

## Article

# Modeling the Interactive Patterns of International Migration Network through a Reverse Gravity Approach

Xuesong Yu , Kun Qin <sup>\*</sup>, Tao Jia <sup>\*</sup> , Yang Zhou  and Xieqing Gao

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China; yxs95@whu.edu.cn (X.Y.); zhouyang0612@whu.edu.cn (Y.Z.)

<sup>\*</sup> Correspondence: qink@whu.edu.cn (K.Q.); tao.jia@whu.edu.cn (T.J.)

**Abstract:** The international migration network, comprising the movements of people between countries, is one of the most important global systems of interaction, which can reflect the complex international relations of economics, cultures, and politics and has huge impacts on global sustainability. However, the conventional gravity model cannot model its complicated interactions accurately. In this article, we propose a novel reverse gravity model using genetic algorithm to reconstruct the complicated interaction patterns with high accuracy. To verify the feasibility of our method, it was applied to a series of international migration networks. We found that the derived node attractions were highly correlated with socioeconomic factors and network metrics, and the calculated node positions outperformed the geometric centers from the perspective of human migration that related to economy and demography. Our approach could be a preferred choice to investigate the spatial-temporal interactive patterns in geographical space, facilitating comprehension of the mechanisms underlying their generation and evolution.

**Keywords:** reverse gravity model; international migration; human geography; genetic algorithm; spatial interaction networks



**Citation:** Yu, X.; Qin, K.; Jia, T.; Zhou, Y.; Gao, X. Modeling the Interactive Patterns of International Migration Network through a Reverse Gravity Approach. *Sustainability* **2024**, *16*, 2502. <https://doi.org/10.3390/su16062502>

Academic Editor: Giuseppe T. Cirella

Received: 2 January 2024

Revised: 29 February 2024

Accepted: 11 March 2024

Published: 18 March 2024



**Copyright:** © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

International migration is a significant matter of interest, not only in the study of demography but also in the domains of human geography, human mobility, and spatial analysis [1,2]. It has a complex impact on global sustainability [3,4]. Constructing international migration networks is an often-used approach [5,6]; the network structure is a good representation of migration flows, and the complex network theory offers a rich diversity of methodologies to examine the characteristics of population interactions. By analyzing the small-world and scale-free properties of the network, as well as the centrality of nodes and community division, researchers can evaluate the structural features of international migration [7–10].

The international migration networks, apart from the topological properties, are also distinguished by their spatial interaction features and are typical instances of spatial interaction networks. On the one hand, by taking spatial factors into account, the community structure and node centrality of the network are different from general complex networks [11,12]. On the other hand, the simulating and modeling of spatial interactions is also a research focus which could help to yield a better understanding of the spatial-temporal patterns within [13,14].

There are many models that migration networks could be fitted into [15], among which the most famous are the gravity model [16], radiation model [17], and intervening opportunity model [18]. The gravity model draws inspiration from Newton's law of gravitation, in that the strength of the interaction is in proportion to the mass of both sides and the inverse of the squared distance between them; it was originally introduced to study the correlations among population, distance, and intercity migrations [19]. Now this

law has been empirically found to exist in various systems, including international trade networks, communication networks, social tagging networks, and others [20–22]. From the perspective of modeling, gravity is distinguished by its easily representable nature of interaction [23]. There is a rich body of academic literature in this field, focusing on how to implement this framework and get a better estimation [24]. The radiation model and intervening opportunity model, drawing on the radiation and absorption process in solid state physics, are also widely used and have proven to be effective in predicting human mobility [25,26]. However, these two models are mainly used in urban and interurban migration [27], while the gravity model is a more preferable choice on the international scale. Despite their effectiveness, these three models still face challenges in finding the appropriate instrumental variables [16]: a poor choice in instrumental variables may lead to bad-fitting results. Another research gap for them is the calculation of distance [28]: the suitability of distance representation between two irregular geographical units also affects the modeling effect. Although the Euclidean distance between the centroids of geographical units is the usual practice, the justification of centroids remains uncertain.

To tackle the first shortcoming, the reverse gravity model, or a reverse-fitting method of the gravity model, is applied [29]. This approach attempts to reconstruct the theoretical attractions of nodes from known spatial interactions and distances; it has also developed many different algorithms [30]. This method is an easily conducted data-driven approach aiming to obtain the best fit and is useful for interaction networks where the relations are complex and factors are hard to identify. However, the proper representation of distance remains a difficult task.

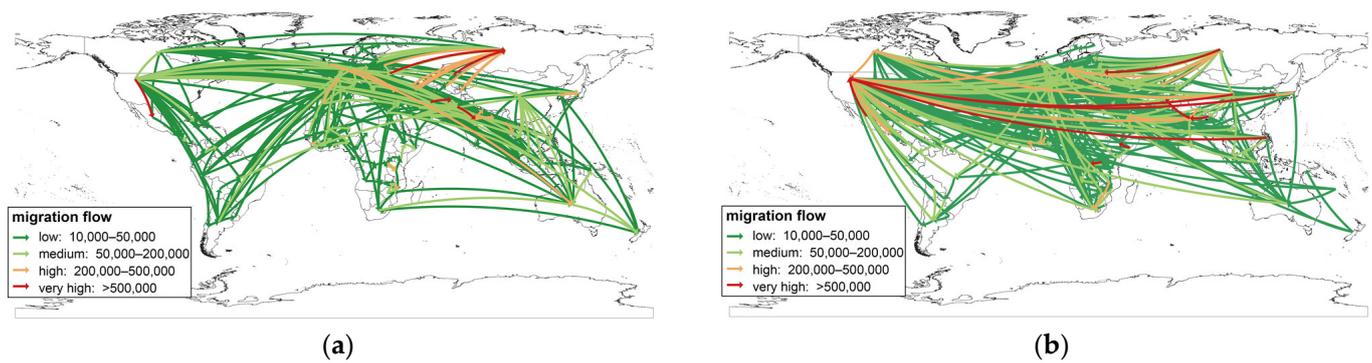
In this article, we proposed a modified reverse gravity model based on genetic algorithm, where not only bidirectional attractions but also a practical position for each node were identified, from which the distances could be calculated. This method was applied to a series of international migration networks: the results revealed the latent spatial-temporal interactive patterns and were compared to other socioeconomic observations. In Section 2, the dataset and our proposed method are elucidated. In Section 3, the results on network properties and reverse gravity fitting are presented, including patterns of node attractions and positions. Sections 4 and 5 present the discussion and conclusions of our study.

## 2. Materials and Methods

### 2.1. Data

The international migration networks we used in this article were constructed from a set of five-year estimations of international migration flows, ranging from mid-year 1990 to mid-year 2020; the bilateral flow data were estimated based on migration stock data using a closed demographic accounting method with a pseudo-Bayesian approach [31]. This representative dataset has been widely used in recent studies, where the international migration is defined based on changes in place of residence [32,33]. For example, if the intensity of the flow from Country A to Country B in one period is 100, then there are 100 people who were first living in Country A but at the end were living in Country B. To match up with other data, we chose 194 countries. We constructed six directed weighted networks where nodes represented countries and edges indicated the migration flow from the source country to the target country, and the weights of edges were equal to the amount of people. These networks represented international migration in six periods: 1990–1995, 1995–2000, 2000–2005, 2005–2010, 2010–2015, and 2015–2020. All networks were strongly connected. The visualization and basic statistics of these networks are shown in Figure 1 and Table 1.

The administrative areas of countries were primarily obtained from the GADM project (<https://gadm.org/data.html>, accessed on 10 March 2024), while the administrative area of China was acquired from the Resource and Environment Science and Data Center (<https://doi.org/10.12078/2023010103>, accessed on 10 March 2024). These data were used to calculate the geometric center.



**Figure 1.** Example of the international migration network, with (a,b) displayed flows from west to east and east to west presented separately in the same network. The geometric center of the country is used as the position and the arrow indicates the direction of the edge. The color of the arrow represents the weight. Edges with very low weight (<10,000) are discarded for better visualization.

**Table 1.** Basic statistics of international migration networks.

Period	Number of Nodes	Number of Edges	Sum of Weights
1990–1995	194	24,937	66,160,782
1995–2000	194	25,372	66,025,147
2000–2005	194	25,729	71,575,382
2005–2010	194	26,088	83,411,284
2010–2015	194	26,371	90,879,826
2015–2020	194	26,598	93,155,493

The gross domestic product (GDP) and population data of all countries were downloaded from the World Bank. The inward and outward flow of foreign direct investment (FDI) were downloaded from the United Nations Conference on Trade and Development. These data could reflect the socioeconomic status of a country and could be used as the determining factor in international migration [34,35].

## 2.2. Complex Network Theory

### 2.2.1. Scale-Free and Small-World Properties

For complex networks, there are two main properties of concern: scale-free and small-world. The scale-free network is the one with a power-law degree distribution [36]:

$$Pr(X) = kX^{-\alpha}, \quad (1)$$

where  $X$  is the degree (or weighted degree),  $\alpha$  is the scaling parameter, and  $k$  is a normalization constant. This distribution means that a large fraction of the links (or weights) fall to a small fraction of nodes, resulting in heterogeneity of the network [37]. A fitting technique designed for empirical data is always used [38].

The small-world effect reflects the fact that the distances between pairs of nodes are short. Mathematically speaking, if a network has a greater average clustering coefficient as well as an equal or smaller average shortest path length compared to random networks of the same size, it is a small-world network. The small-world-ness can be quantitatively measured [39]. In a directed weighted network, the clustering coefficient is [40]

$$cc_i = \frac{\sum_j \sum_h (w_{ij}^{1/3} + w_{ji}^{1/3}) (w_{ih}^{1/3} + w_{hi}^{1/3}) (w_{jh}^{1/3} + w_{hj}^{1/3})}{2[s_i^{tot}(s_i^{tot} - 1) - 2s_i^{\leftrightarrow}]}, \quad (2)$$

where  $w_{ij}$  denotes the weights of the edge from  $i$  to  $j$ ,  $s_i^{tot}$  is the sum of in-strength and out-strength of node  $i$ , and  $s_i^{\leftrightarrow}$  is the reciprocal strength of  $i$ , which is calculated as

$s_i^{\leftrightarrow} = \sum_{j \neq i} w_{ij} w_{ji}$ . This equation captures the fraction of all possible geometric means of subgraph edge weights.

### 2.2.2. Betweenness Centrality

There are many centrality measures to evaluate the importance of a node [41]; of them, the betweenness centrality has been proven to have a vital role in community detection [42], finding influential nodes [43], and other tasks. Betweenness centrality is calculated as the sum of the fractions of shortest paths that go through a given node [44]:

$$bc_i = \sum_{s,t} \frac{\sigma_{s,t|i}}{\sigma_{s,t}}, \quad (3)$$

where  $\sigma_{s,t}$  is the number of shortest paths from node  $s$  to  $t$ , and  $\sigma_{s,t|i}$  is the number of shortest paths from  $s$  to  $t$  that go through  $i$ .

### 2.2.3. PageRank

PageRank was initially designed to rank web pages based on links between them; a web page is more important if it can be redirected from other important web pages [45]. This idea can be represented by the iterative formula below:

$$pr_i = \delta * \sum_{j \in nei_i} \frac{pr_j}{d_j} + \frac{1 - \delta}{N}, \quad (4)$$

where  $nei_i$  is the set of neighbors of node  $i$ ,  $d_j$  is the degree of node  $j$ ,  $N$  is the number of nodes, and  $\delta$  is the dangling factor. For weighted networks, the scaling factor  $\frac{1}{d_j}$  should be  $\frac{w_{ji}}{\sum_h w_{jh}}$ . The iterative process can be seen as a Markov process, and the PageRank vector is then defined as the stationary density of a discrete-time random walk on the network, leading to an algebraic approach to the problem [46].

## 2.3. Reverse Gravity Model Based on Genetic Algorithm

### 2.3.1. The Gravity Model

There exist numerous variations of the gravity model; in this article, a basic form was used. The gravity model considers differences between origins and destinations, and the distance decay parameter  $\beta$  is not fixed [19,23]. The model can be written as

$$G_{ij} = k P_i^{\alpha_{out}} P_j^{\alpha_{in}} d_{ij}^{-\beta}, \quad (5)$$

where  $G_{ij}$  is the geographical flow,  $P_i^{\alpha_{out}}$  and  $P_j^{\alpha_{in}}$  are the theoretical importances for the origin giving out flows and the destination taking in flows, and  $d_{ij}$  is the distance between them. The two factors are referred to as node attractions and are always denoted as  $P_i^{\alpha_{out}}$  and  $P_i^{\alpha_{in}}$ , indicating the same variable with different exponents.

When applying the gravity model to analyze international phenomena such as migration and trade, where  $i$  and  $j$  stand for countries, the variables typically encompass population size, GDP, or GDP-per-capita [47,48]. Additionally, distances are always calculated between geometric centers or capital cities [28,49]. Using the logarithm form, Equation (5) then becomes linear:

$$\ln(G_{ij}) = \ln(k) + \ln(P_i^{\alpha_{out}}) + \ln(P_j^{\alpha_{in}}) - \beta \ln(d_{ij}) = \ln(k) + \alpha_{out} \ln(P_i) + \alpha_{in} \ln(P_j) - \beta \ln(d_{ij}), \quad (6)$$

then ordinary least squares (OLS) and other linear regression methods can be used to solve the parameters  $\alpha_{out}$ ,  $\alpha_{in}$ , and  $\beta$ .

In many situations, the positive impact of GDP, population, or other variables on the number of flows is likely to be found, but the model may not be well-fitted due to a complex causal relationship. An alternative approach is to incorporate other factors into the equation, such as the Global Innovation Index, the Environment Index, average life expectancy, and dummy variables indicating whether the two countries are contiguous and share a common official language [28,50].

### 2.3.2. The Reverse Gravity Model

The reverse gravity model refers to the same equation as mentioned above but applies it in reverse to derive node attractions from the observed  $G_{ij}$  and  $d_{ij}$ . This approach simplifies the task of identifying and combining multiple related factors into the calculation of a single factor. Methods like linear programming [51], algebraic calculation [29,52], and OLS regression based on dummy variables [53] have been considered. More recently, a particle swarm optimization (PSO) approach was proposed and was found to achieve a better reconstruction of node attractions [30].

In this article, we proposed a new reverse gravity model using genetic algorithms (GA). GA, like PSO, is a heuristic algorithm aiming to find the best solution through an iterative optimization process and has been widely applied in research and engineering [54]. This algorithm is inspired by the genetic process in nature [55]. There are some terms in this algorithm. A solution or individual is a set of values, known as genes, which can yield a certain fitness. The population is a group of individuals. This algorithm operates as the population evolves into new generations. During the process, individuals are mating and mutating, which means genes are exchanging and changing randomly, leading to different solutions and fitness, and the favorable ones will be kept. Finally, after enough generations of searching, it will find the best solution.

To reverse fit the gravity model is to search for the best logarithm of  $P_i^{out}$  and  $P_i^{in}$  for every node, as well as the parameter  $\beta$ . The goodness of fitness can be represented by the adjusted  $R^2$  between actual and simulated flows. Moreover, we have introduced a technique to overcome the limitations of distance representation. By employing a similar genetic algorithm approach, the positions of nodes are found, and consequently, the distances can be calculated.

In summary, we proposed a reverse gravity approach with two stages of an alternate renewal process, where the two stages represent searching for node attractions and positions, respectively, and GA is used in each stage. The flow chart of our method is shown in Figure 2. (1) For a spatial interaction network of  $N$  geographic units, the first step is to calculate the pairwise distances of all nodes with the geometric center. (2) Next, we prepare a population with  $2N + 1$  genes for each individual, where the genes represent bidirectional node attractions and the distance decay parameter  $\beta$ , and we run GA to obtain the best solution. (3) We prepare a population of identical individuals. Each has  $2N$  genes with the values of node positions, and we run GA given the previous best node attractions and parameter  $\beta$ . (4) We run another GA to search for the best node attractions and parameter  $\beta$ , based on the previous best positions of nodes. (5) We repeat Steps 3 and 4 until the goodness of fitness converges, and the attractions and positions of nodes and the distance decay parameter are obtained. In this article, the geographic units are countries.

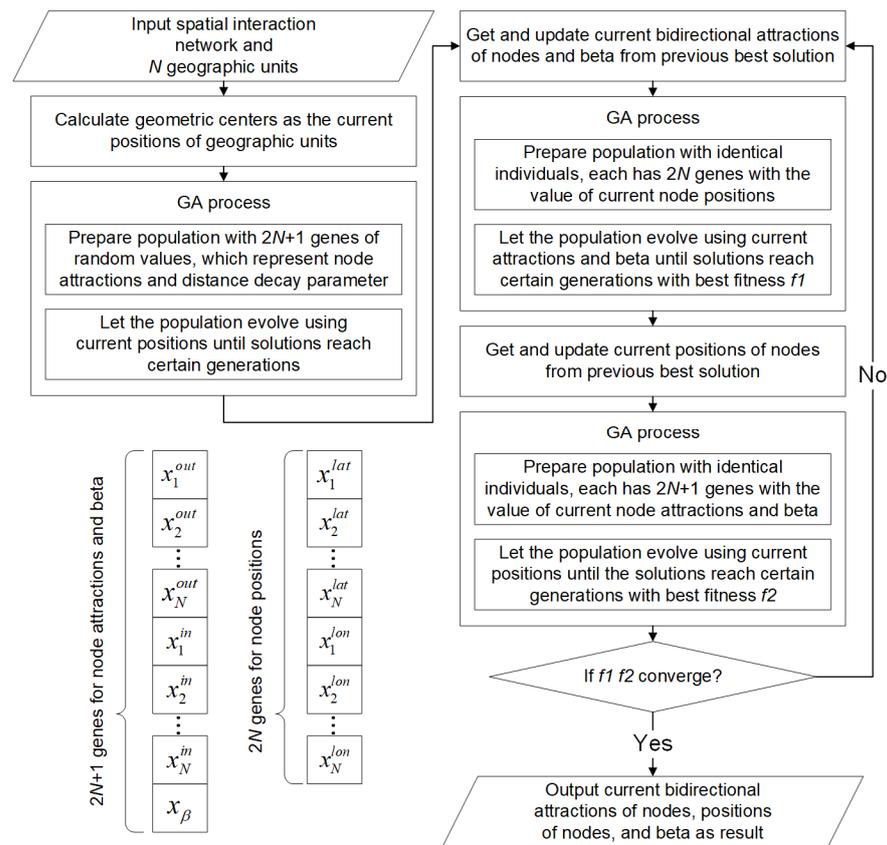


Figure 2. Flow chart of the reverse gravity approach with two stages of alternate renewal process.

### 3. Results

#### 3.1. Structural Properties of the International Migration Networks

##### 3.1.1. Scale-Free

We noticed that the weighted degrees of the United States, Russia, India, Germany, and the United Kingdom added up to 25.5%, 28.0%, 26.2%, 25.3%, 23.1%, and 24.8% of the total weights for six periods respectively; this may suggest a scale-free property. To investigate this, weighted degrees along with weighted in-degrees and out-degrees were fitted. In Figure 3, distributions of all periods are exhibited. We found that all distributions fitted well with the power law, which demonstrates obvious scale-free properties in international migration networks. The result indicates that a large amount of migration was related to only a small number of countries. The heterogeneity of weighted degree distribution declined from 1990 to 2010 as the scaling parameter decreased.

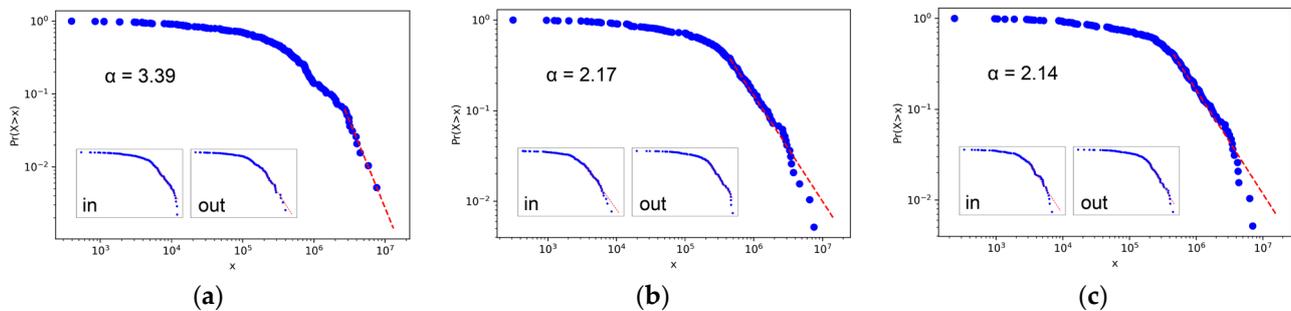
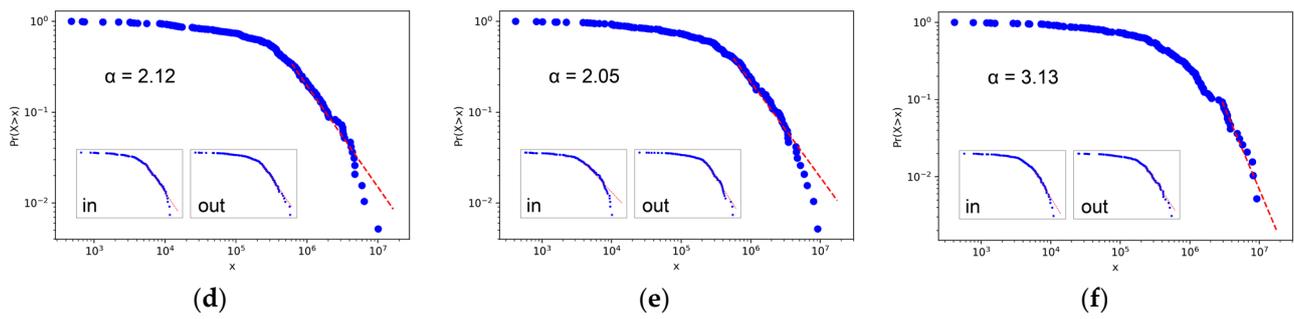


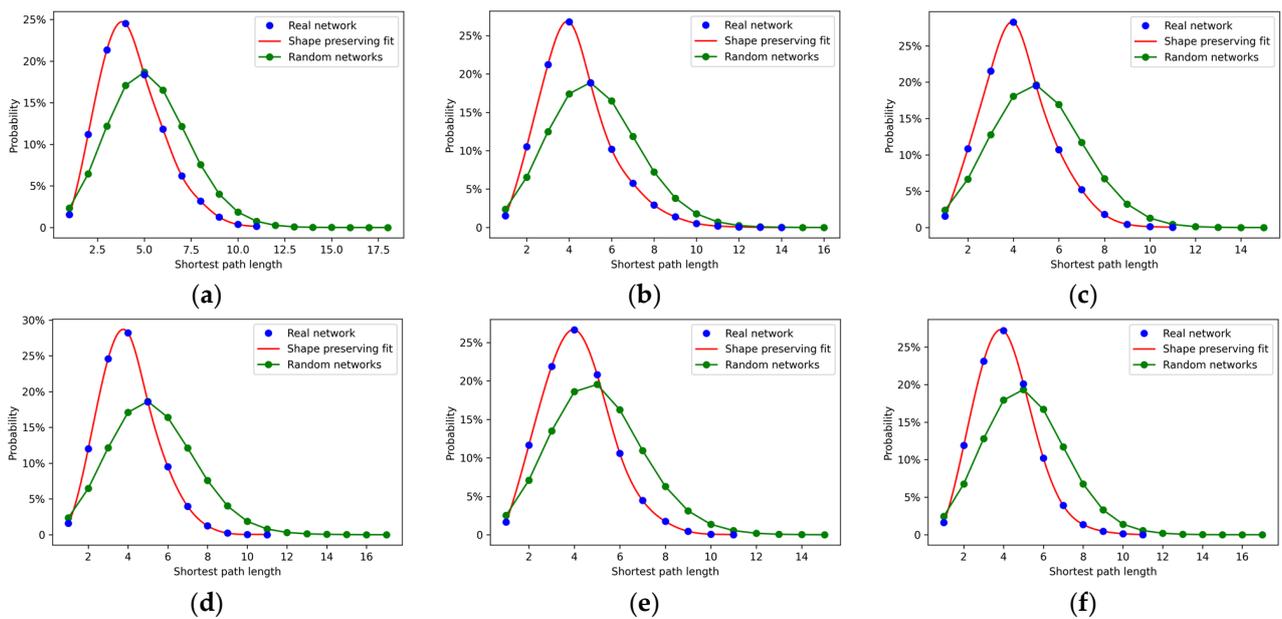
Figure 3. Cont.



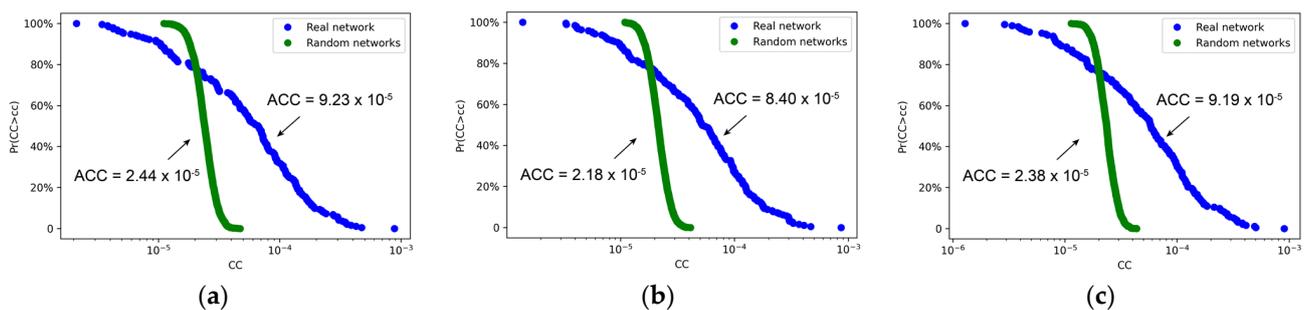
**Figure 3.** The weighted degree distribution (along with the in-degree and out-degree) of international migration networks can be well-fitted by a power-law distribution, indicating scale-free properties in these networks; (a–f) refer to 6 periods.

### 3.1.2. Small-World

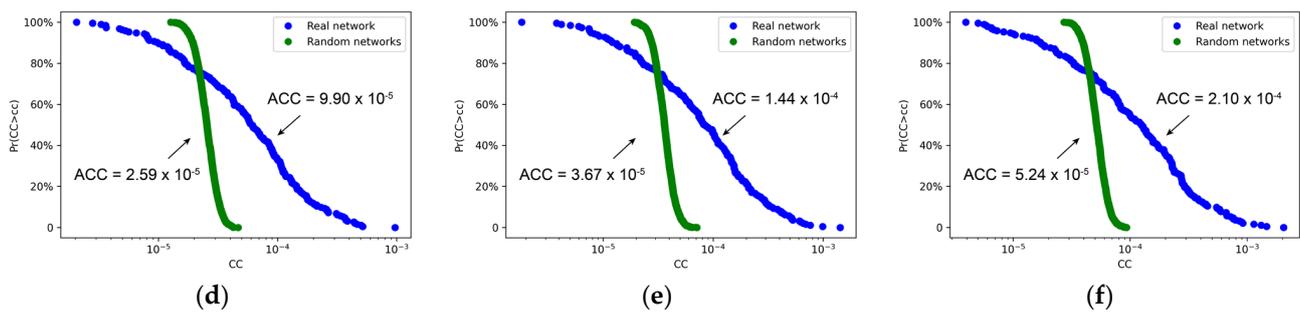
Migration networks should be compared to random networks of the same scale to investigate the small-world property. Here we generated 100 random networks for each period, with not only sums of edge weights but also distributions of edge weights kept the same as the real networks. Then, the distributions of shortest path lengths and clustering coefficients for both real networks and the averages of random networks were measured. Results are shown in Figures 4 and 5.



**Figure 4.** The shortest path length of real international migration networks (blue marker) and the averages of random networks (green marker) of the same scale; (a–f) refer to 6 periods.



**Figure 5.** Cont.



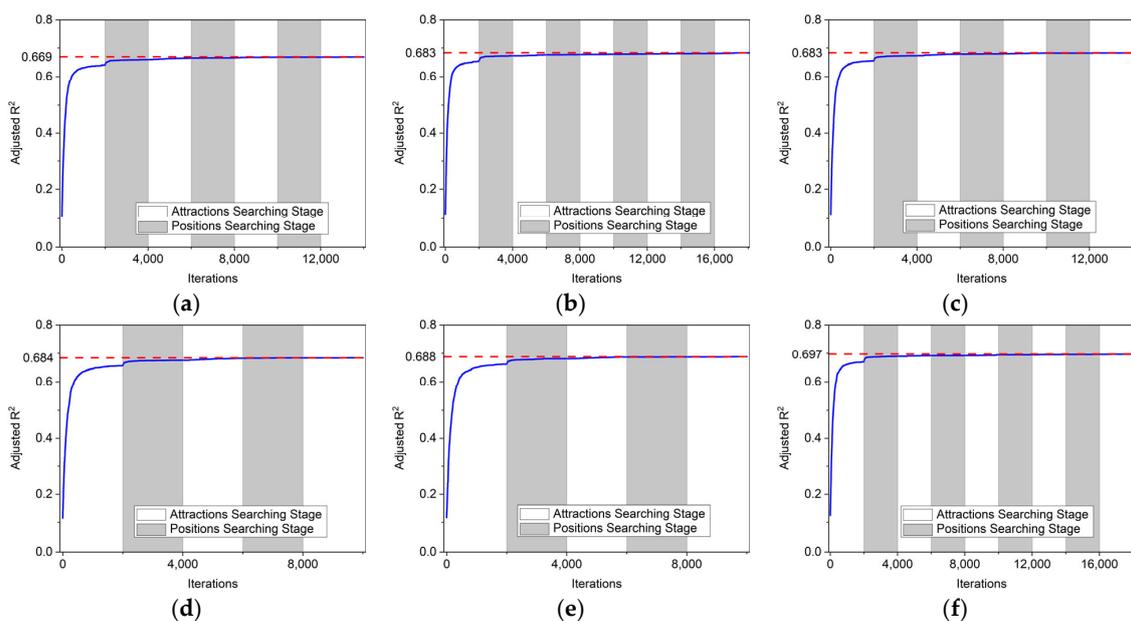
**Figure 5.** The weighted clustering coefficient of real international migration networks (blue marker) and the averages of random networks (green marker) of the same scale; (a–f) refer to 6 periods.

The shortest path lengths of real networks were slightly smaller than those of random networks, while the clustering coefficients were bigger. Using the quantitative measure, the small-world-ness values for each period were 4.58, 4.64, 4.71, 5.01, 4.78, and 5.01. This result showed that the international migration networks had a small-world effect. It should be noted that the reciprocal of edge weight was used for calculating the shortest path length. The existence of both scale-free and small-world properties in international migration networks suggested a complex interactive pattern and emphasized the importance of our subsequent research.

### 3.2. Interactive Patterns of the International Migration Networks

#### 3.2.1. The General Results of Our Method

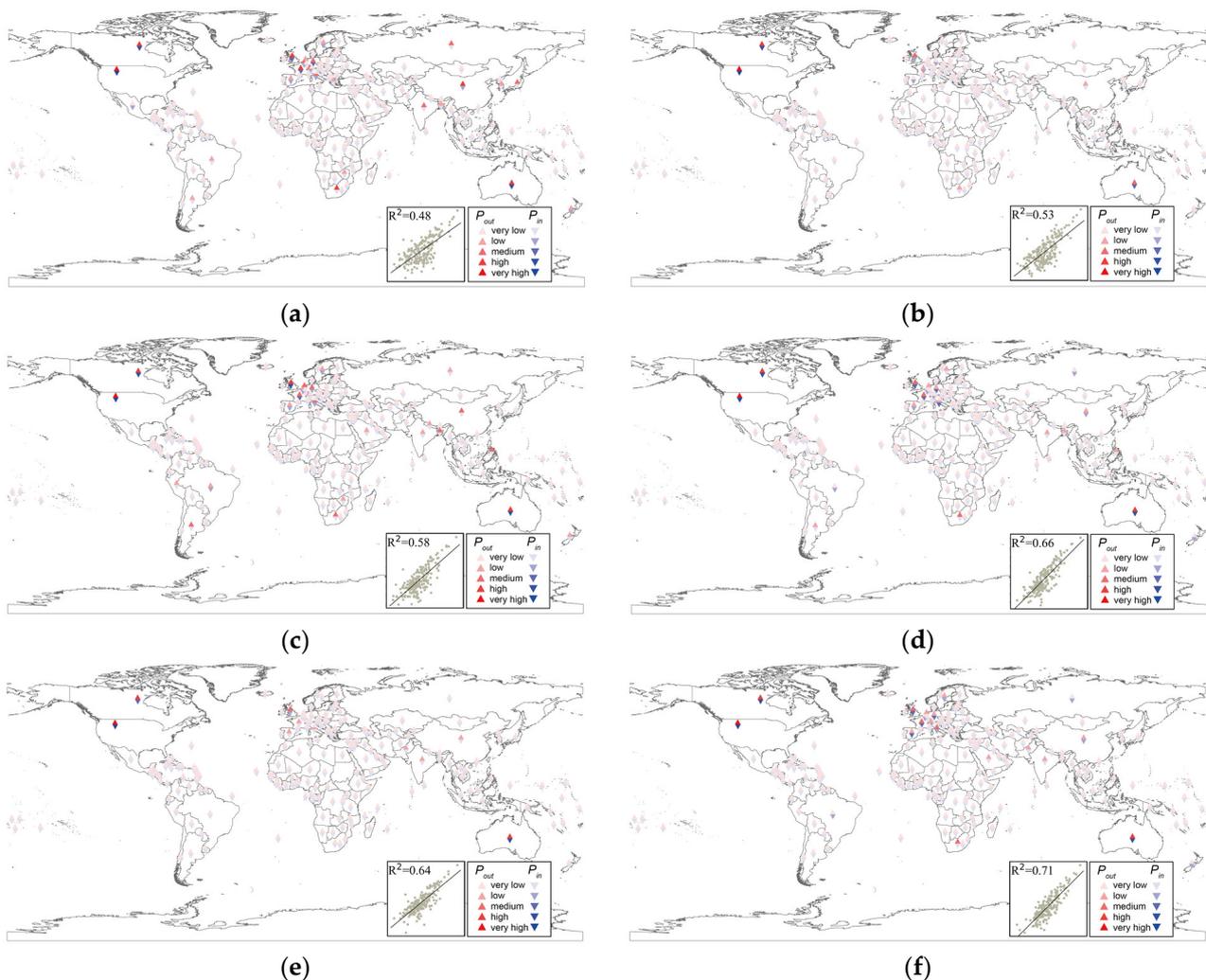
For six international migration networks, our method was used to find the node attractions and positions, as well as the distance decay parameter. For each searching stage in our method, the maximum number of iterations was set to 2000, and the population size was 50, in which 20 individuals were selected for mating, and 10% of genes would mutate. The goodness of fitness curves in Figure 6 demonstrate this procedure. We found that the optimal solutions were likely to be roughly identified in the first two stages and to converge in the subsequent stages. The final adjusted  $R^2$  reached 0.67–0.69, which suggested that the node attractions and positions our method found could interpret a significant portion of the observed data.



**Figure 6.** The goodness of fitness curves: different background colors represent different searching stages, the dashed red line denotes the convergence value, and (a–f) refer to 6 periods.

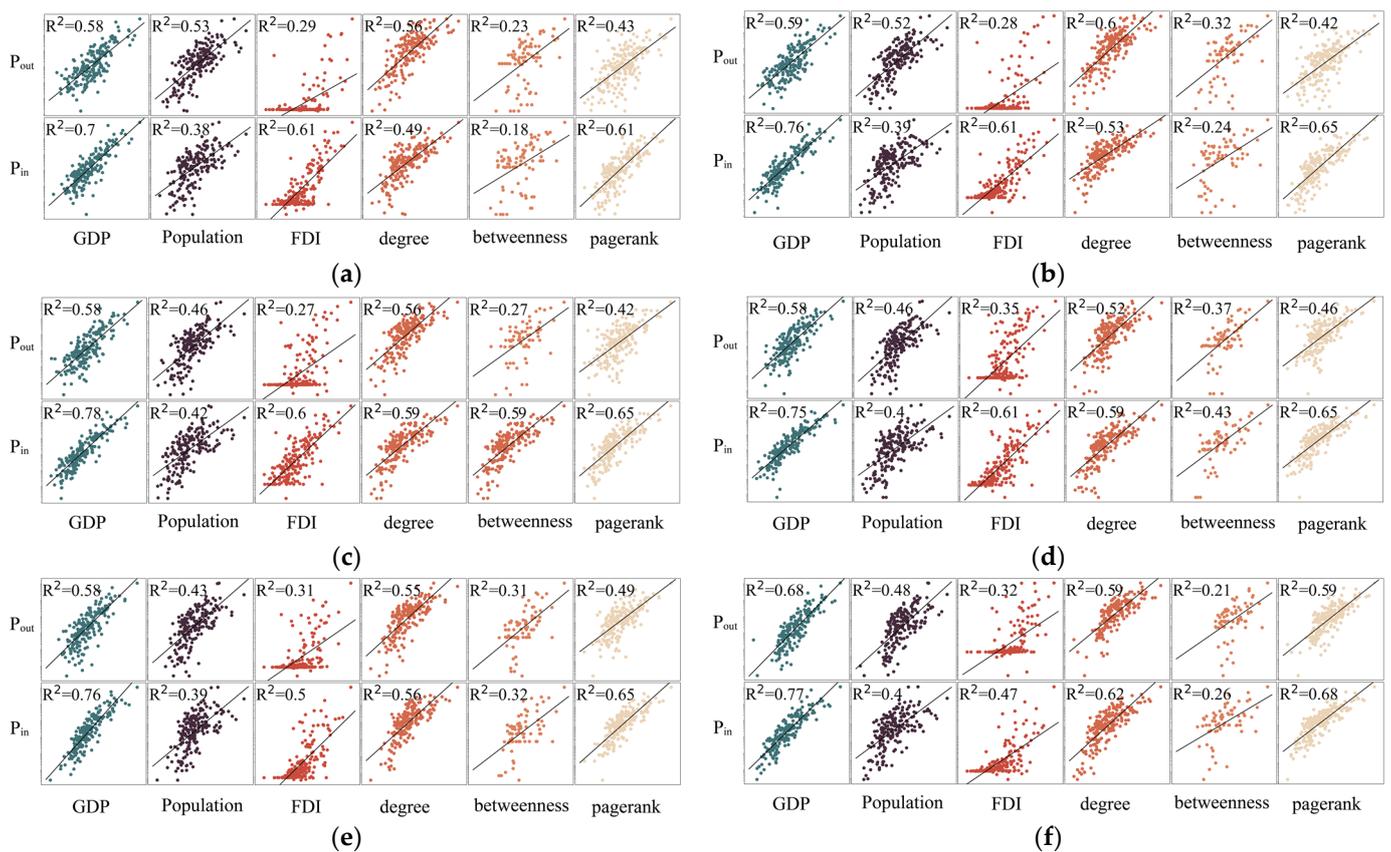
### 3.2.2. Node Attraction Patterns

The derived node attractions, which indicate the ability of pushing and pulling migration flows, are shown in Figure 7 (full results in Table S1). In general,  $P_{out}$  and  $P_{in}$  were correlated, which consisted in the assumption of using the same variable in the classic gravity model [19,56]. From the results, we could find some other patterns, which suggested different roles for different countries in spatial interactions. (1) Countries like the United States, Canada, Australia, and the United Kingdom ranked high in both attractions, indicating their activeness in international migration at all times. (2) Developing countries like China, India, and South Africa in most cases had a relatively stronger outward attraction. (3) Developed countries such as Germany and France had a relatively stronger inward attraction in the periods of 1995–2000, 2005–2010, and 2015–2020.



**Figure 7.** Geographical visualization of bidirectional node attractions. Triangles pointing up and down indicate outward and inward attraction, respectively, and different colors demonstrate different levels of attraction. Correlation plot of  $P_{out}$  and  $P_{in}$  in log–log coordinates are added. (a–f) refer to 6 periods.

Some socioeconomic variables and complex network metrics were used to evaluate the feasibility of the attractions. The correlation matrices are shown in Figure 8.



**Figure 8.** Correlation matrices of node attractions with socioeconomic variables and complex network metrics.  $P_{out}$  and  $P_{in}$  were analyzed using the GDP, population size, FDI, degree, betweenness, and PageRank score, separately. It should be noted that FDI and degree were both directional and were only analyzed with corresponding directed node attraction. All scatter plots are in log–log coordinates, and (a–f) refer to 6 periods.

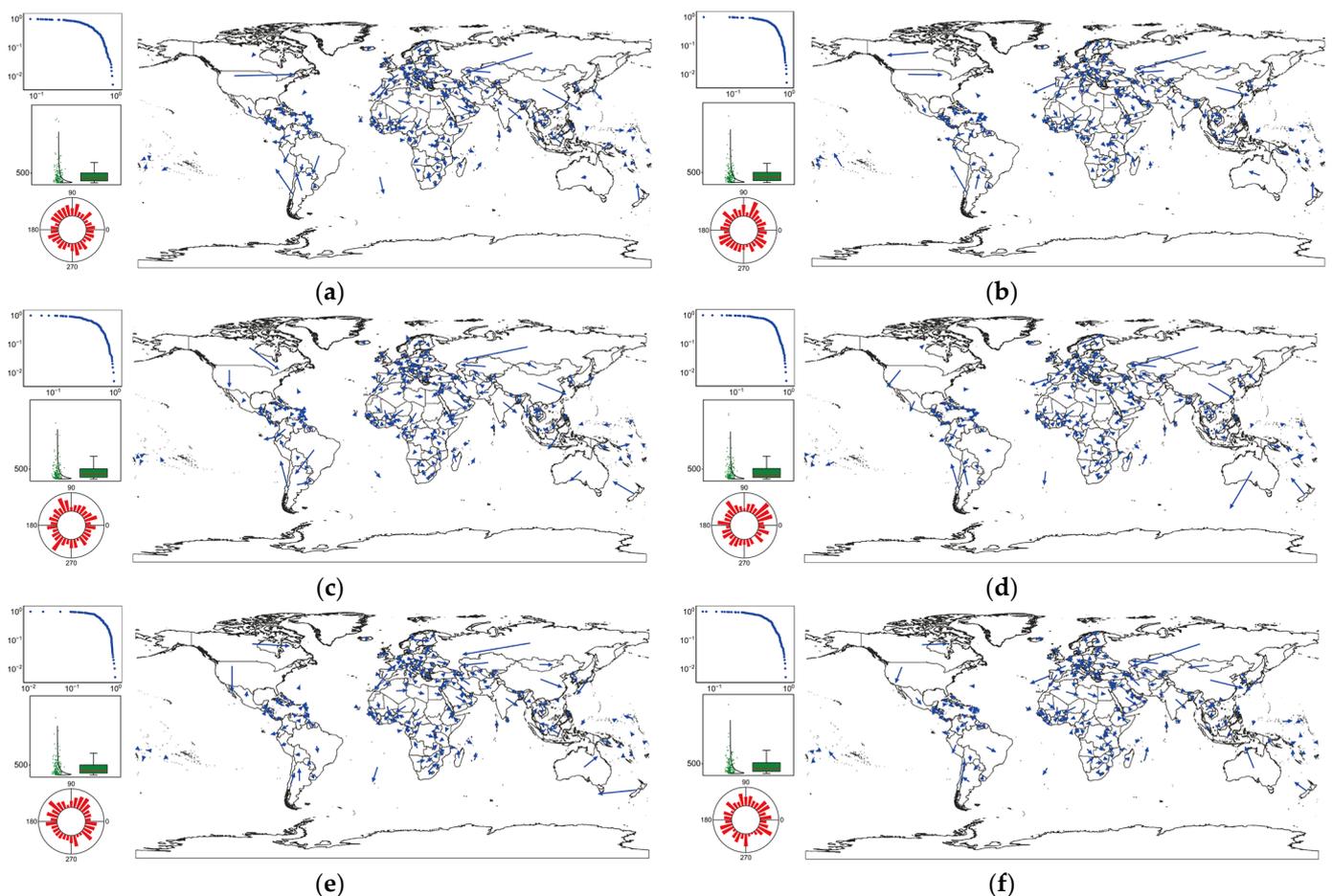
The variable that consistently provided the most accurate interpretation of node attractions for all periods, regardless of the direction, was GDP. The  $R^2$  values were most likely to be the greatest, which suggested that GDP is an effective and comprehensive measure of a country. The correlations between population sizes and node attractions were slightly poorer, and population sizes showed a stronger ability to interpret outward attractions compared to inward attractions. This can be attributed to the fact that population reflects not only opportunities but also competition and is not always the pulling factor. Moreover, although the outward FDI underperformed in most cases, the inward FDI could be a more effective factor in explaining inward attractions than population.

Regarding network metrics, the weighted in-degree and out-degree proved to be the most effective. To test the potential of this metric, they were next used to fit the gravity model, and the goodness of fitness values were 0.409, 0.425, 0.434, 0.437, 0.425, and 0.453 for all networks. This suggested that degrees were still not a good proxy for node attractions. Betweenness, although proven to be a useful measure of centrality, did not appear helpful in this case. It was interesting that the PageRank score performed well, surpassing the population in most periods, particularly in interpreting inward attractions.

To summarize, utilizing a single variable proved insufficient to represent node attractions. The concept of node attractions, which represents the inherent significance of a node, was found to be associated with numerous factors and metrics. Only by using our method can this intrinsic importance be discovered.

### 3.2.3. Node Position Patterns

The positions of nodes were another important output of our method. We found this position represented the actual migration centroid of a country. Due to the heterogeneous distribution of population and economic activity, this centroid of migration did not match the geometric center. In Figure 9, arrows starting from the geometric center and pointing to the calculated position were drawn, indicating the migration bias in the country. It should be noticed that the searching areas for node positions were set as the bounding box of the administrative area. Full results are listed in Table S2.

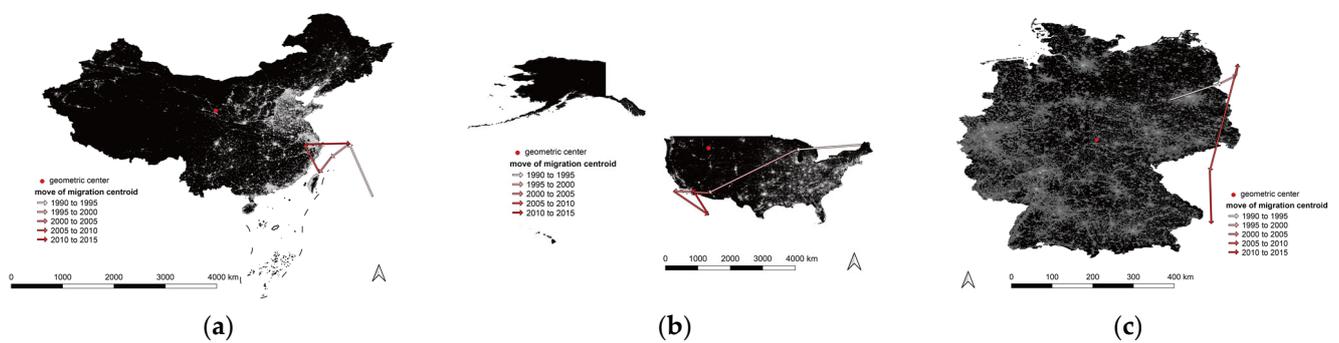


**Figure 9.** Geographic visualization and statistical analysis for node positions. Blue arrows from geometric centers to migration centroids were drawn on the map. On the left of the map, and from top to bottom, we have (1) the complementary cumulative distribution of the distance from the geometric center, normalized by the scale of the country; (2) the box plot of the distance in kilometers; (3) the distribution of orientation away from the geometric center. (a–f) refer to 6 periods.

Several conclusions could be drawn from the results. Statistically speaking, the normalized distance from the geometric center followed a power law distribution, over 70% of the distances were between 0.2 to 0.7. The box plots of distance reported a similar result, revealing that the median distance was about 230 km, with about 27% exceeding 500 km. The distributions of orientations away from the geometric center, on the other hand, showed a random distribution.

For different countries, the calculated position revealed different patterns. Three representative countries were selected as examples to illustrate the varying patterns and meanings of migration centroids. Using nighttime light data of 2020 as background, these three countries are shown in Figure 10. (1) China, as a country covering vast territory, has its population concentrated in southeastern coastal areas, and we found the trajectory

of the migration centroid moved around the same area, especially around Shanghai, the financial and foreign trade center of China. Similar countries were Russia (Figure S1a), Egypt (Figure S1b), Argentina (Figure S1c), etc.; these suggested the ability of migration centroids to indicate the imbalanced demography distribution within a country. (2) For the United States, the migration centroids moved across the country from the northeast, where the traditional industrial area was located, to the southwest, where the electronics industry has increased over the years. This example showed that migration centroids can be used to study the imbalanced developing patterns within a country. (3) The migration centroids of Germany always lay in the east part of the country, which was the area with a relatively low GDP; this may indicate that more migration happened in the less wealthy regions of the country. Other countries such as the United Kingdom (Figure S2a) and India (Figure S2b) shared similarities; here, the migration centroids moved to the north and southeast coastal regions, respectively. For these countries, the migration centroids revealed the heterogeneous economic distribution.



**Figure 10.** Spatial-temporal patterns of migration centroids of (a) China, (b) the United States, and (c) Germany. Nighttime light data of 2020 were taken as background, and arrows with different colors indicate the movement of the migration centroid in different years.

### 3.3. Comparative Analysis of Our Method

Next, the accuracy of our method was tested. Compared with different methods, the goodness of fitness of our method is the highest. For the traditional gravity model, we tested the performance of GDP and population as instrumental variables. The results in Table 2 show that GDP and population size can only account for 24–31% of migration, and population always serves as a better factor than GDP. Putting together GDP and population to get a multivariable model, the goodness of fitness elevated slightly. It is interesting that in this case the parameter  $\alpha_{out}$  was always bigger than  $\alpha_{in}$  for population variables, but for GDP variables it was always the opposite. For example, in 2015,  $\alpha_{out}^{pop}$ ,  $\alpha_{in}^{pop}$ ,  $\alpha_{out}^{GDP}$ , and  $\alpha_{in}^{GDP}$  were 0.389, 0.345, 0.075, and 0.106, which indicated that the population of the source country and the GDP of the target country had a greater influence on migration. This correlated with the common-sense idea that people tend to seek fortune and avoid competition via migration. Our method had also been compared with a reverse gravity method with dummy variables [53]; by introducing  $O_i$  and  $D_j$  such that  $O_i = D_j = 1$  when  $G_{ij} > 0$  and  $O_i = D_j = 0$  otherwise, Equation (6) can be converted into a linear regression problem:

$$\ln(G_{ij}) = \sum \ln(P_i^{out}) * O_i + \sum \ln(P_j^{in}) * D_j - \beta \ln(d_{ij}) + e_{ij}, \quad (7)$$

where  $e_{ij}$  is the error. To avoid multicollinearity and overfitting, ridge regression was used. The goodness of fitness in this method was better than that in the gravity model mentioned before. Our method also outperforms an improved gravity model including contiguity, sharing a common official language, colonial relationships, and other factors, as it was reported to reach a  $R^2$  of 0.604 for migration flows between 200 countries [28].

**Table 2.** The comparison of our method and other methods.

Period	Gravity Model			Reverse Gravity Model	
	GDP	Population	GDP & Population Multivariable	Dummy Variables	Our Method
1990–1995	0.249	0.310	0.327	0.444	0.668
1995–2000	0.249	0.316	0.329	0.454	0.683
2000–2005	0.271	0.319	0.335	0.461	0.682
2005–2010	0.285	0.314	0.335	0.463	0.684
2010–2015	0.305	0.309	0.334	0.464	0.688
2015–2020	0.248	0.308	0.322	0.476	0.696

#### 4. Discussion

By reconstructing the attractions and positions of countries, the gravitational relation that determines migration flow was found. This simple yet effective model could be used to simulate the complex international migration networks. As verified in the geographical preferential attachment model and geographical non-growing network model with vertex weights, this relation may also be the reason for scale-free and small-world properties [57,58].

The distance decay parameter  $\beta$  is an important indicator of the model. This parameter differs in spatial networks of different scales: for intra-urban trips collected from taxi trajectory data and social media check-in data,  $\beta$  is around 1.2–1.5 [59,60]; for inter-urban trips via flights or social media check-in data,  $\beta$  decreases to 0.8–0.9 [30,61]. For migration systems, this value varies from 1 to 4 for different countries [62,63], and it suggests that migration, unlike daily commuting, is more likely to be held back by the cost associated with distance. Here, we found  $\beta$  values were 2.98, 3.06, 2.76, 2.82, 3.03, and 2.92 for each period, respectively, implying a rapid decay in the amount of migration when moving distance increased.

Based on analysis of the results and the comparison with other methods, it can be concluded that the method we proposed is effective and of high accuracy. However, in terms of  $R^2$ , there is still room for improvement in this approach. On the one hand, the gravity model has inherent limitations. This model does not take into account factors like cultural or historical relations between countries, which are distance-independent and may influence the interaction. In addition, the radiation model [17], spatial Durbin model [64], and other improvements of the gravity model are widely used in spatial interaction research, which may also be used in the reverse gravity model. On the other hand, international migration as a whole is a complicated system, and there are community and other mesoscale structures in the network. This suggests that a single global model may not be enough, and instead, different models should be fitted to different sub-network structures. For example, if we extract the subgraph of 35 countries of the Organization for Economic Co-operation and Development from the migration network and apply our method to it, the goodness of fitness could reach about 0.78 for all periods, which is significantly superior to the global fitness.

Node attractions from our calculations correlated differently with various socioeconomic variables and network metrics; these six factors could only partly interpret the node attraction. A previous study also pointed out that in spatial interaction networks, degrees are the result of both attractions and locations [65]. This suggests that node attraction might be used as a criterion for evaluating node centrality. There are many centrality measures; however, the lack of ground truths on the importance of nodes makes it difficult to compare different metrics. The node attractions, as the determinant of interactive strength between nodes, participate in the generation of the network, which makes it a persuasive indicator.

Node positions may also have further implications. Using Equation (5), we found that the position of a node might not only be influenced by its domestic heterogeneity but could also be pulled by other nodes of great power. This phenomenon is more commonly seen in small countries. In other words, the meaning behind node positions may not only be the migration centroid, and further effort should be paid to this conception.

## 5. Conclusions

In this article, we proposed a reverse gravity approach based on genetic algorithm and applied this method to international migration networks from 1990 to 2020. Our method would model the interactive patterns with high accuracy, and the reconstructed node attractions and positions could reveal the characteristics of the country. The node attractions were highly correlated with socioeconomic factors and network metrics, and the node positions outperformed the geometric centers from the perspective of human migration. This method can be used in any other spatial interaction network, which are found extensively in sustainable systems in the real world. Our research presents a tentative solution to the problem of distance calculation when fitting gravity models in migration networks. The limitation of this method is that the node position is influenced both by the domestic heterogeneous distributions of economy and demography and foreign attractions; this complexifies interpretation, especially for small countries. In future research, the implications behind node position should be further investigated, aiming to obtain a more accurate model fitting and a deeper understating of interactive patterns.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su16062502/s1>, Figure S1: Spatial-temporal patterns of migration centroids of (a) Russia, (b) Egypt, and (c) Argentina. Nighttime light data of 2020 were taken as background, and arrows with different colors indicate the movement of migration centroid in different years; Figure S2: Spatial-temporal patterns of migration centroids of (a) the United Kingdom and (b) India. Nighttime light data of 2020 were taken as background, and arrows with different colors indicate the movement of migration centroid in different years; Table S1: Node attractions; Table S2: Node positions; Table S3 distance away from geometric center.

**Author Contributions:** Conceptualization, X.Y., T.J. and K.Q.; methodology, X.Y. and T.J.; software, X.Y.; formal analysis, X.Y.; writing—original draft preparation, X.Y.; writing—review and editing, X.Y. and Y.Z.; visualization, X.Y. and X.G.; supervision, T.J. and K.Q. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (grant numbers: 42171448; 41971332) and National Social Science Foundation of China (No. 21&ZD332).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. The international migration data can be found at [https://figshare.com/collections/Bilateral\\_international\\_migration\\_flow\\_estimates\\_for\\_200\\_countries/4470464](https://figshare.com/collections/Bilateral_international_migration_flow_estimates_for_200_countries/4470464), accessed on 10 March 2024.

**Acknowledgments:** We thank all authors for their contributions and the anonymous reviewers for helpful comments and suggestions. We are grateful for the funding support.

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

## References

1. Raymer, J.; Willekens, F.; Rogers, A. Spatial demography: A unifying core and agenda for further research. *Popul. Space Place* **2019**, *25*, e2179. [[CrossRef](#)]
2. Lewis, G.J. *Human Migration: A Geographical Perspective*; Routledge: Oxfordshire, UK, 2021.
3. Likić-Brborić, B. Global migration governance, civil society and the paradoxes of sustainability. *Globalizations* **2018**, *15*, 762–778. [[CrossRef](#)]
4. Gavonel, M.F.; Adger, W.N.; de Campos, R.S.; Boyd, E.; Carr, E.R.; Fabos, A.; Fransen, S.; Jolivet, D.; Zickgraf, C.; Codjoe, S.N. The migration-sustainability paradox: Transformations in mobile worlds. *Curr. Opin. Environ. Sustain.* **2021**, *49*, 98–109. [[CrossRef](#)]
5. Tranos, E.; Gheasi, M.; Nijkamp, P. International migration: A global complex network. *Environ. Plan. B Plan. Des.* **2015**, *42*, 4–22. [[CrossRef](#)]
6. Leal, D.F.; Harder, N.L. Migration networks and the intensity of global migration flows, 1990–2015. *J. Ethn. Migr. Stud.* **2023**, *49*, 445–464. [[CrossRef](#)]
7. Davis, K.F.; D’Odorico, P.; Laio, F.; Ridolfi, L. Global Spatio-Temporal Patterns in Human Migration: A Complex Network Perspective. *PLoS ONE* **2013**, *8*, e53723. [[CrossRef](#)]

8. Fagiolo, G.; Mastrorillo, M. International migration network: Topology and modeling. *Phys. Rev. E* **2013**, *88*, 012812. [[CrossRef](#)]
9. Fagiolo, G.; Mastrorillo, M. Does Human Migration Affect International Trade? A Complex-Network Perspective. *PLoS ONE* **2014**, *9*, e97331. [[CrossRef](#)]
10. Abel, G.J.; DeWaard, J.; Ha, J.T.; Almquist, Z.W. The form and evolution of international migration networks, 1990–2015. *Popul. Space Place* **2021**, *27*, e2432. [[CrossRef](#)]
11. Gao, S.; Liu, Y.; Wang, Y.; Ma, X. Discovering spatial interaction communities from mobile phone data. *Trans. GIS* **2013**, *17*, 463–481. [[CrossRef](#)]
12. Li, Z.; Tang, J.; Zhao, C.; Gao, F. Improved centrality measure based on the adapted PageRank algorithm for urban transportation multiplex networks. *Chaos Solitons Fractals* **2023**, *167*, 112998. [[CrossRef](#)]
13. Haynes, K.E.; Fotheringham, A.S. *Gravity and Spatial Interaction Models*; Regional Research Institute, West Virginia University: Morgantown, WV, USA, 2020.
14. Oshan, T.M. The spatial structure debate in spatial interaction modeling: 50 years on. *Prog. Hum. Geogr.* **2021**, *45*, 925–950. [[CrossRef](#)]
15. Ramos, R. Modelling migration. In *The Econometrics of Multi-Dimensional Panels: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 455–478.
16. Ramos, R. *Gravity Models: A Tool for Migration Analysis*; IZA World of Labor; Institute for the Study of Labor (IZA): Bonn, Germany, 2016.
17. Simini, F.; González, M.C.; Maritan, A.; Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **2012**, *484*, 96–100. [[CrossRef](#)]
18. Liu, E.-J.; Yan, X.-Y. A universal opportunity model for human mobility. *Sci. Rep.* **2020**, *10*, 4657. [[CrossRef](#)]
19. Zipf, G.K. The P1 P2/D Hypothesis: On the Intercity Movement of Persons. *Am. Sociol. Rev.* **1946**, *11*, 677–686. [[CrossRef](#)]
20. Dueñas, M.; Fagiolo, G. Modeling the International-Trade Network: A gravity approach. *J. Econ. Interact. Coord.* **2013**, *8*, 155–178. [[CrossRef](#)]
21. Krings, G.; Calabrese, F.; Ratti, C.; Blondel, V.D. Urban gravity: A model for inter-city telecommunication flows. *J. Stat. Mech. Theory Exp.* **2009**, *2009*, L07003. [[CrossRef](#)]
22. Liu, J.H.; Zhang, Z.K.; Chen, L.J.; Liu, C.; Yang, C.C.; Wang, X.Q. Gravity Effects on Information Filtering and Network Evolving. *PLoS ONE* **2014**, *9*, e91070. [[CrossRef](#)]
23. Anderson, J.E. The Gravity Model. *Annu. Rev. Econ.* **2011**, *3*, 133–160. [[CrossRef](#)]
24. Shepherd, B. *The Gravity Model of International Trade: A User Guide*; United Nations ESCAP: Bangkok, Thailand, 2013.
25. Alis, C.; Legara, E.F.; Monterola, C. Generalized radiation model for human migration. *Sci. Rep.* **2021**, *11*, 22707. [[CrossRef](#)]
26. Niu, F. A push-pull model for inter-city migration simulation. *Cities* **2022**, *131*, 104005. [[CrossRef](#)]
27. Yan, X.Y.; Zhao, C.; Fan, Y.; Di, Z.; Wang, W.X. Universal predictability of mobility patterns in cities. *J. R. Soc. Interface* **2014**, *11*, 20140834. [[CrossRef](#)] [[PubMed](#)]
28. Ramos, R.; Suriñach, J. A Gravity Model of Migration Between the ENC and the EU. *Tijdschr. Voor Econ. Soc. Geogr.* **2017**, *108*, 21–35. [[CrossRef](#)]
29. Shen, G. Reverse-fitting the gravity model to inter-city airline passenger flows by an algebraic simplification. *J. Transp. Geogr.* **2004**, *12*, 219–234. [[CrossRef](#)]
30. Xiao, Y.; Wang, F.H.; Liu, Y.; Wang, J. Reconstructing Gravitational Attractions of Major Cities in China from Air Passenger Flow Data, 2001–2008: A Particle Swarm Optimization Approach. *Prof. Geogr.* **2013**, *65*, 265–282. [[CrossRef](#)]
31. Abel, G.J.; Cohen, J.E. Bilateral international migration flow estimates for 200 countries. *Sci. Data* **2019**, *6*, 82. [[CrossRef](#)]
32. Benveniste, H.; Oppenheimer, M.; Fleurbaey, M. Climate change increases resource-constrained international immobility. *Nat. Clim. Chang.* **2022**, *12*, 634–641. [[CrossRef](#)]
33. Petropoulos, F.; Apiletti, D.; Assimakopoulos, V.; Babai, M.Z.; Barrow, D.K.; Taieb, S.B.; Bergmeir, C.; Bessa, R.J.; Bijak, J.; Boylan, J.E. Forecasting: Theory and practice. *Int. J. Forecast.* **2022**, *38*, 705–871. [[CrossRef](#)]
34. Mayda, A.M. International migration: A panel data analysis of the determinants of bilateral flows. *J. Popul. Econ.* **2010**, *23*, 1249–1274. [[CrossRef](#)]
35. Garas, A.; Lapatinas, A.; Poullos, K. The Relation between Migration and Fdi in the Oecd from a Complex Network Perspective. *Adv. Complex Syst.* **2016**, *19*, 1650009. [[CrossRef](#)]
36. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
37. Li, Y.L.; Chen, B.; Chen, G.Q. Carbon network embodied in international trade: Global structural evolution and its policy implications. *Energy Policy* **2020**, *139*, 111316. [[CrossRef](#)]
38. Clauset, A.; Shalizi, C.R.; Newman, M.E.J. Power-Law Distributions in Empirical Data. *Siam Rev.* **2009**, *51*, 661–703. [[CrossRef](#)]
39. Humphries, M.D.; Gurney, K.; Prescott, T.J. The brainstem reticular formation is a small-world, not scale-free, network. *Proc. R. Soc. B Biol. Sci.* **2006**, *273*, 503–511. [[CrossRef](#)] [[PubMed](#)]
40. Fagiolo, G. Clustering in complex directed networks. *Phys. Rev. E* **2007**, *76*, 026107. [[CrossRef](#)]
41. Freeman, L.C. Centrality in Social Networks Conceptual Clarification. *Soc. Netw.* **1979**, *1*, 215–239. [[CrossRef](#)]
42. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [[CrossRef](#)] [[PubMed](#)]
43. Jia, T.; Yu, X.S.; Li, X.; Qin, K. Identification and analysis of urban influential regions using spatial interaction networks. *Trans. GIS* **2021**, *25*, 2821–2839. [[CrossRef](#)]

44. Brandes, U. A faster algorithm for betweenness centrality. *J. Math. Sociol.* **2001**, *25*, 163–177. [[CrossRef](#)]
45. Brin, S.; Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. Isdn Syst.* **1998**, *30*, 107–117. [[CrossRef](#)]
46. Masuda, N.; Porter, M.A.; Lambiotte, R. Random walks and diffusion on networks. *Phys. Rep. Rev. Sect. Phys. Lett.* **2017**, *716*, 1–58. [[CrossRef](#)]
47. Barbosa, H.; Barthelemy, M.; Ghoshal, G.; James, C.R.; Lenormand, M.; Louail, T.; Menezes, R.; Ramasco, J.J.; Simini, F.; Tomasini, M. Human mobility: Models and applications. *Phys. Rep. Rev. Sect. Phys. Lett.* **2018**, *734*, 1–74. [[CrossRef](#)]
48. Morley, C.; Rosselló, J.; Santana-Gallego, M. Gravity models for tourism demand: Theory and use. *Ann. Tour. Res.* **2014**, *48*, 1–10. [[CrossRef](#)]
49. Windzio, M. The network of global migration 1990–2013 Using ERGMs to test theories of migration between countries. *Soc. Netw.* **2018**, *53*, 20–29. [[CrossRef](#)]
50. Wang, Y.Q.; Luo, H.; Shi, Y.Y. Complex network analysis for international talent mobility based on bibliometrics. *Int. J. Innov. Sci.* **2019**, *11*, 419–435. [[CrossRef](#)]
51. Okelly, M.E.; Song, W.; Shen, G.Q. New Estimates of Gravitational Attraction by Linear-Programming. *Geogr. Anal.* **1995**, *27*, 271–285. [[CrossRef](#)]
52. Shen, G.Q. Estimating nodal attractions with exogenous spatial interaction and impedance data using the gravity model. *Pap. Reg. Sci.* **1999**, *78*, 213–220. [[CrossRef](#)]
53. Song, W. Nodal Attractions in China’s Intercity Air Passenger Transportation. In Proceedings of the 29th Annual Applied Geography Conference, Tampa, FL, USA, 11–14 October 2006.
54. Garg, H. A hybrid PSO-GA algorithm for constrained optimization problems. *Appl. Math. Comput.* **2016**, *274*, 292–305. [[CrossRef](#)]
55. Holland, J.H. Genetic algorithms. *Sci. Am.* **1992**, *267*, 66–73. [[CrossRef](#)]
56. Tinbergen, J. *Shaping the World Economy; Suggestions for an International Economic Policy*; Twentieth Century Fund: New York, NY, USA, 1962.
57. Manna, S.S.; Sen, P. Modulated scale-free network in Euclidean space. *Phys. Rev. E* **2002**, *66*, 066114. [[CrossRef](#)] [[PubMed](#)]
58. Masuda, N.; Miwa, H.; Konno, N. Geographical threshold graphs with small-world and scale-free properties. *Phys. Rev. E* **2005**, *71*, 036108. [[CrossRef](#)]
59. Liu, Y.; Kang, C.; Gao, S.; Xiao, Y.; Tian, Y. Understanding intra-urban trip patterns from taxi trajectory data. *J. Geogr. Syst.* **2012**, *14*, 463–483. [[CrossRef](#)]
60. Noulas, A.; Scellato, S.; Lambiotte, R.; Pontil, M.; Mascolo, C. A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE* **2012**, *7*, e37027. [[CrossRef](#)]
61. Liu, Y.; Sui, Z.; Kang, C.; Gao, Y. Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. *PLoS ONE* **2014**, *9*, e86026. [[CrossRef](#)]
62. Stillwell, J.; Bell, M.; Ueffing, P.; Daras, K.; Charles-Edwards, E.; Kupiszewski, M.; Kupiszewska, D. Internal migration around the world: Comparing distance travelled and its frictional effect. *Environ. Plan. A Econ. Space* **2016**, *48*, 1657–1675. [[CrossRef](#)]
63. Bernard, A.; Rowe, F.; Bell, M.; Ueffing, P.; Charles-Edwards, E. Comparing internal migration across the countries of Latin America: A multidimensional approach. *PLoS ONE* **2017**, *12*, e0173895. [[CrossRef](#)] [[PubMed](#)]
64. Durbin, J. Estimation of Parameters in Time-Series Regression-Models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1960**, *22*, 139–153. [[CrossRef](#)]
65. Liu, Y.; Gong, L.; Tong, Q. Quantifying the Distance Effect in Spatial Interactions. *Acta Sci. Nat. Univ. Pekin.* **2014**, *50*, 526–534.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.