, *6*, e21427.
25. National Bureau of statistics of the People's Republic of China. *China Statistical Yearbook 2012*; China statistical Press: Beijing, China, 2012. (In Chinese)
  26. National Bureau of Statistics of the People's Republic of China. *China Population & Employment Statistics Yearbook 2011*; China statistical Press: Beijing, China, 2011. (In Chinese)
  27. National Health and Family Planning Commission of PRC. *China Health Statistical Yearbook 2011*; Peking Union Medical College Press: Beijing, China, 2011. (In Chinese)
  28. National Bureau of statistics of the People's Republic of China. *China Statistical Yearbook 2011*; China statistical Press: Beijing, China, 2011. (In Chinese)
  29. National Bureau of statistics of the People's Republic of China. *China Statistical Yearbook 2014*; China statistical Press: Beijing, China, 2014. (In Chinese)
  30. Sridharan, S.; Tunstall, H.; Lawder, R.; Mitchell, R., An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in Scotland. *Soc. Sci. Med.* **2007**, *65*, 1942–52.
  31. Wang, S.; Luo, K., Life expectancy impacts due to heating energy utilization in China: Distribution, relations, and policy implications. *Sci. Total Environ.* **2018**, *610–611*, 1047–1056.
  32. Anselin, L., Local Indicators of Spatial Association-LISA. *Geographical Analysis* **1995**, *27*, 93–115.
  33. Wang, J.F.; Wang, Y.; Zhang, J.; Christakos, G.; Sun, J.L.; Liu, X.; Lu, L.; Fu, X.Q.; Shi, Y.Q.; Li, X.M., Spatiotemporal transmission and determinants of typhoid and paratyphoid fever in Hongta District, Yunnan Province, China. *PLoS Negl. Trop. Dis.* **2013**, *7*, e2112.
  34. Bai, L.; Jiang, L.; Yang, D.-y.; Liu, Y.-b., Quantifying the spatial heterogeneity influences of natural and socioeconomic factors and their interactions on air pollution using the geographical detector method: A case study of the Yangtze River Economic Belt, China. *Journal of Cleaner Production* **2019**, *232*, 692–704.
  35. Cao, F.; Ge, Y.; Wang, J.-F., Optimal discretization for geographical detectors-based risk assessment. *GIScience & Remote Sensing* **2013**, *50*, 78–92.
  36. Zha, X.; Tian, Y.; Gao, X.; Wang, W.; Yu, C., Quantitatively evaluate the environmental impact factors of the life expectancy in Tibet, China. *Environ Geochem Health* **2019**, *41*, 1507–1520.
  37. Wang, L.; BingganWei; Li, Y.; Li, H.; Zhang, F.; Rosenberg, M.; Yang, L.; Huang, J.; Krafft, T.; Wang, W., A study of air pollutants influencing life expectancy and longevity from spatial perspective in China. *Sci. Total Environ.* **2014**, *487*, 57–64.
  38. Lin, R.-T.; Chen, Y.-M.; Chien, L.C.; Chan, C.C., Political and social determinants of life expectancy in less developed countries: A longitudinal study. *BMC Public Health* **2012**, *12*, 85.
  39. Duque, A.M.; Peixoto, M.V.; Lima, S.V.M. A.; O.Goes, M.A.; Santos, A.D.; Araújo, K.C.G. M.; Nunes, M.A.P., Analysis of the relationship between life expectancy and social determinants in a north-eastern region of Brazil, 2010–2017. *Geospat. Health* **2018**, *13*, (2).
  40. Wang, S.; Luo, K.; Liu, Y., Spatio-temporal distribution of human lifespan in China. *Sci Rep* **2015**, *5*, 13844.
  41. Chan, M.F., The impact of health care resources, socioeconomic status, and demographics on life expectancy: A cross-country study in three Southeast Asian countries. *Asia Pac. J. Public Health* **2015**, *27*, NP972–83.
  42. Shkolnikov, V.M.; Andreev, E.M.; Jasilionis, D., The changing relation between education and life expectancy in central and eastern Europe in the 1990s. *J. Epidemiol. Community Health* **2006**, *60*, 875–881.
  43. Ming, Y.; Dong, Z., Life expectancy of China' s population analysis of the impact of factors. *Theory Research* **2010**, *4*, 47–50. (In Chinese)
  44. Anand, S.; Ravallion, M., Human Development in Poor Countries: On the Role of Private Incomes and Public Services. *Journal of Economic Perspectiv* **1993**, *7*, 133–150.
  45. Zheng, C., Analysis of regional differences in life expectancy and economic and social factors. *China Collective Economy* **2010**, *07*, 81–83. (In Chinese)
  46. Singh, G.K.; Siahpush, M., Widening rural–urban disparities in life expectancy, US, 1969–2009. *American journal of preventive medicine* **2014**, *46*, e19–e29.
  47. Huang, J.; Wang, J.; Bo, Y.; Xu, C.; Hu, M.; Huang, D., Identification of health risks of hand, foot and mouth disease in China using the geographical detector technique. *Int J. Environ. Res. Public Health* **2014**, *11*, 3407–3423.





© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).