Integrating Spatial and Attribute Characteristics of Extended Voronoi Diagrams in Spatial Patterning Research: A Case Study of Wuhan City in China

Zuohua Miao *, Yong Chen, Xiangyang Zeng and Jun Li

School of Resources and Environmental Engineering, Wuhan University of Science and Technology, No. 947 Heping Road, Wuhan 430081, China; yongchen1968@163.com (Y.C.); prelian@sina.com (X.Z.); larkiner@gmail.com (J.L.)

* Correspondence: miaozuohua@wust.edu.cn; Tel.: +86-027-6886-2892
Academic Editors: Qiming Zhou and Wolfgang Kainz
Received: 26 April 2016; Accepted: 11 July 2016; Published: 15 July 2016

Abstract: Rapid urbanization has caused numerous problems, and the urban spatial structure has been a hot topic in sustainable development management. Urban spatial structure is affected by a series of factors. Thus, the research model should synthetically consider the spatial and non-spatial relationship of every element. Here, we propose an extended Voronoi diagram for exploring the urban land spatial pattern. In essence, we first used a principal component analysis method to construct attribute evaluation indicators and obtained the attribute distance for each indicator. Second, we integrated spatial and attribute distances to extend the comparison distance for Voronoi diagrams, and then, we constructed the Voronoi aggregative homogeneous map of the study area. Finally, we make a spatial autocorrelation analysis by using GeoDA and SPSS software. Results show that: (1) the residential land cover aggregation is not significant, but spatial diffusion is obvious; (2) the commercial land cover aggregation is considerable; and (3) the spatial agglomeration degree of the industrial land cover is increased and mainly located in urban fringes. According to the neo-Marxist theory, we briefly analyzed the driving forces for shaping the urban spatial structure. To summarize, our approach yields important insights into the urban spatial structure characterized by attribute similarity with geospatial proximity, which contributes to a better understanding of the urban growth mechanism. In addition, it explicitly identifies ongoing urban transformations, potentially supporting the planning for sustainable urban land use and protection.

Keywords: geospatial proximity; attribute similarity; generalized Euclidean distance; Voronoi diagrams; spatial pattern

1. Introduction

Although urban surfaces currently cover only between 3% and 5% of the total land surface of the Earth, environmental impacts in relation to urban growth have become a major concern around the world [1]. Rapid urban growth due to large-scale land use/cover change, particularly in developing countries, has drawn considerable attention since urbanization drives environmental change at multiple scales [2–7]. Recent decades have witnessed unprecedented urban landscape changes, rapid land use changes found in urbanization processes, loss of ecosystems and biodiversity and the exploration of available resources, leading to scarcity [5–11]. These transformations raise a series of questions: Is our urban spatial structure sustainable development? How can policy most effectively shape urban morphology and manage urban spatial structure, expansion and decline or agglomeration and dispersion? What guidelines should we provide for policymakers and regional planners to note the future development of our urban environment? Addressing such questions...
requires an explicit model to discover the aggregation tendency of spatial entities, the distribution rules, the development and change tendency and the driving forces for shaping the urban spatial structure from spatial data [12,13].

Urban spatial structure (also known as urban internal spatial structure or urban space) refers to the arrangement of land use in urban areas, which concerns the arrangement of public and private space in cities and the degree of connectivity and accessibility [14]. The characterization and modeling of urban spatial structure has been highlighted as an important research activity in recent years, especially in those countries with rapid urbanization and industrialization, such as China [15–17].

In the past, various theories, such as central place theory, core-periphery theory, Losch's demand cone theory, spatial diffusion theory, growth pole theory, concentric rings theory, sector theory, Zipf's law and multiple nuclei theory were extensively used to study urban structure patterns and dynamics [1,18–22]. However, urban growth is a complex spatial system, comprised of numerous acting and interacting elements with feedbacks in between, which inevitably determine the overall spatial form of the city [23]. Therefore, these conventional urban geography theories are unable to characterize the spatiotemporal complexity of urban dynamics [1,24,25]. In response, a variety of techniques, such as catastrophe theory, fractals theory, chaos or chaotic attractors theory, self-organizing theory, entropy theory, landscape metrics, cellular automata and agent-based modeling, have emerged since the 1960s to analyze the urban structure and dynamics [1,23,26–31].

The modeling of urban spatial structure has occupied researchers for decades, and an abundant literature has emerged. The advancements of mathematical theories and computer-based simulations based on equations to seek a static or equilibrium solution have helped us to better understand cities [24]. Most common mathematical models are sets of simultaneous joint equations based on theories of population growth and diffusion and economic theory that specify cumulative urban spatial structure over time [32–34]. A major drawback of such mathematical models is that a numerical or analytical solution to the system of equations must be obtained, limiting the level of complexity that may practically be built into such models [35]. The complex systems science modeling techniques offer a potentially appropriate toolbox, notably agent-based modeling and cellular automata, potentially linked to Geographic Information System (GIS) and Remote Sensing (RS) information on land parcels [1,6,23,36–38]. Techniques, such as Cellular Automata (CA), as rule-based models construct a ‘bottom-up’ approach, where the structure evolves from a Moran neighborhood interacting between neighboring cells using different types of transition rules [31,39–41]. Moreover, the conventional urban CA model assumption that a regular cell represents the basic unit of a land use entity is not always reasonable [42]. Several studies deal with the development of vector-based or irregular CA models; however, these new kinds of CA models are still in their infancy [43,44].

The spatial structure of some social and economic variables may be a reflection of spillover effects, which usually are evident from the form of spatial association. Agglomeration economies, on the other hand, may produce differential effects over space [45]. In urban studies, this usually means that high variable values tend to concentrate near other high values, and low values appear in geographical proximity. In this sense, spatial association is the non-random ordering of data values when they are arranged by spatial location and can be analyzed by using distance statistics [45–47]. Geospatial clustering is a series of processes that groups a distinct point set into a number of groups exhibiting similar characteristics according to the geospatial proximity. Points belonging to the same geospatial cluster are more similar to each other than points belonging to different clusters [48]. Analysis of spatial clustering has been a major topic of research in geographical pattern research [49,50]. Spatial statistics methods were introduced in geology and geography in the 1950s, but have only recently started to serve the purposes and needs of urban analysis, specifically applied to the analysis of urban land prices [51–53]. In spatial statistics, spatial data analysis is often categorized into three types, namely point data analysis, lattice data analysis and geostatistics [54]. Geostatistics concerns the spatial pattern of an attribute on a regular or irregular spatial lattice, with an additional goal of predicting the values of the attribute at unsampled locations [54]. Geostatistics uses distance-based functions rather
than neighborhood structures to represent spatial autocorrelation, and the detection of clusters is based on local spatial statistics, such as the local Moran’s I statistic. Geostatistical data refer to spatial data sampled at point locations that are continuous in space [54–56]. Given a set of a finite number of points in the Euclidean plane, we associate each location in the space with its closest member(s) of the point set with respect to the Euclidean distance. The result is a tessellation of the plane into a set of regions associated with individual members of the point set. This tessellation is called the “ordinary planar Voronoi diagrams”, generated by the sample point set [57]. As one of the spatial data analysis tools, Voronoi diagrams can be used to implement the spatial subdivision and reveal the scope of spatial influence generated by urban areas.

The objective of this study is to identify the urban spatial structure in Wuhan of P.R. China by applying spatial data statistics, Voronoi diagrams, GeoDA and GIS. We consider that our analysis is useful in the following sense. First, we provide an extended Voronoi diagram method, extending the traditional spatial comparison distance to generalized Euclidean distances for Voronoi diagrams; a social, economic and spatial index is built to describe the geospatial proximity distance and attribute similarity distance, for detecting urban land use clustering hubs, centers and edges using spatial and attribute characteristics data. Second, we present our methods through an appropriate workflow of the way these techniques are applied. Third, we propose to measure the spatial autocorrelation of the urban spatial structure using Local Moran’s I based on GeoDA. Finally, although the conventional urban science explains urban phenomena partially with factors of technology and population migrations, we briefly analyze the driving forces for shaping the urban spatial structure according to the neo-Marxist theory. To summarize, the cluster concentrations of urban areas are one of the key dimensions of the differences between regions; the spatial statistic defines spatial clusters based on the correlation between each region and a weighted average of its neighbors. This paper reports that the method could be beneficial to relevant urban studies and urban land use/cover management projects.

2. Study Area and Data

2.1. Study Area

Wuhan is located centrally in the People’s Republic of China, the biggest inland port city and the most important city in central China. Wuhan’s urban civilization dates back 3500 years to Panlong Town, which is the oldest city unearthed to date in the Yangtze River basin. As the capital of Hubei province, Wuhan serves as the political, economic and cultural center of the province. The Yangtze is the world’s third longest river, and its largest tributary, Hanjiang, meets it in Wuhan [58,59] (Figure 1a).

Yangtze and Hanjiang divide the city into the three parts of Hankou, Hanyang and Wuchang, namely the “three towns of Wuhan”. The three towns have different spatial characteristics and different developmental tendencies for historical reasons. Among them, Wuchang is a center for culture and education, and the residential land cover is widely distributed. Hankou is a commercial center and a commercial land cover concentrated distribution. Hanyang is an industrial center, and industrial land cover is distributed in the urban fringe. As a typical garden city featuring mountains and water, Wuhan is home to hundreds of hills and nearly 200 lakes of various sizes. With water comprising 25.8% of its area, Wuhan is ranked first among major Chinese cities in water resources [60] (Figure 1b).

As the largest water, land and air transportation hub in inland China, Wuhan is a major transportation hub, with dozens of railways, roads and expressways passing through the city. The road traffic networks in Wuhan are greatly affected by the Yangtze and Hanjiang, showing the multi-core concentrate networks mode. With the political, economic and commercial center as a core, the road traffic network’s density declines with spatial distance. The road network topological class is a layer-circle decreasing distribution [58] (Figure 1c).

Geographically, Wuhan is located between latitudes 29°58’N and 31°22’N and longitudes 113°41’E and 115°05’E. Topographically, Wuhan belongs to the transitional belt from the southeast highlands to foothills in the south piedmont of Dabie Mountain. The middle part is low and flat, and the south part
and north part are hills. The landform types belong to an alluvial plain of rivers and lakes, with many lakes and rivers [4,15] (Figure 1d).

Geographically, Wuhan is located between latitudes 29°58′N and 31°22′N and longitudes 113°41′E and 115°05′E. Topographically, Wuhan belongs to the transitional belt from the southeast highlands to foothills in the south piedmont of Dabie Mountain. The middle part is low and flat, and the south part and north part are hills. The landform types belong to an alluvial plain of rivers and lakes, with many lakes and rivers [4,15] (Figure 1d).

Figure 1. The spatial patterns of the study area. (a) Spatial location of Wuhan in the whole nation; (b) spatial distribution of land uses in the study area; (c) road network topology distribution of the study area; and (d) topography distribution of the study area.

2.2. Data

The original data used in this paper are vector map data, raster map data and interrelated statistical data. The spatial attribute information of the study area, such as land cover type, road accessibility, topographic gradient, building density, etc., is interpreted from the vector and raster map by using ESRI’s ArcGIS 9.3 software (ESRI Corp. Redlands, CA, USA, 2013). The non-spatial data are also the foundation of spatial pattern analysis, which are processed and calculated by using Excel and VC++ from the interrelated statistical data, such as the Wuhan statistical yearbook. To reduce the computational complexity in the construction of the Voronoi diagram and the large-scale spatial data analysis, we select Hankou as the study area. Hankou has a land area of 139.57 km², with water accounting for approximately 4%.

3. Methodology

3.1. Voronoi Diagrams Used to Compare the Distance Extension

Voronoi diagrams and Delaunay triangulation are very important theoretical methods in computational geometry. The study of Voronoi diagrams can be traced back to the seminal work by Voronoi [49]. Consider some d-dimensional space in which a number of given points
(sometimes referred to as seeds, attractors or generators) is located. To each seed, we assign a Voronoi set that includes all points that are closer to the seed to which it is assigned than to any other seed. The collection of all Voronoi sets is then a Voronoi diagram \[61\]. Voronoi diagrams are a very simple geometrical construct with numerous applications, such as biology, ecology, epidemiology, materials science, machine learning, geography and geology, because of the adjacency similarity and systematic theoretical system \[62–65\]. The Voronoi diagram is an important spatial interpolation method because of its geometric structure. Indeed, it can be used to determine the value of any unknown point based on the nearest known point’s value \[66–68\]. The concept of distance is central to Voronoi diagrams. While there are many types of distances, such as polyhedral distance, Gauges distance and Euclidean distance \[61,69\], we improve the distance by integrating the geospatial proximity distance and attribute similarity distance.

In this paper, the Voronoi diagrams’ character is used to construct a Voronoi diagram model for spatial pattern research based on the generalized Euclidean distance (integrating the geospatial proximity distance and attribute similarity distance) of sampling point datasets. Then, the urban land use spatial clustering homogeneous segmentation of the study area is performed based on the sample data. The extended, weighted Voronoi diagrams can be described as follows:

We define the center point set as \( P = \{p_1, p_2, p_3, \ldots, p_n\} \), where \( p_i \) is any given center point, \((x_i, y_i)\) is \( p_i \)'s spatial location and \( A(p_i) = (a_{i1}, a_{i2}, \ldots, a_{in}) \) expresses its attribute vector. In the conventional Voronoi diagram model, the comparison distance is the geometry distance in the Euclidean plane, and the effect of the attribute similarity distance is not considered. In this paper, the geospatial proximity distance is calculated by the following equation:

\[
D_{\text{Spat}}(p_i, p_j) = w_x (x_i - x_j)^2 + w_y (y_i - y_j)^2
\]

In this paper, \( D_{\text{Attr}}(p_i, p_j) \) is the attribute similarity distance between points \( p_i \) and \( p_j \) and is calculated by the following equation:

\[
D_{\text{Attr}}(p_i, p_j) = \sum_{k=1}^{m} w_k (a_{ik} - a_{jk})^2
\]

The geospatial proximity distance and attribute similarity distance are integrated to define the generalized Euclidean distance of the Voronoi diagrams as follows:

\[
D_{ij} = \sqrt{D_{\text{Spat}}(p_i, p_j) + D_{\text{Attr}}(p_i, p_j)}
\]

where \( w_x \), \( w_y \) are the weight coefficients of the spatial coordinates \( x \), \( y \) and \( w_k \) is the weight coefficient of the \( k \) feature in the attribute feature sets.

In this paper, evaluation indices are established by principal component analysis by using SPSS Statistics 19 (IBM Corp. Armonk, NY, USA, 2010) \[70\]. The evaluation index system indicators include the residential floor area ratio (U1), building density (U2), population density (U3), green area (U4), road accessibility (U5), pipe network status (U6), healthcare (U7), universities and colleges (U8), stadium completion degree (U9), noise index (U10), land price (U11) and land idle rate (U12). In this paper, we index the data into point, line and polygon indices according to the respective influences on the spatial pattern base of the study area’s grid map.

- Point index (including the healthcare index, universities and college index and stadiums index): This index is calculated using the linear model based on the grid center. The basis of the service radius from the inside to the outside is used to calculate the attribute value according to the distance linear attenuation model:

\[
F_i = f_0 (1 - r_i) \quad r_i = D_i/D_0 \quad 0 \leq r_i \leq 1
\]
where $F_i$ is the action mark of the index, $f_0$ is the comprehensive scale index of the numerical standardization involved in Geographic Information System (GIS) buffer analysis, $D_0$ is the maximum impacted distance of this index and $D_i$ is the geometric distance between this index and another spatial point.

- **Line index (including road accessibility):** This index is calculated using the exponential model along the linear target. The basis of the service radius from the inside to the outside is used to calculate the attribute value according to the distance decay exponential model:

$$F_i = f_0(1-r_i) \quad r_i = D_i/D_0 \quad 0 \leq r_i \leq 1$$  \hspace{1cm} (5)

- **Polygon index (including the noise index):** This index is obtained by calculating the action mark directly.

$$F_i = 100 \times \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$  \hspace{1cm} (6)

where $F_i$ is the action mark of the index, $X_i$ is the evaluation value of the grid index and $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of this index, respectively.

Index weights reflect the importance of each indicator in the index system, and they are often determined by, for example, the Delphi method or the Analytic Hierarchy Process (AHP) [71–73]. In this paper, we use AHP to obtain the weight for each index (Table 1).

### Table 1. Evaluation index weighting coefficient table.

<table>
<thead>
<tr>
<th>Index Weight</th>
<th>Weight</th>
<th>Index Weight</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>residential floor area ratio (U1)</td>
<td>0.1592</td>
<td>building density (U2)</td>
<td>0.1304</td>
</tr>
<tr>
<td>green area (U4)</td>
<td>0.1052</td>
<td>road accessibility (U5)</td>
<td>0.0706</td>
</tr>
<tr>
<td>health care (U7)</td>
<td>0.0544</td>
<td>universities and colleges (U8)</td>
<td>0.0476</td>
</tr>
<tr>
<td>noise index (U10)</td>
<td>0.0537</td>
<td>land price (U11)</td>
<td>0.0821</td>
</tr>
<tr>
<td>population density (U3)</td>
<td>0.1001</td>
<td>pipe network status (U6)</td>
<td>0.0583</td>
</tr>
<tr>
<td>pipe network status (U6)</td>
<td>0.0583</td>
<td>stadium completion degree (U9)</td>
<td>0.0424</td>
</tr>
<tr>
<td>pipe network status (U6)</td>
<td>0.0583</td>
<td>land idle rate (U12)</td>
<td>0.0960</td>
</tr>
</tbody>
</table>

#### 3.2. Spatial Aggregation and Separation Index of Voronoi Diagrams

The Aggregation Index (AI) describes the randomness or aggregation degree of landscapes and was proposed by O’Neill in 1988 under the name ‘Contagion Index’ (CI) [74]. The aggregation index reflects the spatial configuration of landscape components [75,76]. Consequently, the land use type Voronoi polygonal neighborhood adjacent relationship is analyzed by calculating the Voronoi polygonal neighborhood aggregation index. Then, the overall spatial agglomeration and dispersion degree of land use are investigated.

- **The Voronoi polygon self-neighborhood aggregation index** is calculated with the following equation:

$$\alpha_{ik} = (\hat{\beta}_{ik} \cdot x_{ik}) / n_i \quad i \in N_k$$  \hspace{1cm} (7)

where $x_{ik}$ is one Voronoi polygon; $\hat{\beta}_{ik}$ is the number of polygons whose land use type is $k$ in the Voronoi polygon neighborhood; $n_i$ is the direct neighborhood polygon number of polygon $x_{ik}$, which is obtained by recording the Voronoi vertex topology; and $N_k$ is the number of Voronoi polygons whose land use type is $k$.

- **The Voronoi polygon regional aggregation index** is calculated as follows:

$$\delta_k = \left( \sum_{i=1}^{N_k} \alpha_{ik} \cdot x_{ik} \right) / N_k$$  \hspace{1cm} (8)
where $\delta_k$ is the sum of $\alpha_{ik}$ and $\alpha_{ik}$ is the Voronoi polygon self-neighborhood aggregation index obtained using Equation (7). $\delta_k$ is used to standardize the treatment: as $\delta_k \to 0$, the land use spatial distribution becomes more dispersed, whereas as $\delta_k \to 1$, the land use spatial distribution tends to coalesce.

### 3.3. Technical Research Flowcharts

The technical research flowchart and main technical processes are presented in Figure 2.

- **Step 1:** The study map is divided into 100 m × 100 m grids by using ESRI’s ArcGIS 9.3 software.
- **Step 2:** A Voronoi diagram based on the center point dataset of sample points is constructed. Then, spatial overlay analysis between the Voronoi diagram and the current land use map is performed to obtain the land use type area for every Voronoi polygon. The spatial and attribute character parameters are saved into a spatial and attribute data table (Table 2). In Table 2, the coordinate is the Voronoi polygon center point, and the triangle ID is the serial number of the Voronoi polygon Delaunay triangle. The vertex and edge datasets represent the vertex and edge of every Voronoi polygon. The land use type reflects the status of each land use type, such as 0, 1 and 2. Aggregation index and clustering level data are obtained in Step 3.
- **Step 3:** The model is used to calculate the neighborhood aggregation index for every Voronoi polygon using Equations (7) and (8). To obtain the clustering grade data, the model performs a cluster analysis based on the shortest distance method using SPSS software. Finally, the model produces the experimental Voronoi diagram map based on the principle of "similar category incorporation and heterogeneous merging".

![Figure 2. Research and Voronoi diagram flowcharts. (a) Technical flowchart used for the experiment; (b) Voronoi diagrams flowchart based on the C# program language and the Visual Studio 2010 platform.](image)

### Table 2. Land use grid spatial and attribute structures.

<table>
<thead>
<tr>
<th>No.</th>
<th>Field</th>
<th>Type</th>
<th>No.</th>
<th>Field</th>
<th>Type</th>
<th>No.</th>
<th>Field</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Voronoi ID</td>
<td>Long</td>
<td>4</td>
<td>Vertex Datasets</td>
<td>CArray</td>
<td>7</td>
<td>Area</td>
<td>Float</td>
</tr>
<tr>
<td>2</td>
<td>Coordinate</td>
<td>TPoint</td>
<td>5</td>
<td>Edge Datasets</td>
<td>CArray</td>
<td>8</td>
<td>Aggregation Index</td>
<td>Float</td>
</tr>
<tr>
<td>3</td>
<td>Triangle ID</td>
<td>CList</td>
<td>6</td>
<td>Land Use Type</td>
<td>Int</td>
<td>9</td>
<td>Clustering level</td>
<td>Int</td>
</tr>
</tbody>
</table>
3.4. Voronoi Diagrams’ Building Process

In this paper, the sample datasets are used as the central point of the Voronoi diagrams. This novel model constructs Voronoi polygons by the Delaunay triangulation method based on the grid growth algorithm [77,78]. The main algorithm is built with the C# programming language based on the ArcGIS Engine and Visual Studio 2010 platform. The experimental process is shown in Figure 2b. The novel model defined the following data types for experimental program development: TPointSetArry as the center point set array, TPointsetList as the point set list, TriSetList as the edge set list, nTriCount as the number of Delaunay triangles and TPointSetListD as the Delaunay triangulation point set list.

3.4.1. Calculating the Initial Centre Points and Edges

- **Step 1**: Although any point can be used as the initial point P1, to improve program efficiency, a point near the center of the study area was chosen as the initial point of the grid. The numbers of initial center points and edges are stored in TPointSetArry and TPointSetList, respectively.

- **Step 2**: The model comparison distance is defined by Equation (3).

- **Step 3**: The initial value of $D_{\text{min}}$ is defined as the larger value. In this paper, the diagonal length of the graphic in the entire study area is defined as the initial value of $D_{\text{min}}$.

- **Step 4**: The distance between two points is calculated by using Equation (9) if the grid cell of the initial point also contains other points. This distance is then compared to $D_{\text{min}}$ to identify the nearest point to the initial point $P_1$. Selecting a small value reassigns it to $D_{\text{min}}$.

- **Step 5**: The adjacent grid cells are searched until the shortest side of $D_{\text{min}}$ is found, and the search pattern moves from left to right and from top to bottom. Then, the shortest side is taken as the initial Delaunay edge. The number of initial Delaunay edges is stored in TriSetList.

3.4.2. Constructing the Delaunay Triangulation

- **Step 1**: Taking $P_iP_j$ as the current processing edge, the intersect calculation is then performed with the bottom line of the grid cells to obtain the value of $i_1$. The intersection grid column gives the value of $j_1$. Similarly, the intersect calculation is conducted with the right bottom line of the grid cells to obtain the value of $j_2$. The intersection grid column provides the value of $i_2$. Finally, the cells $(i_1,j_1)$ and $(i_2,j_2)$ are obtained at the right grid.

- **Step 2**: The unit triangle is formed by three vertices: $(i_1,j_1)$, $(i_1,j_2)$ and $(i_2,j_2)$. The triangle circumcircle is then constructed.

- **Step 3**: Repeat Step 1 in the area covered by the circumcircle. The point that creates the largest angle with the current edge $P_iP_j$ is selected as the vertex. Then, the triangle is constructed in the circle formed by the points by repeating the previous steps until all of the grid units covered by the circumcircle are searched and no other vertex is available. The resulting triangle is a Delaunay triangle.

- **Step 4**: The center point and Delaunay triangle are numbered. The numbers of the three central points (vertices) that comprise the Delaunay triangle are recorded. These data are stored in TPointSetListD, and the numerical value of nTriCount is incremented by 1.

3.4.3. Constructing the Voronoi Homogeneous Diagram Map

- **Step 1**: As reported previously [79,80], all Delaunay triangle numbers adjacent to each center point are recorded and sorted in the anticlockwise direction.

- **Step 2**: A convex hull boundary with the peripheral boundary connects each center point of the adjacent triangles in the circumcircle. The entire experimental data area is searched using the novel model, and then, the Voronoi homogeneous diagram map of the study area is produce.
3.5. Spatial Clustering and Autocorrelation Analysis Method

The authors perform a clustering analysis of the Voronoi polygonal neighborhood aggregation index of the study area using SPSS software [70,81]. The clustering analysis applies the shortest distance method and the correlation coefficient as the cluster sign. The mutation of the correlation coefficient is the end mark of the cluster, and the absolute distance threshold is defined as 0.09.

Spatial autocorrelation (also known as spatial dependence, spatial interaction or local interaction) can be loosely defined as a similarity (or dissimilarity) measure between two values of an attribute that are nearby spatially [54]. Spatial autocorrelation can be measured by various indexes, of which probably the most well-known is Moran’s I statistic. Statistics like the global Moran’s I are useful for analyzing datasets in a relatively homogeneous region; it may not be as informative to compute the Moran’s I value for data across a region that could have several spatial regimes [55,82]. One solution is to develop a set of Local Indicators of Spatial Association (LISA), such as the local Moran’s I [55,82]. Therefore, we calculate the local Moran’s I data in the study area by using the spatial autocorrelation analysis (univariate Moran) function of the GeoDA software [83,84]. The generalized form for the local Moran’s I can be defined as follows [55]:

\[
I_i(d) = Z_i \sum_{j \neq i} W_{ij} Z_j
\]

where the observations \( Z_i \) and \( Z_j \) are in standardized form (with a mean of zero and a variance of one). The spatial weight \( W_{ij} \) is in row-standardized form. Therefore, \( I_i \) is a product of \( Z_i \) and the average of the observations in the surrounding locations.

The Moran scatter plot visually illustrates the spatial transformation of a variable (y-axis) on the original standardized variable (x-axis). The slope of the scatter plot corresponds to the value for Moran’s I. According to the Moran scatter plot, we can find the extent of linear association between the values in a given location (x-axis) with the values of the same variable in neighboring locations (y-axis) [82,85].

4. Results and Discussion

4.1. Spatial Pattern Clustering Results

The clustering analysis results and frequency histogram are shown as Table 3 and Figure 3, respectively. In Figure 3, the horizontal coordinate is the value of each land aggregation index, and the vertical coordinate is the ratio of each land use frequency to the group distance.

<table>
<thead>
<tr>
<th>Level</th>
<th>Residential Land</th>
<th>Commercial Land</th>
<th>Industrial Land</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Index Value</td>
<td>Area (hm²)</td>
<td>Ratio (%)</td>
</tr>
<tr>
<td>1st level</td>
<td>≥0.83</td>
<td>16.58</td>
<td>12.52</td>
</tr>
<tr>
<td>2nd level</td>
<td>0.75~0.82</td>
<td>27.32</td>
<td>20.63</td>
</tr>
<tr>
<td>3rd level</td>
<td>0.49~0.74</td>
<td>47.60</td>
<td>35.94</td>
</tr>
<tr>
<td>4th level</td>
<td>0.33~0.48</td>
<td>40.94</td>
<td>30.91</td>
</tr>
</tbody>
</table>

Table 3. Statistical analysis of the Voronoi polygon aggregation index for each land use type.
We merge Voronoi polygons according to the cluster level, merge smaller polygons based on the principle of the longest shared boundary and then obtain the spatial agglomeration pattern Voronoi homogeneous map of the study area (Figure 4a).

**Figure 3.** The frequency histogram of the Voronoi polygon aggregation index.

**Figure 4.** The spatial clustering classification results of the study area. (a) The Voronoi polygons of the study area; (b) spatial clustering classification results of residential land use; (c) spatial clustering classification results of commercial land use; and (d) spatial clustering classification results of industrial land use.
Clustering results of residential land use: The concentrated clustering classification land use area (first level) accounts for 12.52% of the total area and is mainly located along Jiefang road, Hangkong road, Baofeng road and Huangpu road, where municipal establishments are complete and transportation is convenient. The sub-concentrated clustering classification land use area (second level) accounts for 20.63% of the total area and is primarily found along Fazhan road, where new contiguous residential areas, building density and building volume rates are higher. The generally concentrated clustering classification land use area (third level) accounts for 35.94% of the total area, is mainly located in Changfeng township and mostly consisted of urban peripheries, where the layout is not complete and the environment is general. The dispersed clustering classification land use area (fourth level) accounts for 30.91% of the total area and is primarily noted along Houhu road, where the building density and building floor area ratio is low, the building distribution is dispersive, the municipal establishments are not complete and the environment is poor.

Clustering results of commercial land use: The concentrated clustering classification land use area (first level) accounts for 13.94% of the total area and is mainly located in the street business district on both sides of Zhongshan avenue, where abundant, densely-packed financial, insurance and high-rise office buildings, luxury hotels and passenger flow are located. The sub-concentrated clustering classification land use area (second level) accounts for 27.92% of the total area and is typically located in the peripheral zone of Zhongshan avenue, hosting regional traffic, a high building density and construction volume and completed municipal facilities, which attract a large passenger flow. The generally concentrated clustering classification land use area (third level) accounts for 36.49% of the total area, primarily along Baofeng road, North lake road, Station road and Huangpu road. In these areas, the facilities are complete, and the layout of the environment is mixed use. The greatest potential of these regions is for mining because of the rapid economic development and the improvement of infrastructure. The dispersed clustering classification land use area (fourth level) accounts for 21.65% of the total area and is mainly comprised of the Hankou railway station and Huangpu Avenue East theatre. The commercial network density is relatively low, and the passenger flow is not significant, mainly corresponding to regional business services. These regions have the additional advantages of comprehensive exploitation and application.

Clustering results of industrial land use: The generally concentrated clustering classification land use area (third level) and the dispersed clustering classification land use area (fourth level) account for 81.61% of the total industrial land use area. The industrial lands are mainly located in the city’s peripheral area, such as Chanfeng road, where the Qiaokou economic and technological development zone is. Additionally, it is mainly located in north Hankou districts, containing the Panlong economic and technological development zone and the north Hankou metropolitan industrial park. The average efficiency of the industrial land is low because the government restricts industrial areas in cities to improve the urban environment and the economic, social and ecological benefits of urban land use. The area of industrial land declined under the policy of “industrial factory relocation from the second ring zone and replacement with green plants”.

Spatial autocorrelation results: The Moran scatter plot in Figure 5 illustrates the relationship between each spatial clustering patch. In which, the horizontal coordinate is the standard value of each unit’s intensive degree and the vertical coordinate is the average value of the properties of neighboring units. In the Moran scatter plot, the slope of the regression line corresponds to Moran’s I; the four quadrants of the coordinate system correspond to the four types of local space between the regional unit and its neighbors.
The slope of the regression line in the Moran scatter plot is 0.0537, which represents the residential land cover present spatial positive correlation (Figure 5a). The Moran scatter plots are mainly distributed in the first, second and third quadrant, belong to high-high, low–high and low–low spatial clustering. For the Moran scatter plot, the first quadrant of the coordinate system represents the spatial connectivity of the high observed value area unit surrounded by the high observed value region (high–high); the second quadrant of the coordinate system represents the spatial connectivity of the low observed value area units surrounded by the high observed value region (low–high); and the third quadrant of the coordinate system represents the spatial connectivity of the low observed value area unit surrounded by the low observed value region (low-low) [83]. The high–high clustering units in the first quadrant are few compared to the low-low and low-high clustering units in the second and third quadrants. Therefore, the aggregation advantage of residential land use in the study area has an accumulative effect, but the aggregation advantage is not significant and the spatial difference between the overall levels of stacking utilization not large. In Figure 4b, the high–high clustering areas of residential land cover are mainly located in Jiefang road, Hangkong road, Fazhan road, Laodong road, Yongqing road and Wansong road, which are adjacent to the commercial districts, with perfect facilities and convenient traffic; the spatial clustering classes are mainly including the first and second level. The low–high clustering areas are widely distributed in the study area, showing the spatial dispersion characteristics and the spatial clustering class at the third level. Low–low clustering areas are mainly located in urban fringes; the spatial clustering class is the fourth level.

The slope of the regression line in the Moran scatter plot is 0.1658, which represents that the commercial land cover aggregation presents considerable spatial autocorrelation (Figure 5b). The Moran’s I of commercial land is maximized. The maximum value of the gradient of the intensive degree value indicates that the spatial agglomeration of commercial land is considerable compared to those of other land uses and is subject to a spatial agglomeration effect. The Moran scatter plots are mainly distributed in the first and third quadrant, belong to high–high and low–low spatial clustering. The high–high clustering units in the first quadrant are large compared to the low–low clustering units in the third quadrant. Therefore, the commercial land use in the study area has high utilization efficiency and a high intensive utilization degree. In Figure 4c, the high–high clustering areas of commercial land cover are mainly located in the street business districts along Zhongshan Avenue on both sides, with extremely dense commercial networks and extremely high land use efficiency and intensive degree; the spatial clustering classes mainly include the first, second and third level. The low–low clustering areas are mainly located in Hankou railway station with low

![Figure 5. Moran scatter plot chart denoting the heterogeneity of local space. (a) Moran scatter plot for residential land cover spatial pattern agglomeration indexes; (b) Moran scatter plot for commercial land cover spatial pattern agglomeration indexes; (c) Moran scatter plot for industrial land cover spatial pattern agglomeration indexes.](image-url)
density commercial networks; the spatial clustering class is the fourth level. The spatial difference of commercial network density in other regions is obvious and shows dispersion characteristics.

The slope of the regression line in the Moran scatter plot is 0.1473, which represents that the industrial land cover presents a spatial positive correlation (Figure 5c). The Moran scatter plots are mainly distributed in the first and third quadrant, belonging to high–high and low–low spatial clustering. The high–high clustering units in the first quadrant are few compared to the low–low clustering units in the third quadrant. Therefore, the industrial land use in the study area has increased utilization efficiency. The Moran’s I of industrial land is compared to those of other types of land use. The degree of spatial agglomeration is greater than that of residential land and less than that of commercial land. Additionally, this value responds to the actual situation of industrial land, which is mainly distributed in several development zones. In Figure 4d, the high–high clustering areas of industrial land use are mainly located on Chanfeng road, which is the Qiaokou economic and technological development zone; the spatial clustering classes mainly include the first and second level. The low–low clustering areas are mainly locate in the north Hankou districts, which contain the Panlong economic and technological development zone; the spatial clustering class is the fourth level.

4.2. Discussion

An urban area is not just a single agglomeration of people in space, but rather consists of numerous large and small clusters of households, firms and infrastructures that exhibit spatial patchiness in their social and economic structures [86]. The spatial structure and temporal dynamics of urban areas are complex and controlled by various driving forces, such as the regional economy, population movements, the policy environment and socio-cultural processes [87].

According to neo-Marxist theory, conventional urban science explains the urban phenomena partially with the factors of technology and population migrations, which are considered as principal causes of urban structural change [88]. In a neo-Marxist analysis, however, these are not the underlying causes, but instead are intermediate factors produced by something more fundamental, such as the basic requirements and social relations of capitalist production. Neo-Marxists present the structure of urban structure and land use/cover as the result of capitalism in pursuit of profit [88–90]. “Since the process of capital accumulation unfolds in a spatially structured environment, urbanism may be viewed provisionally as the particular geographical form and spatial patterning of relationships taken by the process of capital accumulation” [54,90].

Through different prices of land rent, land tax theory determines the spatial distribution of land use. According to the bidding land tax curve model, the land tax of commercial, residential and industrial land use is negatively related to the distance from the city center. Under the equilibrium condition of land tax and distance, the commercial land should be arranged in the nearest place from the city center, then residential land, and the farthest is the industrial land. In a spatial pattern, the concentric circle layer structure is distributed around the city center [91]. In the study area, the price of land tax is decreasing from the city center to the periphery. In Figures 4c and 6a, a high density commercial district is mainly located in the first grade land tax area, which is the oldest and largest commercial district at present. Residential location patterns are characterized by significant spatial interrelationships across race, income, neighborhood quality, traffic accessibility and distance from a functional area (such as shopping malls, schools, hospitals, parks, etc.) [87]. No matter how large the scale of the functional area is, the spatial diffusion range is limited and close to the functional area to become the first choice for residential land. In Figures 4b and 6a, most residential land is located in the second and third grade land tax area, a contiguity commercial district. Industrial land use is located in the urban fringe, where land resources and labor resources are rich, and it provides convenience for industrial organization, layout and optimization. In sum, the influence of land tax theory on the urban spatial structure of the study area is very obvious, and the urban spatial structure is characterized by a concentric layer structure from the urban center to the urban fringe. The urban spatial structure will transfer from a single center structure to a multi-center decentralized structure.
The government’s macro control policies have affected the urban spatial structure. In the study area, “industry quits from the second ring zone and enters into the third ring zone” is the urban management policy to adjust the industrial structure. The formation of the Qiaokou economic and technological development zone, which is located in the urban fringe, stimulated the construction of the surrounding facilities. With the gradual improvement of living facilities, commercial facilities and the transportation network, the surrounding real estate industry has been stimulated. The new center of the city has been developing and growing, which will eventually develop into a sub-center of the city.

Though land tax is believed to be the primary cause for shaping the urban spatial structure in the study area, other causes, such as population growth, economic investment and development, need to be examined. In Figure 6b, the permanent residents increase slowly with an annual average growth under 1%. The reason is that the study area was an old urban district, and the number of permanent residents tends to be saturated. However, in China, the difference of the urban and rural economic structure causes the income difference between urban and rural areas, and a large number of floating population migrates into cities, as commerce cluster areas can provide them considerable employment opportunities. In response, the floating population promotes a further agglomeration in this area.

At the same time, with an increase of the income, the level of residents’ consumption and demand improving gradually, the demand on the living environment, scale and level of commercial facilities continues to improve. Since the 1990s, China has implemented a series of land system reform policies including “an auction way to implement the transfer of land-use right” and “abolishing the welfare housing system”. Along with the housing commercialization carried out in China, more and more urban residents begin to escape from the “unit community” and purchase marketable housing based on their own abilities and preferences. The real estate industry has made outstanding contributions to the urban economy development, and the growing investment in real estate development has driven the synchronous development of the urban spatial structure agglomeration at the same time. For example, in 2009, Hankou’s investment in real estate development was 23.05 billion RMB with an increase rate of 59.44% over 2008, and the annual per capita disposable income of urban residents was 18,663 RMB with an increase rate of 15.60% over 2008. For the corresponding period, the area of commercialized housing sold was 2.84 million m² with an increase rate of 71.83% over 2008. The correlation between urban structure aggregation and disposable income of urban residents/investment on real estate development was examined, and it reveals a very strong association (Figure 6b).

5. Conclusions and Future Work

Wuhan is now at a stage of rapid urbanization, and urban studies are needed for a better understanding of the urban growth mechanism, to support land use planning for sustainable urban
land use and protection. The spatial pattern of urban land use is affected by a series of factors, including both spatial factors (spatial relationships) and non-spatial factors (attribute relationships), such as political and economic factors. This research focuses on how to use modern theory and technology to describe these factors reasonably, to obtain results that reflect reality and to avoid separating theory from practice.

In this paper, a spatial agglomeration analysis of urban land use based on spatial adjacency and similar attribute characteristics is performed. To this end, the spatial and attribute characteristics are integrated into the generalized Euclidean distance to extend the traditional Voronoi diagram. This novel model constructs a Voronoi homogeneous map based on sample point datasets and the spatial attribute generalized Euclidean comparison distance. The spatial neighborhood aggregation indexes of the Voronoi polygons in the study area are calculated based on landscape ecology theory. The model generated the spatial Voronoi homogeneous map based on the spatial clustering grade of the neighborhood clustering indexes using SPSS software. In addition, the spatial autocorrelation of the urban spatial structure in the study area is analyzed by calculating the local Moran’s I using the GeoDA software.

Although population growth triggered by rural-urban migration generally was believed to be the dominant cause for urban spatial structure changes [7], according to the neo-Marxist theory, we briefly analyze the driving forces for shaping the urban spatial structure, and the analysis indicates that the different price of land rent, the government’s macro control policies and investment in real estate development are the major factors for the rapid urban spatial structure changes in the study area. The present study is expected to have significant implications in rapidly urbanizing cities of this developing country.

This research is only one example of using extended Voronoi diagrams to study the urban spatial pattern, and as such, this is a prototype. There is still much potential for developing this research further. Future research will focus on using the self-organizing theory of neural networks to cluster the classifications of attribute information, using a gravity model to improve the quantitative expression of physical space and attribute correlation [92]. The four forks tree method will be used to study the spatial data structure, and the efficiency of spatial data insertion and queries for improving the triangular surface construction in Voronoi diagrams will be evaluated. In addition, the self-organizing maps (SOM) may have a profound and significant role in merging decision makers based on multi-dimensional spatial data analysis [92,93]. The complexity of human movements has redefined the usage of urban space and the arrangement of resources; changes in the lifestyles of urban residents have prominent and visible effects on urban spatial metrics [92]. Most of these changes have been brought by human impact on the environment and excessive exploitation of resources. In this sense, the urban spatial patterning research approach should be “intelligent” and yield important insights into urban phenomena generated by human movements (social, natural and economic) [9].

**Acknowledgments:** This study was funded by the National Natural Science Foundation of China (41071242) and the Educational Commission of Hubei Province of China (D20131104). We also express our gratitude to the Wuhan municipal bureau of statistics for providing the statistical data and Qiaozhi Wang for proofreading part of the sections. Additionally, we appreciate the anonymous reviewers for their constructive comments, which helped to substantially improve our work.

**Author Contributions:** This research was mainly performed and prepared by Zuohua Miao, Yong Chen and Xiangyang Zen. Zuohua Miao and Yong Chen contributed ideas and conceived of and designed the experiments. Zuohua Miao and Yong Chen wrote the paper, and Xiangyang Zeng reviewed and edited the manuscript. Yong Chen supervised the study, and his feedback was considered throughout the paper. Xiangyang Zeng and Yun Li analyzed the experimental data.

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


9. Vaz, E. The future of landscapes and habitats: The regional science contribution to the understanding of geographical space. Habitat Int. 2016, 51, 70–78. [CrossRef]


16. Ma, Y.; Xu, R. Remote sensing monitoring and driving force analysis of urban expansion in guangzhou city, china. Habitat Int. 2010, 34, 228–235. [CrossRef]

17. Hsu, W.-T. Central place theory and city size distribution*. Econ. J. 2012, 122, 903–932. [CrossRef]


38. Cromley, G.R.; Hanink, M.D. Coupling land use allocation models with raster gis. J. Geogr. Syst. 1999, 1, 137–153. [CrossRef]


42. Wang, F.; Marceau, D.J. A patch-based cellular automaton for simulating land-use changes at fine spatial resolution. Trans. GIS 2013, 17, 828–846. [CrossRef]


63. Sharifzadeh, M.; Shahabi, C. Vor-tree: R-trees with voronoi diagrams for efficient processing of spatial nearest neighbor queries. *Proc. VLDB Endow.* 2010, 3, 1231–1242. [CrossRef]


