Spatial Patterns and Development Characteristics of China’s Postgraduate Education: A Geographic Information System Approach

Ping Li, Zhejiang Wanli University, China
Haidong Zhong, Ningbo University of Technology, China
Justin Zuopeng Zhang, University of North Florida, USA*

ABSTRACT

Using four types of publicly available datasets and ArcGIS software, the authors identify the spatial characteristics of postgraduate education in China at three scales: comprehensive economic zone, provincial, and city. They also employ geographically weighted regression and ordinary least squares to study the factors influencing the spatial pattern of postgraduate education in China at the city scale. The findings show that the number of postgraduate education institutions increases as the longitude of a city increases, but the number decreases from coast to inland. Second, postgraduate education institutions tend to group together in provincial capitals and megacities. Finally, GDP, per capita GDP, population size, local income, and total retail sales of consumer goods significantly impact postgraduate education development. The study contributes to the literature and provides insights for practitioners in promoting urban planning and infrastructure development.

KEYWORDS

China, Geographic Information System, Postgraduate Education, Spatial Analysis, Spatial Pattern

INTRODUCTION

Higher education resources have become one of the most important factors of competitiveness in a country or region due to the rapid development of the knowledge economy. Many studies have shown that rich higher education resources promote the development of human resources and improve labor quality (Jiang et al., 2019; Lao & Xue, 2016). The uneven regional distribution of the resources gathered or brought about by higher education is deemed an important reason that leads to the development disparity in different regions of a country. Since the 1980s, China has increased its investment in higher education, leading to a stable development trend. However, due to differences in geographic location, policies, and economic development between regions, an uneven distribution of higher education resources has gradually emerged.

Postgraduate education in China is the highest level of national higher education. Postgraduate students fall into two types, divided into master and doctoral students. Unlike undergraduate education,
which is usually completed in universities, postgraduate education is generally conducted in various postgraduate education institutions, such as research institutes, well-known enterprises, colleges or universities, and government organizations. During the undergraduate stage, students typically receive general education and have a theoretical basis after obtaining their bachelor’s degree, whereas the postgraduate stage is more likely to focus on professional education and on improving students’ professional skills (Borodako et al., 2021; Grierson & Munro, 2018; Olivier et al., 2020). Therefore, the geographical distribution of postgraduate cultivation institutions reflects the general information on postgraduate education-related resource allocation over a given historical period (Baillette & Barlette, 2021; Guo & Wang, 2020; Gibbons & Vignoles, 2012; Kumar et al., 2021; Liu et al., 2021; Rouibah et al., 2020).

On the one hand, postgraduate education institutions are usually located where higher education activities occur. These locations are chosen by various research institutes, well-known enterprises, colleges, or universities based on the natural and social environment on which they rely. Therefore, the spatial distribution pattern of postgraduate education institutions reflects the relationship between social politics, economy, culture, and education in a certain space (Ali et al., 2021; Al-Hasan et al., 2021; Islam et al., 2016; Kumar et al., 2021; Meso et al., 2021; Sohaib, 2021; Zhang & Srite, 2021). On the other hand, the characteristics and rules of the spatial distribution of postgraduate education institutions provide a reference for the development of regional social, political, economic, cultural, and educational activities, which influence not only current but also future social development, particularly many aspects of the development of higher education (Ahluwalia, & Merhi, 2020; Chang et al., 2020; Ensslin et al., 2020; Ge et al., 2020; Liu & Du, 2020; Sengupta, 2020).

Compared with Western developed countries, China’s postgraduate education system had a late start, but is developing rapidly. Since 1949, the country has gone from decline to prosperity and blazed a path of fast and steady development. The path is proved to be in accordance with China’s economic and social characteristics during every development stage (Halawani et al., 2020; Ifitikhar & Khan, 2020; Rialp-Criado et al., 2020; Vatanasakdakul et al., 2020; B. Zhang et al., 2021). China’s Ministry of Education’s annual statistical data bulletin shows that the total number of postgraduate education institutions reached 0.8 thousand in 2018 for the first time. Meanwhile, the number of in-school postgraduate students went up to 2700 thousand, and the number of overseas postgraduate students reached about 80 thousand due to China’s rapid economic and social development and continued investment in postgraduate education.¹ Behind this large-scale postgraduate education development, China has transformed its focus from quantity to quality, and has made many detours.

It is generally believed that a postgraduate degree is more valuable than an undergraduate degree, as the contribution of postgraduates to regional economic and social development is more remarkable than that of undergraduates. However, most studies only pay attention to the spatial pattern, educational reasons, and factors influencing undergraduates in China. There are few research results in this field on the spatial distribution of postgraduate education and its formation mechanism. This paper focuses on the spatial pattern of postgraduate education and the causes of spatial aggregation in China.

This paper has two main contributions. First, the authors analyze the spatial distribution characteristics of China’s postgraduate education from a geographical point of view. Based on the dataset collected from related government departments and authoritative research institutions, the authors determined the postgraduate education institutions’ location (in [latitude, longitude] format) with the help of the Baidu Maps development platform.² Following the Geographic Information System (GIS) approach (Changchit et al., 2021; Gallego-Gomez et al., 2021; Huang et al., 2021; Kesharwani et al., 2021; Lee et al., 2021; Peng, Chen et al., 2021; Srivastava & Eachempati, 2021; Xanthidis & Xanthidou, 2021; ZareRavasan, & Krčál, 2021; Zhang et al., 2021), the authors quantitatively explore the spatial aggregation characteristics of postgraduate education in China, using ArcGIS software and related tools or documents, at three scales: the comprehensive economic zone, the provincial, and the city scale. Second, the authors follow the social and economic geography viewpoint to quantitatively explore factors that may influence the spatial development pattern of China’s postgraduate education.
Based on the existing research conclusions, the authors collect many widely adopted indices to measure regional economic and social development, such as per capita GDP, population size, local income, etc. With the publicly available data, the authors follow the geographical analysis method and jointly use GWR, autocorrelation, and spatial development trend analysis in ArcGIS software to study possible reasons that may lead to the spatial aggregation pattern of China’s postgraduate education.

The rest of the paper is structured as follows. The second section titled Related Work reviews prior research on the analysis of higher education spatial distribution characteristics, development level evaluation, and other related aspects. The third section titled Data Description and Study Approaches details the data source and research methods. The fourth section titled Spatial Development Characteristics of Postgraduate Education in China investigates the spatial patterns and China’s postgraduate education. The fifth section titled Factors Influencing the Spatial Pattern of Postgraduate Education in China analyzes influencing factors in the spatial pattern of postgraduate education in China. The final section concludes the research and proposes future research directions.

RELATED WORK

Higher education agglomeration refers to a large number of interrelated schools and their supporting institutions coming together in space, with one or several well-known universities as the core to create a strong and sustainable competitive advantage (Gulson & Symes, 2007; Mao et al., 2019; Walsh, 1992). The rise of new economic geography provides a unique perspective for studying the spatial aggregation of human capital. Several studies have theoretically and empirically shown that education is a determinant of regional economic growth in many countries, for example, Romania, Malaysia, the US, the UK, etc. (Islam et al., 2016; Lauder, 2015; Simionescu et al., 2017). Fujita and Thisse (2003) explored the difference between human capital and ordinary workers in economic growth and found regional economic development and welfare changes were clearly correlated to human capital agglomeration. Considering entrepreneurs as a highly mobile factor of production in the new economic geography model, Forslid and Ottaviano (2003) affirmed that market expansion could further attract entrepreneurs to congregate in a particular region. Based on a quantitative analysis of the internal mechanism of how education fosters social cohesion, Camilleri and Camilleri (2016) claimed that education cultivates high-quality talents to improve the knowledge level of the regional population and economic and social development circumstances (Camilleri & Camilleri, 2016; Law et al., 2021; Liu et al., 2021; Mengesha et al., 2021; Trappey et al., 2021; Uniyal et al., 2021; Yao et al., 2021).

Compared with many developed countries, the spatial agglomeration of human capital and higher education in China is peculiar. Many Chinese scholars have focused on the spatial aggregation of higher education in terms of causes, patterns, and distribution characteristics (Bai et al., 2018; Gao et al., 2018; Liu, 2019; You et al., 2016). For example, Qian & Smyth (2008) calculated the Theil index and the Gini coefficient of relevant Chinese data from 1990 to 1998. They concluded that the gap between urban and rural educational resources and inland and coastal educational resources was increasingly significant with an expanding trend (Qian & Smyth, 2008; A. Zhang et al., 2021). A. Zhang et al. (2021) compared and analyzed the unbalanced distribution of China’s educational resources between rural and urban areas and inland and coastal areas using the Gini coefficient and other methods. They found that the unbalanced distribution of China’s educational resources was mainly caused by the uneven distribution of educational resources between urban and rural areas (Zhang & Kanbur, 2009). Jacob (2006) studied the balance between higher education resources and the fairness of students’ right to education in China and identified a disparity in the geographic distribution of higher education and a discrepancy in students’ right to education.

Regarding the causes of higher education agglomeration, many researchers have suggested that the main reason is the need for resource sharing and discipline integration among universities. Many studies have shown that the concentration of higher education contributes to the agglomeration of scientific research, improves the overall strength of a university, and reinforces the overall strength
of a city and the interactive development of universities and industries (Wen & Shi, 2005). Moreover, many research findings have indicated that urbanization is an important factor affecting higher education agglomeration. For instance, Ge (2008) found that the distribution of urban space gradually developed from a sporadic “dot” to a “belt,” and then to a “net” in China over the past half-century, inevitably leading to higher education agglomeration (Ge, 2008). Li (2021) applied the multiple regression analysis methods and pointed out that the main reasons for the change in talent distribution included higher talent cultivation, urban development pattern, urbanization level of a nation, the wage level, and government policies (Almuraqab et al., 2021; Han et al., 2021; Jiang & Li, 2006; Li, 2021; Mutambik et al., 2021; Olesen et al., 2021; Talukder et al., 2020). Based on publicly available statistics from mainland China, Liu et al. (2013) selected five indices: the number of universities, the average expenditure per student, the investment in education, the teacher-student ratio, and the number of students, and conducted an empirical study of the disequilibrium and polarization of university student cultivate institutions, capital investment, etc. They found the following results: 1) the distribution of universities and scientific research institutes exhibit significant spatial disequilibrium characteristics; 2) the regional gap in the distribution of universities and scientific research institutes measured by the total index is much larger than the regional gap measured by the relative index and; 3) the interregional economic development gap is the most crucial cause that leads to the overall disparity (Liu et al., 2013).

The spatial agglomeration model of Chinese higher education is another hot research topic. For example, Wang (2010) investigated the process of urban agglomeration expansion in the Yangtze River Delta region and explored whether there was a kind of intrinsic relationship between economic development and higher education in this area using spatial analysis in geography. Based on spatial statistics, they found that the spatial distribution characteristics of higher education in the Yangtze River Delta region were highly consistent with regional economic and social development (Wang, 2010). Gu (2010) calculated the degree of spatial autocorrelation of educational resources for undergraduates and identified four educational resource allocation clusters: high–high, low–low, low–high, and high–low, respectively. This study suggested that higher education agglomeration may have started 10 years ago (Gu, 2010). Hou & Xue (2009) studied the spatial pattern of universities, research institutions, and other higher education-related resources in China systematically and quantitatively using GIS technology and discovered a severe imbalance appearance (Abdou & Jasimuddin, 2020; Almomani et al., 2021; Chen et al., 2021; Gholami et al., 2021; Hou & Xue, 2009; Jia et al., 2021; Liu & Yu, 2021; Mondal & Chakrabarti, 2021). The pattern of agglomeration of universities is mainly reflected in the aggregation of political centers, economic centers, and cities with rapid economic development (Chen, 2008). From the perspective of population and GDP, Zhao et al. (2007) studied the imbalance and differences between the distribution of Chinese universities in provincial and prefectural regions. They pointed out that the distribution of Chinese universities in the provincial areas was relatively balanced, but the distribution in provincial capitals and prefectural cities was unbalanced, with a high density in provincial capitals (Zhao et al., 2007).

DATA DESCRIPTION AND STUDY APPROACHES

Data Description

In the research, four types of datasets are used. Their types and sources are summarized as follows:

1. Basic geographic data of China (including areas, rivers, cities, and county-level administrative boundaries). The data was downloaded from the national basic geographic data release and sharing platform.
2. Postgraduate student statistics, including the number of degree awardees and students, enrolled in all provinces (except the Hong Kong Special Administrative Region, the Macao Special
Administrative Region, and the Taiwan province) in China in 2018. The data was gathered through the official release website of the Ministry of Education of the People’s Republic of China.  

3. Full names of postgraduate cultivation institutions in China in 2018. The data was collected and sorted from the Graduate Enrollment Information Network, the only comprehensive postgraduate recruitment service platform designated by the Ministry of Education of the People’s Republic of China. However, the data did not include postgraduate cultivation institutions in the Hong Kong Special Administrative Region, the Macao Special Administrative Region, or the Taiwan province.

4. Geographic coordinates of all postgraduate education institutions. The data was obtained using a self-developed application based on the Baidu Maps Open Platform. In this self-developed application, the geographic coordinates of all postgraduate education institutions are stored in .txt format and exported to ArcGIS 10.5 for spatial pattern analysis.

Study Approaches

Very limited research has explored the spatial distribution characteristics of postgraduate education in China from a geographic perspective. This paper uses relevant theories and methods in GIS as a reference and adopts spatial data mining, ArcGIS-based spatial analysis, and geographic statistical analysis to study the development characteristics of postgraduate education in China at three scales: comprehensive economic zone, province, and city scale.

Spatial Autocorrelation and Aggregation Analysis

Existing studies have shown that the Moran’s I is one of the most widely adopted approaches to evaluate the degree of interdependence between data from one location and data from other locations (Tiefelsdorf, 2002; Tillé et al., 2018). It is so popular in spatial autocorrelation statistics that the whole idea and the related functional interfaces of Moran’s I are implemented in ArcGIS software. With the spatial statistics tools in ArcGIS 10.5 or a later version, Moran’s I value can be easily calculated. Meanwhile, another indicator, z-score, which indicates the validity of the computed result, is also reported. Generally, a z-score within \((-\infty, -1.96]\) or \([1.96, +\infty]\) indicates the analyzed result can be deemed effective and credible (with the real approval rating level of 0.95). Otherwise, the analyzed result suggests a clustered spatial distribution. If the z-score value is less than -1.96, the analyzed result indicates a discrete spatial distribution. Otherwise, the analyzed result indicates a spatial distribution that can be considered random. Moran’s I is calculated as follows:

\[
I = \frac{n}{S_0} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j \sum_{i=1}^{n} z_i^2
\]  

In formula (1), \(z_i\) represents the difference between a measured number obtained by various means and the expected value in a specific place \(i\), \(w_{i,j}\) represents the weight of proximity between places \(i\) and \(j\), \(n\) represent the total number of the observed places, \(S_0\) is the total number of \(w_{i,j}\), and \(z_i\) can be calculated according to the formula below:

\[
z_i = \frac{I - E[I]}{\sqrt{V[I]}}
\]
In formula (2), \( E[I] = \frac{1}{n-1} \) and \( [V] = E[I^2] - E[I]^2 \).

Moran’s I cannot reveal whether the spatial data congregate is at high or low values. Given a large amount of previous research, the authors selected the global \( G \) coefficient approach to measure the aggregation extent of postgraduate education in China. The \( G \) coefficient can be calculated using the following formula (Carrijo & da Silva, 2017):

\[
G = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} x_i x_j, \forall j \neq i
\]

In formula (3), \( x_i \) and \( x_j \) are the values measured at locations \( i \) and \( j \), respectively, and \( w_{i,j} \) represents the weight that indicates the importance of the distance between location \( i \) and \( j \). The \( G \) coefficient can be standardized as follows:

\[
z_G = \frac{G - E[G]}{\sqrt{P[G]}}
\]

In formula (4), \( P[G] = E[G^2] - E[G]^2 \). The method of \( G \) coefficient standardization is also integrated into ArcGIS spatial analysis tools.

**GWR Analysis**

To improve the common challenge of embedding the stationarity hypothesis in traditional regression models for geographic reasoning, Fotheringham et al. (2015) proposed a GWR model that directly simulates unstable data and allows for the use of locally estimated values instead of the value that needs to be calculated globally. The GWR model is a new and effective method for studying spatial relationships. In the GWR model, the regression varies, or parameters at specific locations are extracted directly from the test data; the observed values of adjacent locations are utilized to deduce the regression on a small scale, which changes with the change in spatial location. The GWR model can be formalized as follows (Lin et al., 2011):

\[
y_i = \beta_0(u_i, v_i) + \sum_{i=1}^{k} \beta_i(u_i, v_i) x_{ik} + \varepsilon_i
\]

In formula (5), \( (u_i, v_i) \) represents the geographic coordinate value at location \( i \), \( \beta_i \) is used to estimate its adjacent spatial observations, which vary with location, and \( \varepsilon_i \) is a constant.

**Ordinary Least Squares (OLS) Analysis**

It is generally believed that regression analysis is one of the most important statistical methods in social sciences-related studies (Hakulinen & Tenkanen, 1987). According to the basic principle of the idea of the regression analysis model, it has an obvious advantage in evaluating the relationship between two or more varies. With the help of quantitative regression analysis results, it is very easy to explain why something happens in a place. Also, the quantitative regression analysis result can
be used to predict what will happen in the place in the future (Duncan & Hodge, 1963). Ordinary least squares (OLS) is one of the best-known regression approaches and is deemed the basis for different kinds of other spatial regression analysis models. In particular, OLS is considered to have unique advantages in exploring or modeling the relationship between multiple variables (Zhang et al., 2019). According to the existing research, OLS regression analysis can be used in the following three scenarios (Ceccato & Oberwittler, 2008; Martellosio, 2011; Yue et al., 2018):

1. First, to explain a phenomenon with many explanatory varies and evaluate its potential influence on policymaking and decisions on appropriate actions based on that understanding. In this case, the main task is to determine what will happen if one related variable changes or a set of related variables changes differently, such as understanding the main habitat pattern of certain endangered birds to help protect the species through legislation.
2. Second, to build a consistent and accurate predictive model for an event to predict the result with a set of given condition values. For example, how much electricity will be consumed in the following year based on population growth and typical weather conditions in a given year.
3. Third, to explore some hypothetical situations, for example, modeling and simulation analysis of the correlation between regional social development and income level of a given city to better understand economic development and implement strategies to increase wages.

The OLS method is implemented and encapsulated in the ArcGIS spatial statistical analysis toolkit. With the help of ArcGIS software, the OLS analysis tool can generate a series of useful reports and optional tables, which plays a vital role in validating whether the established mode or the result is credible.

SPATIAL DEVELOPMENT CHARACTERISTICS OF POSTGRADUATE EDUCATION IN CHINA

To explore the spatial distribution of postgraduate education in China, the authors use ArcGIS 10.5 to analyze the spatial pattern of postgraduate education institutions by adding data on postgraduate education institutions .txt format to ArcMap 10.5, which is an important component of ArcGIS. With the help of the “Display XY data” function in ArcMap 10.5, postgraduate education institutions can readily be displayed on an administrative map of China according to their coordinates.

Analysis at the Comprehensive Economic Zone Scale
According to China’s coordinated regional development strategy, the nation falls into eight comprehensive-economic-zones (Zeng et al., 2012): the Northeast Comprehensive Economic Zone, which covers Liaoning Province, Jilin Province, and Heilongjiang Province; the Great Northwest Comprehensive Economic Zone, which covers Gansu Province, Qinghai Province, the Ningxia Hui Autonomous Region, the Tibet Autonomous Region, and the Xinjiang Uygur Autonomous Region; the Northern Coastal Comprehensive Economic Zone, which covers Beijing, Tianjin, Hebei Province, and Shandong Province; the Middle Reaches of the Yangtze River Comprehensive Economic Zone, which covers Hubei Province, Hunan Province, Jiangxi Province, and Anhui Province; the Middle Reaches of the Yellow River Comprehensive Economic Zone, which covers Shanxi Province, Henan Province, and the Inner Mongolia Autonomous Region; the Southern Coastal Comprehensive Economic Zone, which covers Fujian Province, Guangdong Province, and Hainan Province; the East Coastal Comprehensive Economic Zone, which covers Shanghai, Jiangsu Province, and Zhejiang Province; and the Great Southwest Comprehensive Economic Zone, which covers Yunnan Province, Guizhou Province, Sichuan Province, Chongqing, and Guangxi Province.
Data on postgraduate education institutions in China were summarized based on their economic zones and then imported into ArcGIS10.5 for spatial analysis. The spatial distribution of the 853 postgraduate education institutions is shown in Figure 1 and reveals a clear spatial agglomeration of postgraduate education institutions. Statistical analysis of detailed data indicates that postgraduate education institutions in the Northern Coastal Comprehensive Economic Zone account for 28% of the country. Postgraduate education institutions in the East Coastal Comprehensive Economic Zone ranked second, accounting for 14% of the national total. The Middle Reaches of the Yangtze River Comprehensive Economic Zone owns 14% of the total postgraduate education institutions in China, while the Southern Coastal Comprehensive Economic Zone has only 6% of the postgraduate education institutions in the nation.

Figure 1. Spatial pattern of postgraduate education institutions in the eight economic zones of China

Analysis at the Provincial Scale

The spatial pattern analysis results of postgraduate education institutions on the provincial scale are shown in Figure 2 and reveal a clear uneven spatial distribution of postgraduate education institutions in China. Quantitative analysis shows the top 10 provinces (provincial cities) in terms of the number of postgraduate education institutions: Beijing, Shanghai, Shaanxi, Hubei, Jiangsu, Liaoning, Shandong, and Sichuan, followed by Guangdong and Heilongjiang. The sum of the number of postgraduate education institutions in these 10 provinces accounts for 61% of the national proportion. Among them, Beijing has about 18% of all postgraduate education institutions in China. The spatial distribution of these institutions at the provincial scale indicates an overall increasing trend from west to east; at the same time, the distribution of the number of postgraduate education institutions increases from south to north.

For further analysis of postgraduate education and the related resource in China at the provincial scale, the authors explored the distribution of postgraduate students in schools and institutions in each province, and the result is shown in Figure 3. The authors found a similar spatial aggregation of postgraduate students in schools and institutions in Beijing, Hubei, and Jiangsu. However, compared with Figure 3, they found that provinces with a large number of postgraduate students in schools and institutions did not necessarily have a large number of postgraduate education institutions. For
example, Beijing, Hubei, and Jiangsu are the top three provinces regarding the number of postgraduate students; however, the provinces with the largest number of postgraduate education institutions are Beijing, Shanghai, and Hubei, respectively.

Figure 2. Spatial distribution of postgraduate education institutions at the provincial scale

![Spatial distribution of postgraduate education institutions at the provincial scale](image1)

Figure 3. Spatial distribution of postgraduate students in schools and institutions

![Spatial distribution of postgraduate students in schools and institutions](image2)
Analysis at the City Scale

Figure 4 illustrates the spatial pattern of postgraduate education institutions in China at the city scale. It reveals that postgraduate education institutions are gathered in many Midwestern cities. There are 365 cities (including autonomous prefectures and counties, not including Hong Kong, Macao, and Taiwan) in China, but only 150 cities have postgraduate education institutions. The top 10 cities for such institutions are Beijing, Shanghai, Wuhan, Xi’an, Tianjin, Harbin, Shenyang, Guangzhou, Chengdu, and Nanjing. Postgraduate education institutions in these cities account for 47.2% of all such institutions in China. Among the top 10 cities, the number of postgraduate education institutions in the four megalopolises of Beijing, Shanghai, Wuhan, and Xi’an is higher than that of other cities, accounting for more than 31% of the country’s total.

To explore a more detailed spatial aggregation pattern of postgraduate education institutions, the authors conducted a spatial aggregation analysis at the city scale using cluster and outlier analysis tools in ArcGIS. Figure 5 shows the overall geographic gathering profile of postgraduate education institutions at the city level, measured by the Anselin Local Moran’s I. In most cities in China, the aggregation characteristics of postgraduate education institutions are not significant. However, we find that Beijing, Tianjin, and Suzhou are “High–High” spatial aggregation areas. According to the index synthesis corresponding to ArcGIS spatial analysis report, it can be judged that the probability of spatial aggregation is greater than 95%. There are also many “Low–High” outliers such as Chengde, Zhangjiakou, and Jiaxing. All these findings suggest that a significant spatial autocorrelation exists between graduate education development and urban economic and social development in China.
FACTORS INFLUENCING THE SPATIAL PATTERN OF POSTGRADUATE EDUCATION IN CHINA

It is generally accepted that higher education institutions cluster in cities with a developed economy and clear regional advantages. To determine the specific influencing factors to the spatial pattern of postgraduate education at the city level, the two commonly utilized models, GWR and OLS, were selected to explore the potential ratio between socio-economic and technological development level and postgraduate education.

Due to data access restrictions, the authors selected GDP, GDP per capita \((PerGDP)\), population, disposable income per capita \((PerCapitaIncome)\), number of college students \((CollegeStudents)\), total retail sales of consumer goods \((TotalRetailSales)\), including junior college students and undergraduates, and number of patents and local income \((LocalFinRev)\) as indicators of economic, social, and technological development.

Analysis of Influencing Factors and Model Building

The authors used the number of postgraduate education institutions \((PostgraduateNum)\) as the dependent variable and \(LocalFinRev, PerGDP, Population, PerCapitaIncome, TotalRetailSales, CollegeStudents, Patents, \) and \(GDP\) as explanatory variables. According to the GWR modeling rules, the model can be established as follows:

\[
y_i = \beta_0 + \sum_{j=1}^{k} \beta_1(x_{ij} \cdot PerGDP) + \sum_{j=1}^{k} \beta_2(x_{ij} \cdot PerCapitaIncome) + \sum_{j=1}^{k} \beta_3(x_{ij} \cdot TotalRetailSales) + \sum_{j=1}^{k} \beta_4(x_{ij} \cdot CollegeStudents) + \sum_{j=1}^{k} \beta_5(x_{ij} \cdot Patents) + \sum_{j=1}^{k} \beta_6(x_{ij} \cdot LocalFinRev) + \varepsilon_i \ (i = 1, 2, 3, ..., n)
\]

Analysis of the Regression Results

Table 1 shows that \(R^2\) adjusted is greater than the strong correlation threshold of 0.7, which indicates the proposed model has good applicability and can explain 86.03% of the spatial pattern of postgraduate education institutions at the city level. In addition, the result implies a strong positive correlation of
postgraduate education with economic, social, and technological development factors. To identify the strong correlation variables, the authors selected the OLS model for multivariate stepwise regression analysis after eliminating the influence of multicollinearity. According to the fit effect, they selected the optimal model. The summary of the OLS result is displayed in Table 2.

Table 2 shows five variables (GDP, PerGDP, Population, TotalRetailSales, and LocalFinRev) that are highly correlated to PostgraduateNum, with the probability of statistical correlation being as high as 99%. This indicates that the established GWR model is highly significant, and the linear correlation between the selected variables is close. According to the regression coefficients and significance of the explanatory variables, postgraduate education institutions (PostgraduateNum) are significantly affected by GDP, GDP per capita, city population, total retail sales of consumer goods, and local income.

Table 1. GWR analysis results using the ArcGIS spatial relationships tool

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>12.70356798</td>
<td></td>
</tr>
<tr>
<td>ResidualSquares</td>
<td>7297.013385</td>
<td></td>
</tr>
<tr>
<td>EffectiveNumber</td>
<td>25.44964682</td>
<td></td>
</tr>
<tr>
<td>Sigma</td>
<td>4.778624666</td>
<td></td>
</tr>
<tr>
<td>AICc</td>
<td>2076.55963</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.877336971</td>
<td></td>
</tr>
<tr>
<td>* PostgraduateNum</td>
<td>0.860300431</td>
<td></td>
</tr>
<tr>
<td>** GDP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** LocalFinRev</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** TotalRetailSales</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** PerCapitaIncome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** Population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** CollegeStudents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** Patent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>** PerGDP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * stands for dependent field and ** stands for explanatory field

The relationship between PerCapitaIncome and PostgraduateNum did not present the expected results. On the contrary, it showed a significant negative correlation, indicating that a high level of disposable income per capita cannot improve postgraduate education. This may be related to the fact that funds for postgraduate education are generally allocated by the Ministry of Education or provincial education departments.

Table 2 also implies the number of college students and the number of postgraduate education institutions is not necessarily related. One possible explanation is that the scope of postgraduate student recruitment is nationwide, but undergraduate student enrollment in one city may be relatively low. Similarly, the relevancy between the number of patent applications and the number of postgraduate education institutions is not significant. This may be because enterprises also apply for many patents in addition to postgraduate education institutions.
Regarding the coefficient of determination, $R^2$, in OLS, the linear fit effect of each equation is good, and the explanation degree of the dependent variable is as high as 87% (Table 3). To examine residuals of the established GWR model, the histogram of the residuals is generated. The authors found it very close to the normal distribution curve in most areas (Figure 6). These results illustrate the overall good fit of the spatial distribution model of postgraduate education institutions.

Table 2. Summary of OLS results

<table>
<thead>
<tr>
<th>Coefficient[a]</th>
<th>StdError</th>
<th>t-Statistic</th>
<th>Probability [b]</th>
<th>Robust_SE</th>
<th>Robust_t</th>
<th>Robust_Pr [b]</th>
<th>VIF [c]</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.338346</td>
<td>3.99571</td>
<td>1.836524</td>
<td>0.068331</td>
<td>3.981332</td>
<td>1.843189</td>
<td>0.067346</td>
<td>--------</td>
</tr>
<tr>
<td>0.002799</td>
<td>0.000663</td>
<td>4.219401</td>
<td>0.000047*</td>
<td>0.001351</td>
<td>2.071809</td>
<td>0.040049*</td>
<td>19.79087</td>
</tr>
<tr>
<td>0.000107</td>
<td>0.000032</td>
<td>3.309298</td>
<td>0.00189*</td>
<td>0.000062</td>
<td>1.74131</td>
<td>0.083756*</td>
<td>3.85077</td>
</tr>
<tr>
<td>0.009137</td>
<td>0.004344</td>
<td>2.103203</td>
<td>0.037168*</td>
<td>0.006794</td>
<td>1.344764</td>
<td>0.018080*</td>
<td>4.925188</td>
</tr>
<tr>
<td>-0.000062</td>
<td>0.000136</td>
<td>-0.45906</td>
<td>0.64889</td>
<td>0.000078</td>
<td>-0.79677</td>
<td>0.442688</td>
<td>3.204052</td>
</tr>
<tr>
<td>0.000703</td>
<td>0.000208</td>
<td>3.381606</td>
<td>0.000936*</td>
<td>0.000697</td>
<td>1.008645</td>
<td>0.031481*</td>
<td>1.491052</td>
</tr>
<tr>
<td>0.012745</td>
<td>0.015697</td>
<td>0.81193</td>
<td>0.418154</td>
<td>0.016622</td>
<td>0.766726</td>
<td>0.44484</td>
<td>1.182913</td>
</tr>
<tr>
<td>1.08644</td>
<td>0.550848</td>
<td>1.973022</td>
<td>0.050392</td>
<td>1.061798</td>
<td>1.023208</td>
<td>0.030790*</td>
<td>6.92917</td>
</tr>
<tr>
<td>0.005373</td>
<td>0.002215</td>
<td>2.42556</td>
<td>0.016505*</td>
<td>0.00426</td>
<td>1.261342</td>
<td>0.0209214*</td>
<td>10.37616</td>
</tr>
</tbody>
</table>

Note: [a] Used to represent type (positive correlation or negative correlation) and intensity of the relationship between the explanatory variable and selected potential dependent variables. [b] * used to imply the variable with statistically significant, general with (p < 0.01). [c] VIF is short for the Variance Inflation Factor.

Figure 6. Histogram of standardized residuals with OLS analysis
CONCLUSION AND FUTURE WORK

Previous studies have shown that the imbalanced development of higher education has long been the focus of the government and related scholars. The spatial distribution of postgraduate education results from multiple factors, and its spatial distribution pattern is the local embodiment of China’s economic and information technology development. This paper presents three main findings.

First, the spatial analysis revealed that the geographic distribution of postgraduate education varies significantly in China. From analysis results at the comprehensive economic zone scale, the distribution of postgraduate education institutions had a noticeable step-like trend from the eastern coastal area to the western inland. In contrast, the difference from north to south was not significant. At the provincial level, postgraduate education institutions were grouped in the central and eastern provinces. A more detailed analysis at the city scale showed that postgraduate education institutions were mainly located in provincial capitals or first-tier megacities, such as Beijing, Shanghai, Wuhan, and Chengdu.

Second, the spatial aggregation pattern of postgraduate education was evident. Analysis at the provincial scale indicated differences in the specific spatial aggregation pattern of postgraduate education institutions and in the spatial distribution of postgraduate students in schools and institutions. The uneven spatial distribution of postgraduate education in cities was also prominent. Only 150 of the 365 cities in China have postgraduate education institutions, and Beijing has about 18% of all such institutions and 14.2% of all postgraduate students in schools and institutions nationally. In addition, the city-level cluster and outlier analysis of postgraduate education institutions revealed “High–High” spatial aggregation areas in Beijing, Tianjin, and Suzhou. There were also many “Low–High” outliers such as Zhangjiakou and Jiaxing.

Third, the spatial regression analysis at the city scale showed that the level of development of postgraduate education was significantly influenced by five indicators related to economic and social development. Specifically, the indicators included per capita GDP, local income, population, GDP, and total retail sales of consumer goods. Meanwhile, the other three indicators: the number of college students, disposable income per capita, and the number of patents, had no significant impact on the level of development of postgraduate education.

Quantitative analysis of the study indicated that many economic and social development-related factors influence the spatial pattern of postgraduate education in China. The government should focus on the balanced development of higher education and increase the support for the development of postgraduate education in underdeveloped areas. To deal with the current complex economic environment, it is necessary to maintain a stable and healthy economic development environment and drive the development of higher education with high-level economic development.

Table 3. OLS diagnostics report

<table>
<thead>
<tr>
<th>Dependent Variable: PostgraduateNum</th>
<th>Akaike’s Information Criterion (AICc) [d]: 1165.209934</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations: 154</td>
<td>Adjusted R-Squared [d]: 0.870664</td>
</tr>
<tr>
<td>Multiple R-Squared [d]: 0.0898341</td>
<td>Joint F-Statistic [e]: 18.005138 Prob(&gt;F): (8,145) degrees freedom: 0.000000*</td>
</tr>
<tr>
<td>Joint Wald Statistic [e]: 37.0433093 Prob(&gt; chi-squared): (8) degrees of freedom: 0.000011*</td>
<td></td>
</tr>
<tr>
<td>Koenker (BP) Statistic [f]: 58.922831 Prob(&gt; chi-squared): (8) degrees of freedom: 0.000000*</td>
<td></td>
</tr>
<tr>
<td>Jarque-Bera Statistic [g]: 13695.73958 Prob(&gt; chi-squared): (2) degrees of freedom: 0.000000*</td>
<td></td>
</tr>
</tbody>
</table>

Note: * used to imply the variable with statistically significant, general with (p < 0.01). [d] Selected by ArcGIS to measure the fit/performance of the OLS model. [e] Selected by ArcGIS to measure the overall significance of the OLS model, with * indicating an overall significant result. [f] Selected by ArcGIS to measure the consistency of the analyzed model. [g] Selected by ArcGIS to measure the accuracy of the model.
postgraduate education institutions need to increase cooperation, complement each other’s advantages, and share high-quality education resources to enhance the comprehensive strength of China’s higher education.

The findings of this study have important reference significance for optimizing the spatial distribution pattern of higher education, narrowing the spatial gap, and promoting the regional coordinated development policy of higher education. However, this study had several limitations. The authors collected publicly available data from 2018 to study the spatial development characteristics and explore its possible impact factors. The development of postgraduate education is a dynamic process; a continuous and long time span data study may produce different findings. In the future, the authors will focus on collecting more data to conduct multiple-year-based studies and compare the factors influencing postgraduate education in China with other related research. Besides, many countries have been undergoing digital transformation supported by emerging technologies. These digital technologies include machine learning (Mirsadeghi et al., 2021), deep learning (Sahoo & Gupta, 2020), big data analytics (Sultana et al., 2021; Zulkefly et al., 2021), and the Internet of Things (Mishra et al., 2021; Peng, Clough et al., 2021), which can help analyze spatial patterns and development characteristics.

**FUNDING AGENCY**

This work is supported by the Key Project of Ningbo Education Science Planning in 2022 (Grant Number 2022YZD013) and the Scientific Research Startup Fund of Ningbo University of Technology in 2022.

**CONFLICTS OF INTEREST**

The authors declare that they have no conflicts of interest to report regarding the present study.
REFERENCES


ENDNOTES

1. http://www.moe.gov.cn/s78/A03/moe_560/jytsj_201
2. https://lbsyun.baidu.com/
5. https://lbsyun.baidu.com/