

Article

Investigating the Impact of Human Activity on Land Use/Cover Change in China's Lijiang River Basin from the Perspective of Flow and Type of Population

Jun Li ^{1,2}, Yuan Zhang ³, Qiming Qin ^{2,*} and Yueguan Yan ^{1,4}

¹ College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China; junli_geo@126.com (J.L.); yyg_0720@163.com (Y.Y.)

² Institute of Remote Sensing and Geographic Information Systems, Peking University, Beijing 100871, China

³ Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; zhangyuan76@gmail.com

⁴ State Key Laboratory for GeoMechanics and Deep Underground Engineering, China University of Mining and Technology, Beijing 100083, China

* Correspondence: qmqin@pku.edu.cn; Tel.: +86-10-6275-1965

Academic Editor: Audrey L. Mayer

Received: 14 November 2016; Accepted: 2 March 2017; Published: 5 March 2017

Abstract: Exploring how human activity impacts land use/cover change (LUCC) is a hot research topic in the field of geography and sustainability management. Researchers have primarily used socioeconomic variables to measure human activity. However, the human activity indexes mainly based on socioeconomic variables have a spatial resolution that is coarser than traditional LUCC datasets, which hinders a deep and comprehensive analysis. In view of these problems, we selected China's Lijiang River Basin as our study area and proposed the use of GPS trajectory data for analyzing the impact of human activity on LUCC from two perspectives: (1) Type of population: we used the kernel density estimation method to extract the spatial distribution of activity intensity of local residents and tourists, investigated their correlation with the LUCC result, and found these two populations have different impacts on each land cover; (2) Flow of population: we used the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm and a network analysis method to build a flow network of population from raw trajectories, conducted regression analysis with LUCC, and found that the flow of population is an important factor driving LUCC and is sometimes a more important factor than the static distribution of the population. Experimental results validated that the proposed method can be used to uncover the impact mechanism of human activity on LUCC at fine-grained scales and provide more accurate planning and instructions for sustainability management.

Keywords: human activity; land use/cover change; impact analysis; trajectory data analysis

1. Introduction

Since the Industrial Revolution in 1760s, human activity has brought significant changes to Earth's natural environment with the improvement of engineering tools. An apparent phenomenon among these is the change of land cover [1]. According to some estimates [2,3], the majority of the terrestrial biosphere was transformed to agricultural and settled anthromes by 2000. For example, during the period from 1880 to 1980, logging resulted in a 47% decline of forest/woodland in tropical Asia [4]. Human-induced land use/cover changes (LUCC), in turn, also affect human survival and development, which has raised widespread concerns in human society. By human activity, we mean all undertakings by humans for survival and improvement of living standards including land reclamation, grazing, water resource use and development, engineering construction, ecological system management, etc. [5].

In recent decades, human activity and socioeconomic variables have become very important factors driving LUCC [6]. Therefore, understanding the impact of human activity on LUCC is necessary for providing scientific instructions for human activity regulation and effective land resource allocation.

Nowadays, remote sensing technology has been widely used as an effective tool to determine LUCC over a large area, so collecting quantitative information of human activity at fine-grained scales becomes essential for studying human-earth interaction. Scholars used to quantify human activity by socioeconomic indexes such as population density, GDP, and then ascertain human activity and its resulting impact on the landscape. For example, Lü et al. (2004) selected road density, ratio of human settlement to nonsettlement, industrial land, and farmland to assess the intensity of human activity. They found that the perimeter density of landscape can be well expressed as the linear combination of the spatial elements of human activity, and this implied that the trend of landscape change could be pre-estimated using the change of spatial elements of human activities [7]. Calo and Parise (2006) developed a Karst Disturbance Index (KDI) based on a categorical framework encompassing physical, biological, and social indicators to evaluate the degree of disturbance in two areas in the Apulian karst of southern Italy. Their experimental results showed that KDI is a useful approach for a preliminary evaluation of the degree of disturbance in karst [8]. Guo et al. (2006) calculated the degree of human disturbance based on the proportion of construction, tourism, town, and country effects, and discovered that the changes in landscape structure varied significantly based on the degree of human activity. Their findings illustrated that anthropogenic disturbance is the primary cause leading to the division of landscape patches [9]. In addition, Wang et al. (2010) calculated the length of roads, number of road crossings, number of residents, number of permitted discharge facilities, number of toxic release sites, etc., and used them to represent the levels of human disturbances [10]. Garbarino et al. (2014) used Euclidean-distance-based factors including distances from buildings, roads, and tourism lodges as assessment indexes of human activity, and their results suggested relationships among forest structure, anthropogenic influences, and topography [11]. Similarly, Roth et al. (2016) analyzed the spatial distribution of human activity based on the locations of road networks and population centers, conducted a land use/cover transition analysis, and found a large proportion of transitions from anthropogenic land use/covers to natural ones [12].

Although previous work has made progress in measuring human activity, the following problems still exist: First, human activity indexes based on socioeconomic variables are usually at the scale of the administrative unit, which makes it difficult to conduct in-depth analysis between human activity and land cover. Second, they only reflect a coarse intensity distribution of human activity and cannot show the flow of populations among regions. Third, they are unable to demonstrate the difference in spatiotemporal distribution among different types of populations. Therefore, finding a direct approach to monitor human activity at fine spatial scales is necessary.

As location and wireless communication technologies gradually become ubiquitous, more and more sensors are used in various fields, and a large amount of data are produced. This type of data collection process is called participatory sensing by Goldman et al. (2009) [13]. Participatory sensing data are collected by GPS receivers, mobile phones, etc. [14–18], and are increasingly being applied in monitoring human activity [19]. We can obtain a high-resolution distribution of activity intensity, flow of population, and also acquire information for different types of populations if the attribute information is available. The integration of participatory sensing and remote sensing can provide technical basis for studying how human activity impacts LUCC in a deeper and more comprehensive way.

This paper is an extension of a previous paper [20] which applied GPS trajectory data to study the relationship of human activity and landscape. Compared to that paper, which explores the interaction from the perspective of activity intensity, this paper investigates the interactions from a different perspective: flow and type of population. We hypothesize that (1) Flow of population has different impacts on LUCC compared with the static distribution of population, and the direction of flow also

plays an important role in driving LUCC; (2) Different types of populations have different impacts on LUCC.

2. Materials and Methods

2.1. Study Area

The study area is Guilin City in China's Guangxi Zhuang Autonomous Region, consisting of 13 counties as shown in Figure 1: Guilin Urban Area (GUA), Lingchuan County (LC), Xing'an County (XA), Yangshuo County (YS), Longsheng County (LS), Ziyuan County (ZY), Quanzhou County (QZ), Guanyang County (GY), Gongcheng County (GC), Pingle County (PL), Lipu County (LP), Yongfu County (YF), and Lingui County (LG). The Lijiang River Basin, which is present in four counties of the study area, is a hotspot region for studying LUCC. On account of the scenic rivers and mountains of the Lijiang River karst landscape, this region has become a well-known tourist destination [21]. Over 40 million tourists visited this region in 2015. While the rapid development of the tourism industry has helped grow the local economy, it has also caused severe environmental degradation during the past two decades, including significant changes in land cover and landscape patterns [22,23].

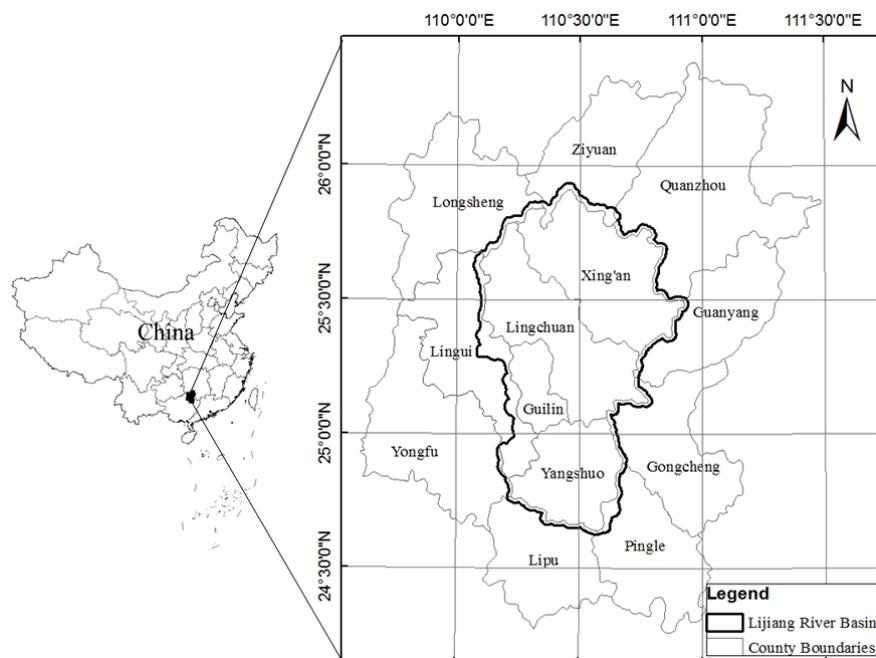


Figure 1. Location of the study area.

2.2. Data Acquisition

The datasets in the previous study [20] are also used in this study including: Landsat images, the Global Digital Elevation Model of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER GDEM), land use and cover products, high-spatial-resolution remote sensing images, and GPS trajectory data. (1) Landsat images of the study area were downloaded from the United States Geological Survey (USGS) data archive with a spatial resolution of 30 m. Four scenes of Landsat images (Path: 124 to 125, Row: 42 to 43) cover the study area. The images for 2009 were collected on 31 October and 23 November, while those for 2013 were collected on 26 October and 4 December; (2) ASTER GDEM product is generated based on the observation data of the Terra Satellite. Its spatial resolution is 30 m; (3) The land use/cover products before 2009 were collected from the local tourism bureau and statistics bureau. Although these products were not collected at the same time as the Landsat images, they can be used as supplementary reference data. Several on-the-spot

investigations were also made using GPS receivers and digital cameras. They are useful for model training and validation for classification; (4) High-spatial-resolution remote sensing images from Google Earth are also used as supplementary reference data for model training and validation; (5) GPS trajectory data were collected by the National Commercial Vehicle Monitoring Platform operated by the Ministry of Transportation of China in the year of 2012. There are 2655 coaches and 1571 tourist shuttles monitored in the dataset. These two types of vehicles are mainly minibuses with a capacity of around 20. Additionally, this paper also uses socioeconomic statistics and basic geographic data; (6) Socioeconomic statistics in 2008 and 2012 were collected from the Guilin Statistical Yearbook including GDP, population, number of tourists, etc.; (7) Basic geographic data includes vector road and administrative divisions at a scale of 1:10,000.

2.3. Methods

The general workflow is composed of three major steps as shown by Figure 2:

- (1) Land use/cover classification based on remote sensing images. Google Earth images and available land use/cover products are split into two parts: one part for classification model training and another for validation. Landsat images and the ASTER GDEM product are used as inputs, and an object-oriented classification method is applied to obtain the land use/cover results in 2009 and 2013.
- (2) Human activity analysis based on GPS trajectory. Map matching is conducted between raw GPS trajectories and road networks to filter out noise GPS points. The kernel density estimation method is used to compute activity intensity maps of local people and tourists from coaches and tourist shuttles, respectively; on the other hand, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is used to extract Stay Regions of vehicles, and then the flow network of humans among the counties is generated according to network theory.
- (3) Human impact analysis on LUCC. By overlaying the human activity of two populations on LUCC, the correlation between them and the difference between the two types of populations are analyzed, while regression analysis is used to investigate the impact of the flow of the population on LUCC, including the differences among land types.

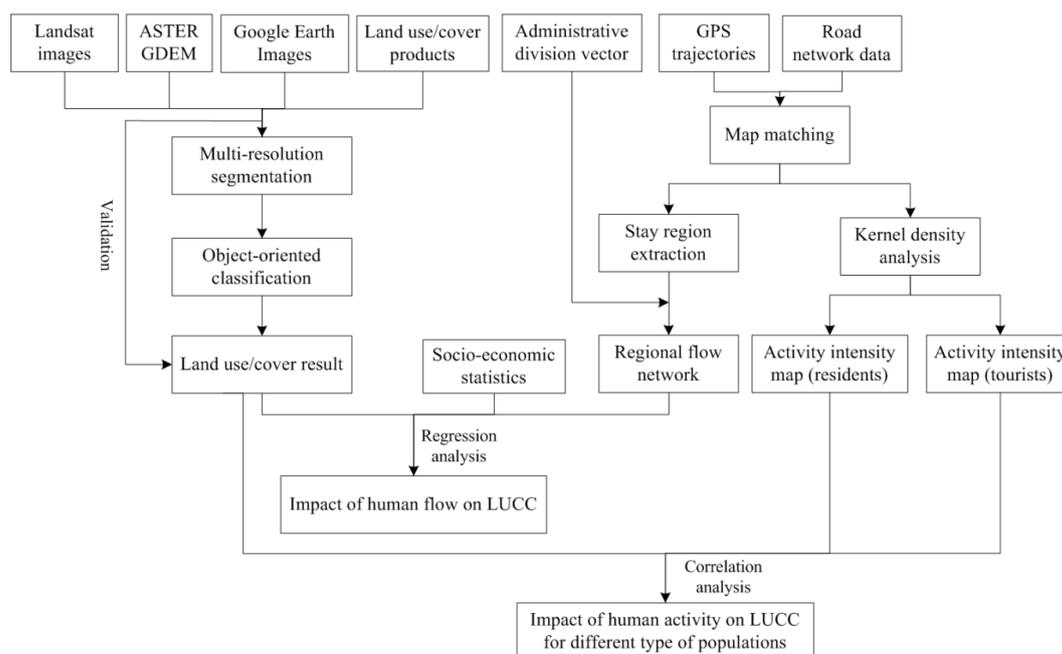


Figure 2. General workflow.

2.3.1. Land Use/Cover Classification

Land use/cover is classified based on remote sensing images through the approach presented in Li et al. (2016). Because the classification result is also used to develop landscape metrics that are sensitive to fragmentation, the supervised object-oriented classification method [24] is chosen for this research. Before that, the software eCognition is applied to implement multi-resolution segmentation [25] for segmenting images into homogeneous regions. Five types of land; woodland, water, built-up land, farmland, and others (mainly shrub and bare land), are subjects of interest in this study. For more information on the land use/cover classification, see Li et al. (2016) [20].

2.3.2. Human Activity Analysis

2.3.2.1. Flow Analysis of the Population

In order to uncover the flow information, the start and end points of each population flow need to be first identified, namely the Stay Regions of vehicles [26]. Stay Regions are the locations where a vehicle stays for a certain amount of time, and usually these are the places where boarding and deboarding occur. In this paper, Stay Regions are identified from raw trajectories by a spatial clustering method according to the characteristics of trajectories within Stay Regions, and then, the flow network of populations among regions are generated based on the pass order of Stay Regions and their association to administrative divisions.

(1) Typical characteristics of Stay Regions

As shown in Figure 3, as a vehicle moves, its trajectory points distribute along smooth lines or curves. However, within Stay Regions, a vehicle would have short-time stops or irregular movements due to boarding and deboarding of passengers. Therefore, the number of trajectory points is relatively large, and they are clustered. It can be concluded that the trajectory points within Stay Regions show typical characteristics different from those in other regions: (a) the distribution density is large, and apparent point clusters exist; (b) the stay duration of vehicles is relatively long; (c) and the moving speed is relatively low.

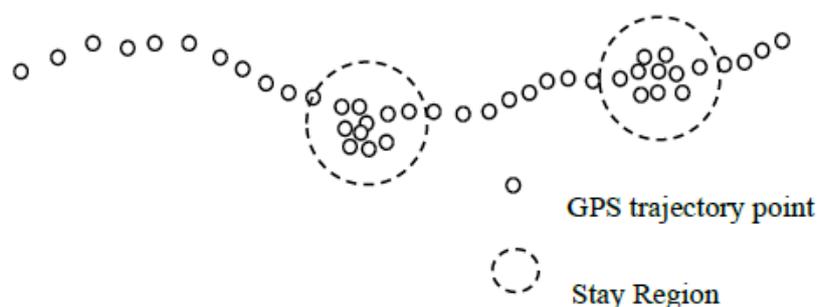


Figure 3. Distribution characteristics of GPS trajectory points within Stay Regions.

(2) Stay Region extraction method

The key to detecting Stay Regions is to find highly dense point clusters existing in raw GPS positioning points. A commonly used method is the spatial clustering technique [27,28]. Popular spatial clustering methods are partitioning methods, hierarchy methods, density-based methods, and grid-based methods [29]. The density-based DBSCAN algorithm is selected to implement clustering since it can detect clusters with arbitrary shapes and is robust against noise points [30]. Two parameters *Eps* and *MinPts* are used to adjust the density of the cluster. *Eps* is a distance parameter used to define the neighborhood of points, while *MinPts* is the minimum number of points required to form a cluster. A small *Eps* and large *MinPts* tend to detect more dense clusters. According

to the characteristics of our trajectory dataset and research needs, the two parameters are set to 30 m and 30, respectively. The change of these two values would bring fluctuations to the number of Stay Regions, but it would not lead to a large change in the flow network result since the generation of the flow network is a global rather than local result. Each detected point cluster represents a Stay Region.

(3) Generation of the flow network of the population among counties

For convenience, the coordinates of the mass point $P_i(X_{P_i}, Y_{P_i})$ of a cluster C_i is calculated using Equation (1).

$$X_{P_i} = \frac{1}{|C_i|} \sum_{O_j \in C_i} X_{O_j}, Y_{P_i} = \frac{1}{|C_i|} \sum_{O_j \in C_i} Y_{O_j} \quad (1)$$

A vehicle trajectory can be split into several segments based on Stay Regions, and each trajectory segment is either a moving segment outside Stay Regions or a stay segment within a Stay Region. The stay segments whose stay duration exceeds a certain value are replaced with the mass points of its corresponding Stay Region. Connecting these mass points chronologically would transform the raw trajectories into high-level moving paths. The pass order of the counties is analyzed to compute the flow of population among regions.

The flow of population is calculated at the scale of the administrative division to correspond to the spatial resolution of other socioeconomic statistics. The thirteen nodes, representing the thirteen counties of Guilin City, were based on locations of county government facilities. Based on these nodes, a directional spatially embedded network (adjacency matrix) is established according to network theory. The links between nodes are directional, and their weights are flow of vehicles from the start node to end node. For each node, the sum of the weights of all links associated to this node is called degree. The sum of the weights of a node is called outdegree if only the links starting from this node are considered, while it is called indegree if only the links ending at this node are taken into account. The established network can reflect the overall flow of populations among counties. A node's degree, indegree, and outdegree represent the total flow, inflow (flow from outside), and outflow (flow to outside) of population in the node's representative county. The weight of a link represents the flow of population between the two counties that are represented by the two nodes of this link.

2.3.2.2. Activity Intensity Analysis of Different Types of Populations

Human activity intensity means the amount of human activity per unit area per unit time. To archive the spatial distribution of human activity intensity from the GPS trajectory data, the measure proposed by Li et al. (2016) is used. The metric is defined as the number of trajectory points per unit area per unit time [20]:

$$f(x) = \frac{1}{\sqrt{2\pi}nh} \sum_{i=1}^n e^{-\frac{(x-x_i)^2}{2h^2}} \quad (2)$$

where h is the bandwidth, n is the number of points within the bandwidth, x_i is the location of point objects, and x is the location to calculate density. The choice of bandwidth is very important. According to Koutsias et al. [31], if the goal is to suggest models about the data, it would be sufficient to choose its value subjectively by visual inspection. Kong et al. (2002) studied the spatial distribution characteristics of human impacted landscape and found the distance of significant impact on the landscape is between 1000 and 1200 m [32], so the bandwidth is set to 1000 m in this research. The grid cell size for calculation is set to 30 m to make the calculation result consistent with the land cover result.

The vehicle trajectory data are then separated according to the type of vehicle: the trajectories of coaches are used to derive local residents' activities while the trajectories of tourist shuttles are used to derive tourists' activities. Through this method, the activity intensity maps of two types of populations are obtained. In this paper, local residents represent the people who reside and live in Guilin regardless of whether they are registered or unregistered, while tourists represent the people who visit Guilin for sightseeing and temporary stays. The categorization of population is different from that in sociology,

but separating these two types of populations and investigating their respective impact on LUCC is very useful especially for the cities such as Guilin with a large number of tourists every year.

2.3.3. Impact Analysis of the Flow and Type of Population on LUCC

Based on the above results, the impact of the flow of population on LUCC and the difference in influence of the two types of populations' activities are explored. As for the former one, a regression analysis method is used to evaluate how the flow of population impacts LUCC compared with other socioeconomic variables. Stepwise regression analysis [33] is a semi-automated method of establishing a model by successively adding or removing variables based solely on the *t*-statistics of their estimated coefficients. It is good at fine-tuning a model and can find a good model fit without redundant variables, compared with other traditional regression methods. Stepwise regression analysis method is often used in LUCC research [34,35], so it is chosen here to pick out the variables that have significant impact on each type of land cover. Similar to previous research [36,37] in which the GDP and population-related variables were selected as major representative socioeconomic variables, the following independent variables are taken into account: GDP (X_1), GDP of the primary industry (X_2), GDP of the secondary industry (X_3), GDP of the tertiary industry (X_4), population (X_5), flow of population (X_6), outflow of population (X_7), and inflow of population (X_8). Because participatory sensing is an emerging technology, the time span of analysis is relatively short. The land cover area instead of the change value is used in the regression analysis, so the dependent variables are: the area of woodland, water, built-up land, crop land, and others. After regression analysis, those independent variables found to be significant are included in the final step. In order to evaluate the relative importance of each impact factor, the standardized regression coefficient is calculated using the following equation:

$$\beta_{X_i} = B_{X_i} \frac{sd_{X_i}}{sd_Y} \quad (3)$$

where X_i is the impact factor to evaluate, Y is the dependent variable, B_{X_i} is the unstandardized regression coefficient, and sd_{X_i} and sd_Y are the standard deviations of the impact factor and the dependent variable, respectively. The higher the standardized coefficient value is, the larger the impact is.

As for the latter one, the human activity of two populations is overlaid on LUCC to investigate their correlation from three aspects: (1) how human activity correlates to LUCC as the activity intensity changes; (2) how the correlation between human activity and LUCC differs from land type to land type; (3) the impact differences of two types of populations on LUCC.

3. Results and Discussion

3.1. Land Cover Change

The land use/cover results of the Lijiang River Basin in 2009 and 2013 were obtained by remote sensing image classification, as shown by Figure 4. By comparing classification results and validation datasets, the classification accuracies for the years 2009 and 2013 are 79.81% and 82.69%, respectively. This is usually acceptable in land cover change research. The areas of each type of land cover in four counties and the percentage of change are shown by Table 1. Overall, both woodland and built-up land experience a certain degree of increase in all four counties, but the area of crop land significantly decreases in all counties except for GUA. Figure 4 and Table 1 also show that in suburban areas, the area of cropland decreases while the area of woodland increases. It can be inferred that China's Grain for Green policy (a program for converting the sloped cropland to forest or grassland) has an effect in the Lijiang River Basin. Another phenomenon it may disclose is that urban sprawl and green engineering cause the reduction of shrub and bare land. This is more obvious in the Guilin Urban Area, the most developed region in Guilin City.

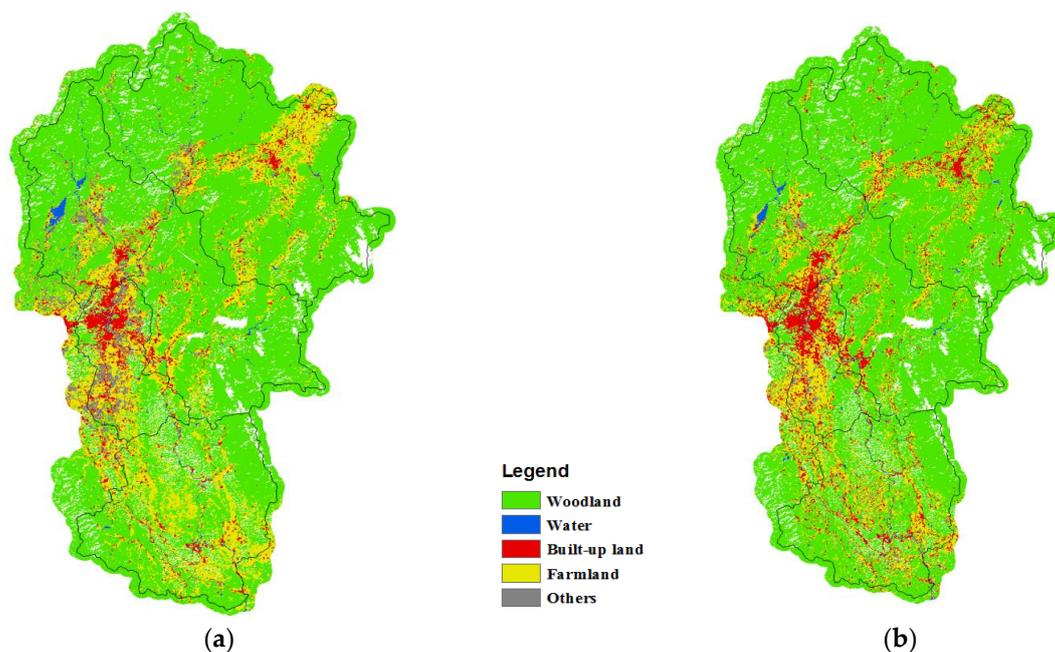


Figure 4. Land use/cover results in the Lijiang River Basin in the year 2009 (a) and 2013 (b) (from Li et al., 2016 [20]).

Table 1. Land cover change in four counties in the study area.

County	Year	Woodland (ha)	Water (ha)	Built-Up Land (ha)	Cropland (ha)	Other (ha)
Guilin Urban Area	2009	15,289.38	920.79	10,708.11	19,514.88	7839.72
	2013	17,338.59	1010.52	13,596.57	19,357.38	2969.82
	change	+13.4%	+9.7%	+27.0%	−0.8%	−62.1%
Xing'an County	2009	172,858	2440.98	6728.04	38,780.10	2298.24
	2013	180,612.6	2236.68	10,070.01	29,173.32	1012.68
	change	+4.5%	−8.4%	+49.7%	−24.8%	−55.9%
Lingchuan County	2009	163,340.1	4218.57	7333.83	33,990.48	6757.92
	2013	168,318.7	3466.98	11,561.85	29,795.76	2497.59
	change	+3.0%	−17.8%	+57.7%	−12.3%	−63.0%
Yangshuo County	2009	88,191.09	1422.54	4676.49	37,541.25	734.31
	2013	88,726.95	1705.86	9169.47	29,284.92	3678.48
	change	+0.6%	+19.9%	+96.1%	−22.0%	+400.9%
Standard Deviation		69,324.3	1173.2	2878.2	7304.6	2564.6

3.2. Flow of Population

Based on the flow analysis method introduced in Section 2.3.2.1, the flow of population in every county of Guilin City is obtained. Table 2 shows the order of the flow value from large to small. The three counties with dense scenic spots rank in the first three places: GUA, YS, and LC. The county with the least flow of population is ZY. This matches well with the facts of Ziyuan County: 93.44% of Ziyuan County is mountainous, and the transportation infrastructure lags behind compared with other counties, which limits the flow of population with the outside.

Table 2. Population flow of every county of Guilin City.

County	Total Flow	Flow to Outside	Flow from Outside
GUA	4769	2452	2317
YS	1690	638	1052
LC	818	422	396
YF	654	346	308
XA	653	358	295
LG	557	268	289
LP	452	220	232
QZ	416	217	199
LS	395	213	182
GC	393	248	145
GY	334	164	170
PL	332	196	136
ZY	189	108	81

We also found that the flow of population shows a large difference in magnitude over space, and the population flow is gradually becoming weak from GUA to the border counties. Figure 5a shows a high level moving path of vehicles based on the Stay Regions, and Figure 5b demonstrates the generated flow network of the population between any two counties. For convenience, the flow network of population is classified to five levels represented by different colors based on the magnitude of the flow. The first level is interchanges between GUA and its neighboring counties. Among all the counties, the link between GUA-YS is the strongest one, and the flow of populations and vehicles between these two counties accounts for a significant proportion of the whole flow network. The second level is mainly the links between GUA and relatively far counties, while the flows at the other levels mainly occur between GUA and border counties or among suburban counties. It can be inferred that the flow of population in the study area centered around GUA, which is determined by its geographic superiority and high economic level.

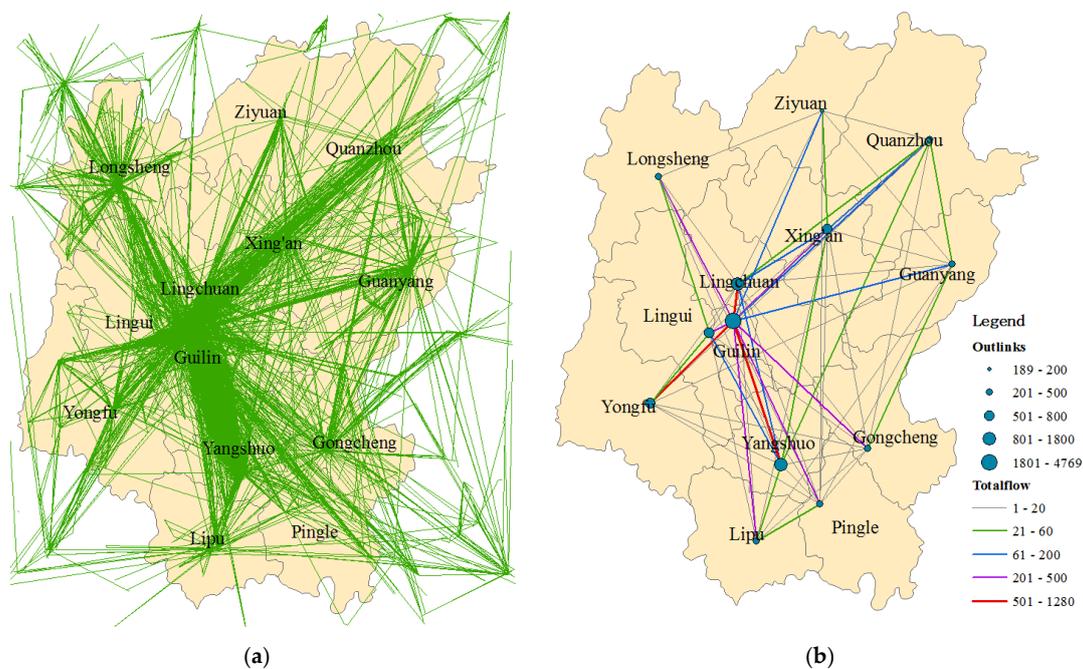


Figure 5. Flow of population among the counties in Guilin City. (a) high level moving paths of vehicles based on Stay Regions, (b) the flow network of the population among counties at five levels: first level (red), second level (purple), third level (blue), fourth level (green), and fifth level (grey).

3.3. Activity Difference between Two Types of Population

The spatial distribution of the activity intensity of local people and tourists is shown by Figures 6a and 6b, respectively. It can be seen that local people's activities concentrate in GUA, along major roads and in some towns, while tourists' activities concentrate in GUA, YS, LC, XA, and major scenic sites. By overlaying the activity intensity of the two types of populations, the intensity difference is calculated as shown in Figure 6c. A positive value indicates that the activity intensity of the local people is stronger than that of the tourists, and the red color represents a much stronger value while yellow represents a slightly stronger value. A negative value indicates that the activity intensity of the local people is weaker than that of the tourists, and the dark blue color represents a much weaker value, and the sky blue color represents a slightly weaker value, while the fruit green color represents a nearly equal value. The sky blue and dark blue areas basically reflect the distribution of the scenic spots. It can be concluded that the number of scenic spots in LC is dominant, and there are also a certain number of scenic spots in GUA, YS, and XA.

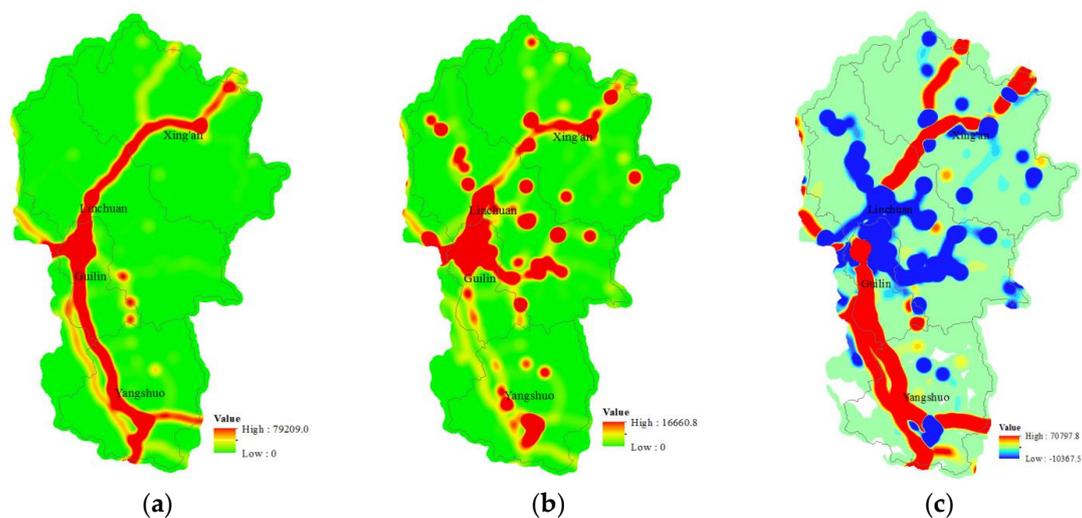


Figure 6. Spatial distribution of activity intensity of two types of population (a) local people, (b) tourists, (c) intensity difference.

3.4. Impact of the Flow of Population on LUCC Based on Regression Analysis

In addition to the population flow variables listed in Table 2, other commonly used socioeconomic variables as shown in Table 3 are also used in the regression analysis. Since a lag time is expected between socio-economic variables and LUCC change [38], the socioeconomic data taken a year before the remote sensing datasets are used.

Table 3. Socioeconomic variables of four counties in the Lijiang River Basin.

County	Year	GDP (Billion Yuan)	GDP 1st (Billion Yuan)	GDP 2nd (Billion Yuan)	GDP 3rd (Billion Yuan)	Population (Thousand)
GUA	2008	28.04498	0.63562	11.24986	16.1595	759.275
	2012	42.71358	0.90363	17.21815	24.5918	731.2
XA	2008	6.2084	1.61461	2.9342	1.65959	371.373
	2012	12.39799	2.56504	7.14563	2.68732	381.8
LC	2008	5.70418	1.63284	2.51892	1.55242	365.923
	2012	10.57084	2.72893	5.11033	2.73158	381.1
YS	2008	3.96708	1.06224	1.18517	1.71967	312.364
	2012	7.62279	1.80504	2.65376	3.16399	317.6
Standard deviation		13.646	0.751	5.489	8.706	182.754

The stepwise regression analysis is conducted between the socioeconomic variables and LUCC. Table 4 shows the regression functions of five types of land cover with determination coefficient, root mean squared error (RMSE), F-statistic, and p -values. The p -values of all five functions are smaller than 0.05, which proves that the correlation is significant.

Table 4. Regression function between land cover and socioeconomic statistics.

Land Type	Regression Function	R^2	Modified R^2	RMSE	F-Statistic	p -Values
woodland	$Y = 0.0101X_3 + 1100.19X_5 - 104.266X_8 + 161554$	0.99982	0.999687	1225.78	7461.84	5.98e-8
water	$Y = -1.05748X_8 + 3251.21$	0.60308	0.536927	798.374	9.1164	0.0234
Built-up land	$Y = 0.00232X_1 + 0.0242X_2 + 1907.51$	0.93050	0.902696	897.813	33.4699	0.0013
cropland	$Y = -0.0766X_2 - 5.963X_6 + 3.559$	0.95269	0.933766	1879.92	50.3426	0.0005
other	$Y = -0.0108719X_3 + 39950.5$	0.81162	0.780222	3105.35	25.8503	0.0023

The regression functions disclose two interesting phenomena: (1) human activity within a region and the flow of population among regions have different impacts on LUCC. The impact varies from land class to land class; (2) the flow of population has directional properties, and the flows in different directions produce different impacts on LUCC.

After stepwise regression, three variables remained for woodlands: GDP of the secondary industry, population, and inflow of population. According to Equation (3), their standardized regression coefficients are 0.078, 0.29, and -1.296 , respectively. It indicates that the area of woodland is negatively correlated to the inflow of population, while it is positively correlated to population and GDP of the primary industry. It can be seen by comparing the coefficient values that the inflow of population has a larger impact on woodland than the population. The population of LC, XA, and YS are 365,923, 371,373, and 312,364, respectively, which are similar, but among these three counties, the smallest area of woodland is 88,191 hectares, and the largest area is 172,858 hectares, which are much different. Therefore, the population is apparently not a major impact factor of the woodland. Among all seven variables, the inflow of population forms a better linear relation to the woodland area. It can be inferred that the large flow of population from outside would be accompanied by the decrement of woodland.

After the regression analysis, the inflow of population was the only remaining variable for water. The standardized regression coefficient is -0.7766 . This shows that similar to woodland, the inflow of population has a larger impact compared with other socioeconomic variables, and the inflow of population and water area are negatively related to each other. It can be inferred that the more population that flows in, the more severe issues the water body faces.

Two variables remained for built-up land: GDP and GDP of the primary industry. Their regression coefficients are 1.1042 and 0.6316, respectively. The first phenomenon that the regression function discloses is consistent with our common knowledge: GDP is a dominant factor for determining the area of built-up land. Another interesting result is that the built-up land area is also related to the agricultural production. This indicates that the urban sprawl of Guilin City is accompanied by the development of the agricultural industry.

Two variables also remained for crop land: GDP of the primary industry and total flow of the population. Their regression coefficients are -0.7878 and -1.4455 , respectively. It can be seen that the area of crop land is negatively correlated to both remaining variables. This indicates that the total flow of population can reflect the change of crop land more than the other variables, including the GDP of the primary industry. This indicates that the more connections a county has with the outside, the less it relies on the agricultural industry.

3.5. Correlation between Two Types of Populations and LUCC

By overlaying the spatial distribution of human activity intensity on land use/cover, Figure 7 was generated. It shows the correlation between the area percentage of each land cover and the activity

intensity of the local people and tourists in 2009 and 2013. Figure 7a–d corresponds to woodland, water, built-up land, and crop land, respectively.

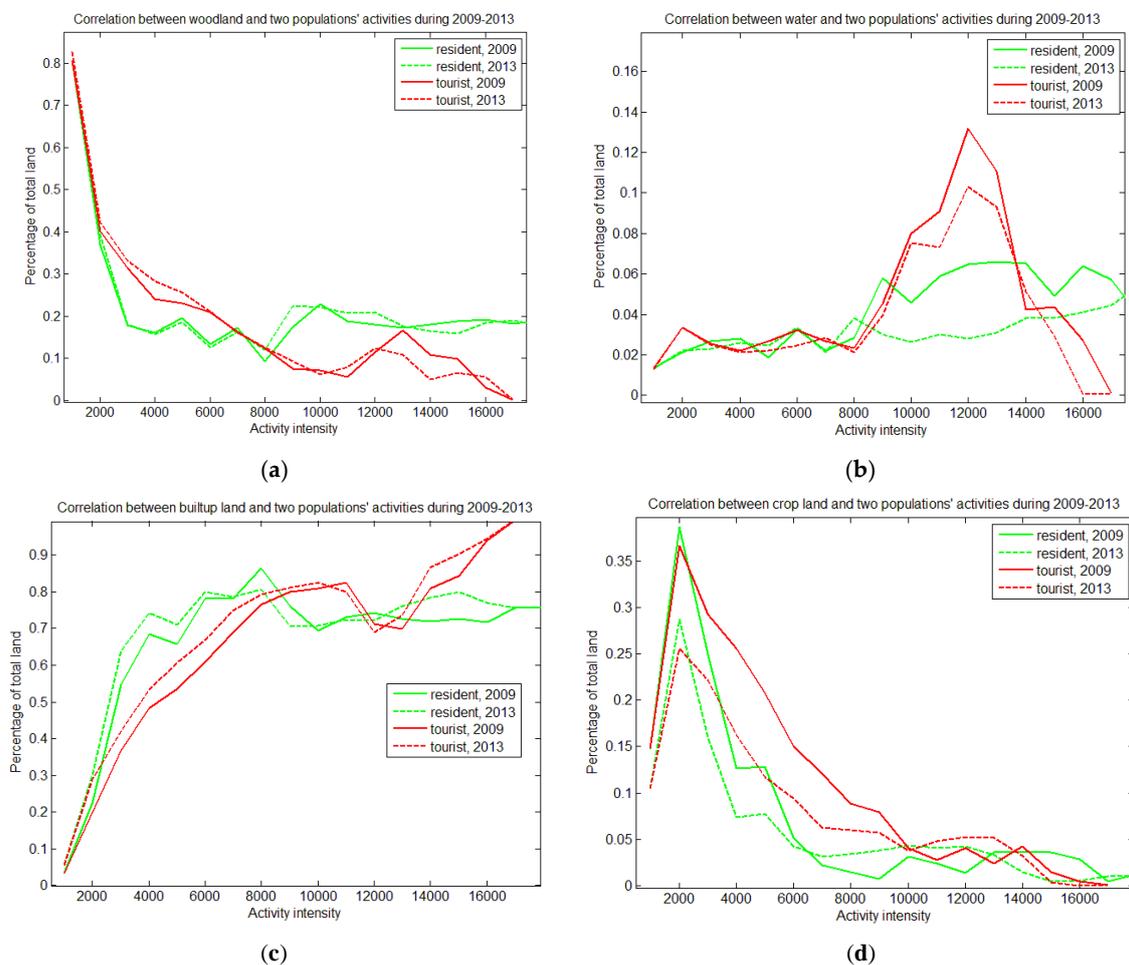


Figure 7. Correlation between two populations' activities and each land type from 2009 to 2013 (a) woodland; (b) water; (c) built-up land; (d) cropland.

The correlation between the activities of the two populations and the same type of land were generally consistent. However, there were differences in details. The experimental results demonstrate the following phenomena: (1) woodland limits human activity in the study area, and the limiting effect on local residents is more severe than that on tourists; (2) humans prefer to live or visit the regions with more water, and tourists have more of this need than the local people; (3) tourists prefer to visit crowded places, but local people are less sensitive to the development degree of a place; (4) most human activity is distant from crop land, and the decrement of crop land is more correlated to tourists' activity than local people's.

Specifically, Figure 7a shows that human activity and woodland percentage are overall negatively correlated. As the percentage of woodland decreases, the intensity of human activity increases. It can be inferred that woodland restrains human activity. By comparing the slope and the declining range of the red and green curves, it can be found that the restraining effect of woodland on local residents is more severe than that on tourists. This could be attributed to several woody scenic spots and nature reserves, such as the Maoershan National Nature Reserve, that attract a certain number of tourists.

Unlike woodland, the percentage of water is generally positively correlated to the activities of the two populations, as shown by Figure 7b. It indicates that both types of populations prefer to live in or visit the regions with more water. The peak value of water percentage for tourists is larger

than that for local people, showing that tourists have more of this need than the local people. This is probably because water is an essential component of the scenic spots in this area. For both populations, the percentage of water drops as time goes on, and this is relatively obvious in the areas with high activity intensity (above 8000). For the same intensity level of activities, the decrement of the water percentage corresponding to tourists is smaller than that corresponding to the local people. Therefore, it can be inferred that the impact of local people's activity on the water decrement is stronger than that of tourists' activity.

The above results about the relationship between woodland, water, and human activity reinforce two previous research findings [39,40] in a quantitative way: water has a high aesthetic preference, while dense forests are less preferred.

As for built-up land, it is also positively correlated to the activity of the two populations, as shown by Figure 7c. For local people, the land percentage evolves to being stable after a small increment of activity intensity, while for tourists, the land percentage continuously goes up. This indicates that tourists prefer to visit crowded places. By contrast, local people are less sensitive to the development degree of a place. Looking at the change of the land percentage with time, it can be found that human activity generally promotes the increment of the built-up land regardless of the type of populations.

Figure 7d shows the correlation between the area percentage of crop land and activities of the two populations. Similar to woodland, at an overall level, human activity and crop land percentage are also negatively correlated. Since crop land is only for farming and cannot provide other functions, most human activity is distant from crop land. The slope of the curves show that local people's activity is more sensitive than tourists' activity. It can also be seen by comparing the curves in two colors that the percentage of crop land decreases during the four year period. An interesting phenomenon is that the decrement amount for tourists is larger than that for local people. A possible reason could be that some villages transform crop land into tourism-related facilities or land to serve tourists.

3.6. Discussion

The Lijiang River Basin is a scenic spot with intensive human activity of both local people and tourists. Much research has been conducted in this area exploring the human impacts on LUCC [23,41–43]. For instance, Chen and Bao (2010) and Lu et al. (2015) analyzed the impact of human activity on land use/cover change processes based on their knowledge of local policies and socioeconomic activities [41,42]. Our method differs from their approaches in that we used GPS trajectory data to directly monitor human activities which does not rely on subjective knowledge. Xiang and Meng (2014) and Mao et al. (2014) calculated the landscape structure variation coefficients of the regions surrounding residences and scenic spots, respectively, to evaluate the disturbance range of residents' and tourists' activities [23,43]. Compared with their approaches, our method can quantify human activity and obtain its distribution difference over space which facilitates fine-grained analysis.

This paper makes an attempt to correlate human activity with LUCC. However, it must be stressed that the proposed method is based on some basic assumptions that may not always hold true: first, the number of passengers is linearly correlated to the trips taken by the bus. According to [44,45], the number of private cars per capita in the study area is extremely small (around 0.034), so the tourist shuttles and coaches are the two dominant transportation tools. This assumption is acceptable in our study area, but it should be performed with great care in regions with a large number of private cars per capita; second, the spatial distribution of the travel activity intensity can reflect that of the entire human activity. Guilin City is a city with a high proportion of tertiary industries, so people's travel activity is highly related to the entire human activity. In addition, this study was based solely on the GPS trajectory data of vehicles and, therefore, excluded other human activity that may be relevant. This would bring a certain amount of bias to the results drawn from analyzing the correlation between human activity and LUCC or from the regression analysis. For example, the regression coefficient could be larger or smaller than the true value, but it would not influence the general pattern too much. The results would be more comprehensive and accurate if more sources of participatory sensing

data (e.g., mobile phone positioning data) were used. Notwithstanding the above simplifications, the proposed method still has good extensibility and can be applied to various kinds of regions by selecting appropriate data sources and adjusting parameters.

4. Conclusions

The human influence on LUCC is increasingly becoming a hot research issue in the field of geography. However, the study thereof is difficult due to the lack of direct monitoring tools to gauge human activity. In order to fill this gap, we selected the Lijiang River Basin as a study area, introduced the emerging participatory sensing technology, and analyzed the impact of human activity on LUCC from the perspectives of flow and type of population: (1) we used the kernel density estimation to extract the spatial distribution of activity intensity of local people and tourists from GPS trajectory data, and overlaid human activity on LUCC to explore the correlation characteristics between the two types of populations and each land cover; (2) we used the DBSCAN algorithm and a network analysis method to build a flow network of population from raw trajectories, and conducted regression analysis to explore the influence of the flow of population on the change of each land cover.

The above experimental results demonstrate that the proposed method can be used to reveal the impact mechanism of human activity on LUCC at fine-grained levels, and provide more accurate planning and instructions for ecological management. Our method also brings to light that the activities of different types of populations have different correlations with land cover change even when the activity intensity is the same, and the correlation varies from land cover to land cover. We also revealed that the flow of population is an important factor driving LUCC, and in some cases, is a greater factor in explaining LUCC rather than the static distribution of population.

Further work is needed to make this work more comprehensive. First, the flow of population is currently only analyzed at the county-level. Using different scales of population flow may produce different results, and finding an optimum analysis scale would be an important aspect in our future work. Furthermore, when a long time period of GPS trajectory data is accumulated, it is also important to investigate how the change of human activity intensity over time influences LUCC.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (No. 41601485), the joint fund of the National Natural Science and Civil Aviation research foundation of China (No. U1533114) and the National Science and Technology Pillar Program of China (No. 2012BAC16B04). The authors would like to thank the two anonymous reviewers whose comments and suggestions greatly improved the manuscript.

Author Contributions: Jun Li and Qiming Qin conceived and designed the experiments; Yuan Zhang performed the experiments; Jun Li and Yueguan Yan analyzed the data; Jun Li wrote the paper.

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

References

1. Pan, Y. Human activities and geographical conditions. *Yunnan Geogr. Environ. Res.* **1990**, *2*, 19–28.
2. Ramankutty, N.; Evan, A.T.; Monfreda, C.; Foley, J.A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **2008**, *22*, 1–29. [[CrossRef](#)]
3. Ellis, E.C.; Goldewijk, K.K.; Siebert, S.; Lightman, D.; Ramankutty, N. Anthropogenic transformation of the biomes, 1700 to 2000. *Glob. Ecol. Biogeogr.* **2010**, *19*, 589–606. [[CrossRef](#)]
4. Lambin, E.F.; Geist, H.J. *Land-Use and Land-Cover Change: Local Processes and Global Impacts*; Springer: Berlin, Germany, 2008.
5. Zheng, H.; Ouyang, Z.Y.; Zhao, T.Q.; Li, Z.X.; Xu, W.H. The impact of human activities on ecosystem services. *J. Nat. Resour.* **2003**, *18*, 118–126.
6. Fu, B.; Lu, Y.; Chen, L.; Su, C.; Yao, X.; Liu, Y. The latest progress of landscape ecology in the world. *China Acta Ecol. Sin.* **2008**, *28*, 798–804.
7. Lü, Y.; Chen, L.; Fu, B. The analysis of human activities and landscape pattern at the county level. *Acta Ecol. Sin.* **2004**, *24*, 1833–1838.

8. Calo, F.; Parise, M. Evaluating the human disturbance to karst environments in Southern Italy. *Acta Carsol.* **2006**, *35*, 47–56. [[CrossRef](#)]
9. Guo, L.; Xia, B.; Yu, S.; Gong, C. Effect of anthropogenic disturbances on the temporal-spatial changes of landscape patterns at Taishan Mountain. *Chin. J. Eco-Agric.* **2006**, *14*, 235–239.
10. Wang, L.; Wehrly, K.; Breck, J.E.; Kraft, L.S. Landscape-based assessment of human disturbance for Michigan Lakes. *Environ. Manag.* **2010**, *46*, 471–483. [[CrossRef](#)] [[PubMed](#)]
11. Garbarino, M.; Lingua, E.; Marzano, R.; Urbinati, C.; Bhujju, D.; Carrer, M. Human interactions with forest landscape in the Khumbu valley, Nepal. *Anthropocene* **2014**, *11*, 544–554. [[CrossRef](#)]
12. Roth, D.; Moreno-Sanchez, R.; Torres-Rojo, J.M.; Moreno-Sanchez, F. Estimation of human induced disturbance of the environment associated with 2002, 2008 and 2013 land use/cover patterns in Mexico. *Appl. Geogr.* **2016**, *66*, 22–34. [[CrossRef](#)]
13. Goldman, J.; Shilton, K.; Jeff, B.; Estrin, D.; Hansen, M.; Ramanathan, N.; Reddy, S.; Samanta, V. *Participatory Sensing: A Citizen-Powered Approach to Illuminating the Patterns that Shape our World*; Woodrow Wilson International Center for Scholars: Washington, DC, USA, 2009.
14. Thiemann, C.; Theis, F.; Grady, D.; Brune, R.; Brockmann, D. The structure of borders in a small world. *PLoS ONE* **2010**, *5*, e15422. [[CrossRef](#)] [[PubMed](#)]
15. Kang, C.G.; Liu, Y.; Ma, X.J.; Wu, L. Towards estimating urban population distributions from mobile call data. *J. Urban Technol.* **2012**, *19*, 3–21. [[CrossRef](#)]
16. Liu, Y.; Wang, F.; Xiao, Y.; Gao, S. Urban land uses and traffic source-sink areas: Evidence from GPS-enabled taxi data in Shanghai. *Landsc. Urban Plan.* **2012**, *106*, 73–87. [[CrossRef](#)]
17. Shen, Y.; Chai, Y.; Kwan, M.P. Space-time fixity and flexibility of daily activities and the built environment: A case study of different types of communities in Beijing suburbs. *J. Transp. Geogr.* **2015**, *47*, 90–99. [[CrossRef](#)]
18. Li, J.; Wang, J.J.; Zhang, J.F.; Qin, Q.M.; Jindal, T.; Han, J.W. A probabilistic approach to detect mixed periodic patterns from moving object data. *Geoinformatica* **2016**, *20*, 715–739. [[CrossRef](#)]
19. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.G.; Zhi, Y.; Chi, G.H.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [[CrossRef](#)]
20. Li, J.; Zhang, Y.; Wang, X.; Qin, Q.M.; Wei, Z.; Li, J.Z. Application of GPS Trajectory Data for Investigating the Interaction between Human Activity and Landscape Pattern: A Case Study of the Lijiang River Basin, China. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 104. [[CrossRef](#)]
21. Guilin Chorography Office. *Guilin Chorography (1991–2005)*; Guilin Chorography Editorial Committee: Guilin, China, 2011.
22. Xiang, W.; Li, X.; Ding, T.; Huang, Y.; He, C.; Lu, S. Analysis on the vegetation landscape pattern of upper reaches in Lijiang River Basin. *Guangxi Sci.* **2009**, *16*, 455–459.
23. Xiang, Y.; Meng, J. Temporal and Spatial Analysis of Tourism Disturbance on Landscape Pattern in the Li River Basin of Guangxi. *Mt. Res.* **2014**, *32*, 11–20.
24. Lucas, R.; Rowlands, A.; Brown, A.; Keyworth, S.; Bunting, P. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 165–185. [[CrossRef](#)]
25. Rahman, M.D.; Saha, R.; Saha, S.K. Multi-resolution segmentation for object-based classification and accuracy assessment of land use/land cover classification using remotely sensed data. *J. Indian Soc. Remote Sens.* **2008**, *36*, 189–201. [[CrossRef](#)]
26. Zheng, V.; Zheng, Y.; Xie, X.; Yang, Q. Collaborative location and activity recommendations with GPS history data. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 1029–1038.
27. Pei, T.; Wan, Y.; Jiang, Y.; Qiao, Y. Detecting arbitrarily shaped clusters using ant colony optimization. *Intl. J. Geogr. Inf. Sci.* **2011**, *25*, 1575–1595. [[CrossRef](#)]
28. Murray, A.; Grubestic, T.; Wei, R. Spatially significant cluster detection. *Spat. Stat.* **2014**, *10*, 103–116. [[CrossRef](#)]
29. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann: Waltham, MA, USA, 2011.
30. Ester, M.; Kriegel, H.P.; Sander, J.; Xu, X.W. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Available online: <http://www.dbs.ifi.lmu.de/Publikationen/Papers/KDD-96.final.frame.pdf> (accessed on 3 March 2017).

31. Koutsias, N.; Kalabokidis, K.D.; Allgöwer, B. Fire occurrence patterns at landscape level: Beyond positional accuracy of ignition points with kernel density estimation methods. *Nat. Resour. Model.* **2004**, *17*, 359–375. [[CrossRef](#)]
32. Kong, N.N.; Zeng, H.; Li, S.J. A study of the spatial distribution characteristics of human landscape impact in Wolong National Reserve, Sichuan Province. *Acta Sci. Nat. Univ. Pekin.* **2002**, *38*, 393–399.
33. Draper, N.; Smith, H. *Applied Regression Analysis*, 2nd ed.; John Wiley & Sons, Inc.: New York, NY, USA, 1981.
34. Huang, K.C.; Huang, T. Simulation of land-cover change in Taipei Metropolitan Area under climate change impact. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *18*, 012106. [[CrossRef](#)]
35. Khaledian, Y.; Kiani, F.; Ebrahimi, S.; Brevik, E.C.; Aitkenhead-Peterson, J. Assessment and monitoring of soil degradation during land use change using multivariate analysis. *Land Degrad. Dev.* **2017**, *28*, 128–141. [[CrossRef](#)]
36. Laurance, W.F.; Albernaz, A.K.M.; Schroth, G.; Fearnside, P.M.; Bergen, S.; Venticinque, E.M.; Da Costa, C. Predictors of deforestation in the Brazilian Amazon. *J. Biogeogr.* **2002**, *29*, 737–748. [[CrossRef](#)]
37. Gong, C.; Yu, S.; Joesting, H.; Chen, J. Determining socioeconomic drivers of urban forest fragmentation with historical remote sensing images. *Landsc. Urban Plan.* **2013**, *117*, 57–65. [[CrossRef](#)]
38. Bonilla-Moheno, M.; Aide, T.M.; Clark, M.L. The influence of socioeconomic, environmental, and demographic factors on municipality-scale land-cover change in Mexico. *Reg. Environ. Chang.* **2012**, *12*, 543–557. [[CrossRef](#)]
39. Ulrich, R. Aesthetic and affective response to natural environment. In *Behavior and the Natural Environment*; Altman, I., Wohlwill, J., Eds.; Springer: New York, NY, USA, 1983; pp. 85–125.
40. Hunziker, M.; Kienast, F. Potential impacts of changing agricultural activities on scenic beauty—A prototypical technique for automated rapid assessment. *Landsc. Ecol.* **1999**, *14*, 161–176. [[CrossRef](#)]
41. Chen, Z.G.; Bao, J.G. Land use change dynamics and its determining mechanism in a typical tourist city: The case of Yangshuo County, Guangxi. *Sci. Geogr. Sin.* **2010**, *30*, 544–550.
42. Lu, D.D.; Zhong, Y.D.; Yang, Y.D.; Chen, S.Y. Research on land use/land cover changes and driving force in upper reach of Li River Basin. *Ecol. Econ.* **2015**, *31*, 128–131.
43. Mao, X.Y.; Meng, J.J.; Wang, Q. Modeling the effects of tourism and land regulation on land-use change in tourist regions: A case study of the Lijiang River Basin in Guilin, China. *Land Use Policy* **2014**, *41*, 368–377. [[CrossRef](#)]
44. National Bureau of Statistics of the People’s Republic of China. Statistical Bulletin of National Economy and Social Development of China in 2012. Available online: http://www.gov.cn/gzdt/2013-02/22/content_2338098.htm (accessed on 1 April 2014).
45. Guangxi Zhuang Autonomous Region Bureau of Statistics. Statistical Bulletin of Economy and Social Development of Guangxi in 2012. Available online: http://www.gxtj.gov.cn/fzlm/zd gz/201304/t20130401_26110.html (accessed on 1 April 2014).



© 2017 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/>).