International Journal of Strategic Property Management
ISSN: 1648-715X / eISSN: 1648-9179
2023 Volume 27 Issue 1: 50–63
https://doi.org/10.3846/ijspm.2023.18708

RESIDENTIAL SPATIAL DIFFERENTIATION AT THE COMMUNITY LEVEL IN HANGZHOU, CHINA

Ling ZHANG *, Baihong WU

College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China

Received 07 August 2022; accepted 26 February 2023

Abstract. Residential spatial differentiation is an important social issue for every large city to explore. The "subcultural" clusters formed by the subjective choices of a city's residents divide the living spaces of the city and create residential differentiation. A study of this phenomenon would help to facilitate the allocation of urban resources and the formulation of relevant development strategies. In recent years, Hangzhou, one of the major developed cities in China, has seen a gradual increase in the proportion of its non-local population. Employing principal component analysis and K-means clustering, this study investigated the city's spatial differentiation of living spaces at the community level and identified the social clustering areas. The Shannon-Wiener index and the index of dissimilarity were used to examine the degree of residential segregation in each administrative district. The results showed that the functional distinctions among the districts of Hangzhou were pronounced and had certain degrees of class solidification. A certain spatial coupling was observed between residential differentiation and housing prices. More reasonable allocations of social resources would help to ameliorate the negative effects caused by the spatial differentiation of living spaces.

Keywords: residential spatial differentiation, principal component analysis, K-means cluster, Shannon-Wiener index, index of dissimilarity.

Introduction

The phenomenon of residential spatial differentiation refers to the fact that residents with particular characteristics live in specific spatial areas of a city, which contains residential differentiation or even mutual segregation. Within relatively isolated areas, homogeneous populations share similar social characteristics, follow common customs and commonly accepted values, or maintain the same subculture. Many scholars have provided answers to the question of what defines a living space. Waddell and Borning (2004) thought of living spaces as resources on which residents depend for their survival and that results from the fact that people who settle in a particular residential space would maximize the utility of the activity space, which is a phenomenon based on location selection (Bhat & Guo, 2004). In turn, this phenomenon leads to the fact that residents with the same characteristics choose the area that best meets their living needs. Modai-Snir and Plaut (2019) indicated that spatial differentiation of urban living is an inevitable result of urbanization in market economic conditions and that the same living background and habits would increase the sense of the identity of the residents.

Today's urbanization trends in China are also gradually leaning toward Western models (Forrest, 2012; Zhao, 2013), in which high degrees of development lead to high degrees of socio-spatial segregation. The main driver of such current segregation in China is also the result of the unequal distribution of resources during the country's transformation into a market economy (Wu, 2002). As a typical economically developed historical city in eastern China, Hangzhou has experienced large-scale urban expansion and urban renewal. Its economy has perennially ranked among the highest in the country in terms of land finance and land dependence (Qian, 2015). Around 2005, Hangzhou gradually transformed its economic focus to new sectors, such as finance and the Internet, information software, e-commerce, and the Internet of Things, which are now leading the economy.

Market economies have been implemented in the major cities of China since the start of economic reforms and liberalization. Among these cities, Hangzhou has been given better development opportunities (Wei & Li, 2002).

*Corresponding author. E-mail: *zlcivil@163.com*

As its strategic position in the country's development has been rising in recent years, its development has gradually accelerated while its resources and scale have expanded. According to the censuses, the total population of the city in 2015 reached 9.018 million, which is an increase of 36% over that in 2010, and it even reached 12.204 million in 2021. The non-local population growth rates from 2010 to 2015 and from 2015 to 2020 were 1.50% and 109.91%, respectively. The abundant resources and recruitment policies have attracted a large number of non-local university graduates to settle and work in this city. Residents with increasingly diverse educational and economic backgrounds, as well as nationalities, gather in particular areas to form diverse social clusters. This pattern creates segregation and class solidification in certain areas in the long run, which may give rise to future social problems (Frankenberg, 2013; Tammaru et al., 2016). Therefore, it is particularly important to accurately identify the distributions of these residential "social areas" for policy making, market analysis, and amenity improvement.

The rapid development of China's urban real estate market in the last two decades has led to exponential growth in housing development and housing prices, which have resulted in the expansion of housing market stock and rapid spatial expansion. The marketization of housing has broken the housing system previously based on work units, providing a way for residents to replace and improve their homes. The built-up area of Hangzhou grew from 412.5 km^2 in 2010 to 706 km^2 in 2022, corresponding to an expansion of 1.7 times. As such, the spatial gradient of housing prices is evident, and price has become an important consideration in residents' housing choices. New classes, such as local homeowners who are compensated for demolitions, homeowners of greatly appreciated school district housing, and newly arrived young university graduates, have emerged and the spatial differentiation patterns of living spaces have been constantly changing. Organized by the Hangzhou government and transportation department, a large-scale household survey in 2015 covered household demographic characteristics, housing status, and transportation. The data provide rare support for the study of residential differentiation. This study used the statistical data of this survey at the level of the smallest inner-city unit (i.e., the community) to reveal the characteristics of the social dimension of residential space and the pattern of differentiation, with the aim to analyze the diversity and the degree of differentiation of social categories at a higher level, (i.e., administrative districts), and to study the association with house prices in the formation of spatial differentiation. By adopting the ecological factor analysis method, this study fully considered the contextual factors of Chinese cities in the process of social dimension analysis, diluted the ethnic status, highlighted the effects of household registration, education, occupation, and housing situation, and constructed a multivariate principal component analysis (PCA) method to reveal the spatial differentiation characteristics of modern Chinese cities, providing diverging empirical evidence from other related studies. As one of the few studies that focused on the interaction between housing prices and location status (commuting time, etc.) and spatial differentiation patterns, this study provided lessons for other regions where the real estate market is still in its expansion stages. This in-depth analysis of Hangzhou's residential spatial differentiation not only explained the city's residential spatial patterns and socio-economic-demographic characteristics, but also provided some policy recommendations for urban development and planning, as well as some experiences for the development pattern of the real estate market in other rapidly urbanizing cities.

The structure of the remainder of this paper is as follows. The next section is a summary of the current research. The following section introduces the methodology and data. The section after that presents a community-level analysis of residential clusters. The penultimate section discusses residential diversity at the district level. The final section presents the conclusion.

1. Literature review

The theoretical basis of this study is the theory of socio-spatial dialectic. The structure of spatial organization does not independently construct and transform itself, nor is it a simple expression of the class structure that originates from the social relations of production. On the contrary, it represents the dialectical unity component of the general relations of production, which are simultaneously social and spatial (Soja, 1980). Although space itself may be initially given, the organization, use, and value of space are the products of social translation, transformation, and experience. Spatial homogeneity exists within the urban built environment in unison with the social class structure and its conflicts and transformations. This homogeneity provides us with theoretical clues to study urban residential differentiation (Wu, 2016). Social area studies were first proposed by Shevky and Williams (1949) and elaborated by Shevky and Bell (1955), who formed the prototype of urban area studies by dividing cities into different social areas according to three dimensions: economic status, family status, and ethnic status. Subsequent studies have adopted many measurement methods and geographers, such as Rees (1970) and Knox (1987), have developed a greater interest in research models about spatial dimensions. They thought that the factor ecological research method of social area analysis could help people understand intra-urban residential differentiation. This method has gradually become an important tool for studying the spatial structure of a city. The existence of these three dimensions was confirmed by studies in several other global cities, including Melbourne, Toronto, and Chicago (Murdie, 1969; Rees, 1970); moreover, this method was successfully applied in several non-Western urban settings (Johnston et al., 2001; Smith & Scarpaci, 2000). Studies on social area analysis have been successful in explaining the spatial variations in urban structure. Many dimensions produce the phenomenon of residential spatial differentiation in cities besides economic development. For example, the segregation of African Americans in the United States is attributed to economic reasons and racial factors (Williams & Collins, 2001). However, the situation in China is not the same because a large portion of economic development has been concentrated in the eastern regions, where the distribution of ethnic minorities is low. Li and Wu (2008), Liu et al. (2014), and Wang (2017) found that the main individual characteristic factors that have led to segregation in China were social status, residency rights, migration status, income levels, and socialist urban policies. Wu et al. (2020) and Xing et al. (2022) analyzed residential spaces in Chinese cities and argued that social policies regarding household registration and education for the mobile population should guide their orderly integration into the urban society.

Previous studies noted that higher socioeconomic households tend to live in neighborhoods with higher housing values (Baer & Williamson, 1988). Some scholars have also argued for the establishment of a link between housing values and other non-residential parameters in housing choice, such as demographics (Benefield, 2009; Shaw, 2002). Reed (2013) formally considered social area analysis and house price changes together. This study identified social dimension factors from a number of demographic variables through PCA, and found that house price divergence in Melbourne suburban areas could be explained by economic and demographic variables. Feng and Han (2021) stated that heterogeneity in housing prices would lead to the heterogeneity in residential space segregation, which would, in turn, intensify uncoordinated regional development. Studying the spatial heterogeneity of residents would be beneficial not only to coordinated regional development but also to the healthy and rational development of the real estate market. Chhetri et al. (2009) explored the Australian housing market and found that the boom in the real estate market had created segregation for individual housing and location choices in Australia's metropolitan areas. Research on urban spatial fragmentation patterns would help understand the various factors, such as personal consumption preferences, variety of local government planning schemes, job accessibility, quality education, council fees, and taxation, of the housing market.

Inequality in income growth varies considerably across countries, depending mainly on the composition of economic sectors. The transition from secondary to tertiary industries' dominance in cities seems to be increasingly accompanied by a corresponding increase in the level of urban social polarization, a reduction of homogeneity among urban social classes, and an increasing structural complexity (Badcock, 1997). In the era of globalization, the demand for unskilled labor has decreased considerably, while the income from new jobs created by producer services and knowledge industries has begun to rise consistently (Reich, 1991). Fikire (2021) and Cox and Hurtubia (2021) have expressed a consistent view of households with different incomes that choose their housing locations according to their income levels. They also believed that different classes had the propensity to choose such housing that would contribute to residential segregation.

The majority of literature investigated residential segregation at the smallest spatial statistical unit scale. An example is Tammaru et al.'s (2016) study of Sweden based on the smallest residential neighborhood SAMS, with an average population of about 1,000 inhabitants. Li and Wu (2008) and Zhang et al. (2022) conducted a sub-district level study, which is equivalent to the town level in Western countries, of this problem in Chinese cities. Li et al. (2010) and Wu et al. (2014), whose study is based on data from an earlier year, mainly used sub-district-level data from first-tier cities and a small portion of data from some towns. The study ultimately showed that better empirical results would be obtained with higher-quality data from smaller geographic units and that studies of large Chinese cities should use data with clearer boundaries.

The diversity index and the dissimilarity index are commonly used to measure the degree of residential differentiation. The diversity index is more suitable to describe the even distribution of social groups, while the dissimilarity index is suitable for variables with significant regional differences, such as housing tenure in the study of residential clusters in Nanjing (Wu et al., 2014). Allen et al. (2015) and Šimon et al. (2020) indicated that the index of dissimilarity is the most reliable measure to quantify the phenomenon of residential differentiation. Harris and Owen (2018); Preston and Ray (2020) and Křížková and Šimon (2022) used the index of dissimilarity to analyze racial segregation in several cities. Their research was based on the fact that there was already a clear racial division in the city, such as that between black people (i.e., the marginal group) and white people (i.e., the dominant group). However, in the case where multiple social groups coexist and the type of inhabitants is not well defined, the calculation of the dissimilarity index becomes quite difficult. To calculate this index, Křížková and Šimon (2022) set a threshold for the proportion of population of {10%, 20%, ..., 100%}, so as to calculate the proportion of those groups that were beyond each threshold in each spatial unit. In a study of housing tenure differences and residential segregation in Shanghai, Li and Wu (2008) used a similar approach to measure the index of dissimilarity, and found that the degree of the residential concentration of each cluster in Shanghai is significant. They concluded that there was no evidence of high levels of segregation among social groups to the extent that exists in Western countries.

In synthesis, social area research based on the ecological factor analysis method needs to keep up with social development and the changes brought by the globalization and economic transformation in large cities. The dimensions of economic status and family status need to be considered in a way that is more in line with local characteristics. In fact, the main factors affecting socio-spatial differentiation vary across different stages of urbanization, and with different degrees of urban development and maturity of real estate markets. At the same time, since the formation of socio-spatial differentiation patterns is the result of the aggregation of the housing choices of countless households, the city's residential history, housing market development, housing policies, and policies on migration influence the location of households, and can eventually affect the residential differentiation. Residential differentiation is an enduring topic that still needs vivid case studies covering different urbanization stages and different socio-cultural contexts, which would enrich the theory of socio-spatial unity, provide empirical evidence for regions in similar development stages, and offer policy recommendations for regional urban planning and housing development.

2. Methods and data

2.1. Data

The scope of this study covers 10 districts under the jurisdiction of Hangzhou: Shangcheng, Gongshu, Xihu, Binjiang, Xiaoshan, Yuhang, Linping, Qiantang, Fuyang, and Lin'an. The data were obtained from a household questionnaire based on a sample of Hangzhou traffic management departments in 2015. The survey data included relevant community-level aggregated data such as family size, age, education, household registration, occupation, income, and travel times (to work, school, shopping, and entertainment) of the respondents, which were collected from 35,000 households in 629 relevant communities. The summary data at the community level are shown in Table 1. House price is one of the most important factors influencing residents' housing choices. To study how they affect the residential spatial distribution, this study additionally collected the transaction prices of housing at the community level in 2015, which were retrieved from the second-hand house prices data recorded by XiTai.com for each year in the city. We took the average of the transaction prices of the communities for which transactions are recorded in a specific year, and used them as house price data of each community. The statistics of house prices are shown at the bottom of Table 1.

Table 1. Descriptive statistics of variables

	Average	Median	SD	Min.	Max.
Family size	2.85	2.77	0.55	2.00	5.00
Housing space per capita (m^2)	35.59	32.29	14.91	18.00	159.00
Housing area (m^2)	101.67	92.48	49.94	46.00	396.00
Number of cars per household	0.72	0.70	0.34	0.00	2.00
Population ratio 6–19 years old	8.30%	7.88%	4.54%	0.00%	32.47%
Population ratio 20–29 years old	14.47%	12.50%	9.44%	0.00%	58.82%
Population ratio 30–59 years old	55.22%	56.68%	10.27%	7.02%	86.21%
Population ratio over 60 years old	22.01%	18.91%	14.00%	0.00%	75.00%
Proportion of population with university or higher education	23.54%	21.06%	14.35%	0.00%	83.08%
The proportion of household registration in local	85.56%	91.25%	15.88%	9.00%	100.00%
The proportion of household registration ratio	90.40%	94.88%	12.47%	18.00%	100.00%
Proportion of population work as head of state institutions and enterprises	3.35%	2.04%	4.10%	0.00%	30.86%
Proportion of professional and technical staff	10.14%	9.03%	7.11%	0.00%	47.18%
Proportion of clerical and related staff	15.84%	14.71%	9.43%	0.00%	69.23%
Proportion of commercial and service workers	15.33%	13.79%	10.30%	0.00%	84.72%
Proportion of personnel in agricultural, forestry, animal husbandry and fishery	0.25%	0.00%	1.08%	0.00%	18.00%
Proportion of operators of production and transportation equipment	2.22%	0.90%	3.89%	0.00%	33.00%
Proportion of other employees	52.87%	51.82%	15.77%	0.00%	98.20%
Proportion of respondents with monthly income below 5k	65.59%	67.87%	15.27%	10.39%	96.15%
Proportion of respondents with monthly income between 5k–10k	17.94%	16.42%	11.57%	0.00%	63.89%
Proportion of respondents with monthly income between 10k–15k	1.57%	0.00%	2.79%	0.00%	32.76%
Proportion of respondents with monthly income above 15k	0.82%	0.00%	2.10%	0.00%	17.00%
Proportion of monthly household income below 10k	53.71%	54.55%	19.18%	0.00%	100.00%

End of Table 1

	Average	Median	SD	Min.	Max.
Proportion of monthly household income between 10k–20k	38.56%	38.89%	14.63%	0.00%	81.82%
Proportion of monthly household income between 20k–30k	6.55%	4.23%	7.73%	0.00%	52.00%
Proportion of monthly household income between 30k–40k	1.03%	0.00%	2.74%	0.00%	24.00%
Proportion of monthly household income above 40k	0.16%	0.00%	0.85%	0.00%	12.00%
Average travel time to work	32.22	31.00	9.67	12.00	79.00
Average travel time to school	24.99	21.67	16.11	3.00	230.00
Average travel time to shopping	21.94	20.13	9.97	5.00	105.00
Average travel time to entertainment	32.88	30.00	17.59	5.00	162.00
Average travel time to home	30.49	29.58	8.65	12.00	70.00
Average travel time for other purposes	31.8	28.17	17.55	3.00	150.00
House price of community (¥/m^2)	17562.40	16750.00	7089.30	3017.74	45030.00

Note: 1 EUR is equal to 7.3718 CNY (Chinese Yuan Renminbi).

2.2. Cluster analysis, diversity analysis and dissimilarity analysis

Factor-based social area research methods have been proposed by previous scholars and have been widely applied gradually. PCA identifies the principal factors after effectively identifying the interrelationships between multiple variable factors, thus achieving dimensionality reduction of the variables. The results of this analysis were used to define and distinguish different housing markets (Chhetri et al., 2009; Wu et al., 2014; Tomal, 2021). The use of PCA and cluster analysis was found as superior for studying social differentiation and spatial mixing in Chinese cities (Wu et al., 2017). Based on multiple variables such as population, education, occupation, and housing, this type of analytical approach enables the identification, classification, and analysis of the spatial separation between different social groups.

Following previous studies, especially Wu et al. (2014, 2017), we collected household and socioeconomic variables for the residential submarket. These variables were designed for PCA. The household variables were family size, age and household registration while the socioeconomic variables were education, income, and occupation. After a detailed transportation survey, we also incorporated individual travel variables such as travel times to schools, shopping centers, and recreational facilities. Cluster analysis can accomplish a further analysis of the spatial differentiation of social residences after the identification of the resident characteristic factors.

The peculiarities of urban zoning in China have led to more possibilities for studying spatial differentiation and residential segregation. Compared to the former, the latter emphasizes the concepts of unevenness, exposure, and concentration (Liu et al., 2018; Massey, 2012). The intra-city area is divided into multiple administrative districts, so it is feasible to further study the distribution of community diversity in each district after the cluster analysis.

In this study, the Shannon-Wiener index was used to measure diversity and further analyze the residential differentiation pattern, as follows:

$$H = -\sum_{i=1}^{R} S_i \ln S_i, \tag{1}$$

where: H is the Shannon-Wiener Index; S_i represents the proportion of the i-th social group in the bounded social region; and R is the total number of social groups in that social region.

Moreover, the cumulative proportion method was used to calculate the dissimilarity index. The degree of residential segregation was examined by counting the percentage of districts where each social group exceeds a certain threshold ratio.

3. Community-level residential differentiation in Hangzhou

The 629 community data points were subjected to PCA and rotated 15 times by varimax rotation to achieve convergence and extract 7 principal components. The KMO value reached 0.689, indicating that the data were suitable for use in factor analysis. The variance contribution of each principal component was relatively average and the final variance explained 64.425%. According to the rotated component matrix (see Table A1 in Appendix), the correlation of each variable in the principal components can be judged. To assign the variables to each factor, the factor discriminant criterion can be formulated based on the absolute value of factor loading (Dolnicar & Grün, 2008; Chhetri et al., 2009). Chhetri et al. (2009) set the criterion at 0.40 to maximize the specificity of their principal components. Sarstedt and Mooi (2019) suggested that if the number of factors to be extracted is very low, the loading should be at least 0.50, while if the number is high, a loading of above 0.3 is acceptable. In this study, the loading level of 0.5 was used. These related variables reflect the structure of the residential

Table 2. Characteristics of principal components and the spatial distribution

Variance contribution	Component name	Component characteristics	Main spatial distribution
Principal Component I 12.51%	Upper middle-income factor	Proportion of monthly income above 15k (0.86), Proportion of monthly household income between 30k–40k (0.833), Proportion of monthly household income between 20k–30k (0.583), Proportion of monthly income between 10k–15k (0.787)	Employment center (Huanglong, near Yuhang Science and Technology Park); Wulin CBD; Qianjiang New CBD
Principal Component II 10.71%	Long-distance travel factor	Average time to travel home (0.838), to entertainment areas (0.684), to work (0.671), and other activities (0.666)	More remote areas on the outskirts of cities
Principal Component III 9.34%	Non-local resident young business service workers factor	Proportion of commercial and service workers (0.675), Population ratio 20–29 years old (0.647), Proportion of other employees (–0.835), Proportion of household registration ratio (–0.304), Proportion of household registration in local (–0.28)	The concentration areas are not very clearly distributed in all areas of the city
Principal Component IV 8.99%	Low- and middle-income population with higher education factor	Proportion of monthly household income between 10k–20k (0.803), Proportion of monthly income between 5k–10k (0.798), Population with university or higher education (0.584)	Clustering in higher education resource areas within Hangzhou
Principal Component V 8.23%	Middle-aged and elderly local residents factor	Proportion of household registration ratio (0.856), Registration in local (0.84), Population ratio 6–19 years old (–0.116), Population ratio 20–29 years old (–0.302)	There is a relatively even distribution in the inner city of Hangzhou
Principal Component VI 7.69%	Family with school-age children factor	Population ratio 6–19 years old (0.719), Population ratio 30–59 years old (0.729), Number of cars per household (0.529)	Mainly located in the inner city and suburban areas around university towns and famous secondary schools
Principal Component VII 6.96%	Primary and secondary industry employees factor	Proportion of personnel in agriculture, forestry, animal husbandry, and fishery (0.689), Proportion of operators of production and transportation equipment (0.675), Population ratio 20–29 years old (0.327)	Urban–rural fringes, including agricultural areas and fisheries, industrial, science, and technology parks, and other areas

Note: CBD means "central business district". According to the industry classification standard of the National Bureau of Statistics of China, agriculture, forestry, animal husbandry, and fishery belong to the primary industry, and production and transportation equipment belong to the secondary industry.

components of the society. The characteristics and spatial distribution of each principal component are summarized in Table 2. The spatial distribution of the factor scores for each principal component is shown in Appendix Figure A1.

Many studies have used cluster analysis (Tu et al., 2022). Yang et al. (2019) stated that K-means clustering was a type of unsupervised learning that generally used Euclidean distances as a measure of similarity between data objects. There are the advantages of simplicity, satisfactory results, and easy implementation with this type of method. However, the number of clusters K in the K-Means algorithm needs to be determined in advance. If the chosen K-value is too small or too large in the empirical analysis, then the results will deviate from reality. In this study, K-means clustering was used for the cluster analysis. To establish the relationships among the community house prices more intuitively, the house price factor was added as the eighth principal component[1]. The Euclidean distances were used as the measure of similarity and 6–9 classes were tested. After the outliers were processed, the best choice for the K-value was 8 to obtain an even distribution of the number of communities in each cluster and to conduct a social cluster analysis more conveniently at a later stage.

The calculation of a more intuitive discriminant coefficient helped to discern the main characteristics of each cluster. Referring to Shoemaker (1994), Wu et al. (2014) and Wu (2016), we calculated the correlation coefficients (γ_{ij}) between the principal components of the clusters. Also, we used the squared mean of the correlation coefficients ($\overline{R}_L^2$) to determine the dominant characteristics of each cluster. The functional form is:

$$\overline{R}_L^2 = \frac{\sum_{i \neq j} \gamma_{ij}^2}{k}, \tag{2}$$

where γ_{ij} represents the correlation coefficient calculated from the sequence of factor scores between principal components i to j. These scores are given by each community in the clusters. k represents the number of correlation coefficients obtained for any principal component. The clustering components and score discriminant data are shown in Table 3. The largest value in each row means that in that cluster, the summed correlation between that factor and the other factors is the strongest. The two factors with the

[1] Referring to Sarstedt and Mooi (2019), Dolnicar and Grün (2008) and Shoemaker (1994), since factor-cluster analysis may discard some variables with low loadings in the extracted factors, which may have important information for identifying niche categories, we added additional house prices to the clustering analysis. The high F-ratio not presented here also shows that it can differentiate clusters well.

Table 3. Clustering components and score discriminations ($\overline{R}_L^2$)

Clusters	Number of communities	Clustering variables							
		PC I	PC II	PC III	PC IV	PC V	PC VI	PC VII	House price
CL 1	74	0.006	0.011	0.010	0.026	0.015	<u>0.044</u>	<u>0.050</u>	0.040
CL 2	24	0.068	0.037	0.105	<u>0.106</u>	0.067	<u>0.126</u>	0.042	0.048
CL 3	50	0.045	0.045	0.054	0.019	0.029	<u>0.097</u>	<u>0.059</u>	0.039
CL 4	72	0.021	0.039	0.025	0.026	<u>0.056</u>	0.016	<u>0.046</u>	0.035
CL 5	118	0.021	0.027	0.010	<u>0.040</u>	0.028	0.011	<u>0.038</u>	0.037
CL 6	128	0.035	0.006	<u>0.055</u>	0.021	<u>0.064</u>	0.004	0.040	0.012
CL 7	40	0.006	0.041	<u>0.043</u>	0.037	0.022	0.040	<u>0.045</u>	0.020
CL 8	21	<u>0.098</u>	0.038	0.076	0.114	0.085	0.063	0.065	<u>0.096</u>

largest coefficients of determination were used to characterize each cluster. The spatial distribution of residences in each cluster is shown in Figure 1.

Cluster 1: The volume of communities in this cluster accounts for 14%. The characteristics of the population are the families with school-age children (Component VI), and the primary and secondary industry employees (Component VII). According to the geographical analysis in Figure 1, the primary and secondary industry employees mainly live in the suburbs. Others are mostly located in Xihu District, which is rich in educational resources. In Xihu District, 144 public and private primary and secondary schools, as well as kindergartens, are equipped with high-standard facilities and equipment. Of the primary and secondary schools, 97.6% are covered by provincial compulsory education standards. Famous universities are also located in this district. The median level of community housing prices in the communities located in the Xihu District is rather high at about ¥23k per square meter (see the last column of Table 4).

Cluster 2: The volume of such communities is 4.5%. The main characteristics of this clustered population are families with school-age children (Component VI), low- and middle-income population with higher education (Component IV), and the non-local residents young business service workers (Component III). Most people among this group live far away from the inner city in suburban areas that are yet to be developed. The median house price of the residential community is the lowest of all the clusters.

Cluster 3: Clustering of families with school-age children. The main feature exhibited by such a clustered population is the number of communities having families with school-age children (Component VI) accounts for 9.4%. Residents are distributed in Xihu and Gongshu Districts, which are quality middle-school districts and have universities. They live in the peripheral areas of these two districts around the inner-city area and the price of housing is at a medium level.

Cluster 4: These communities account for 13.5% of the total number of people. The main characteristics of these communities are the middle-aged and elderly local residents (Component V), and the primary and secondary industry employees (Component VII). They live in the outer suburbs of the city and the urban-rural fringe areas, where there are more primary industries and industrial parks, in which "old Hangzhou people" may be working in related industries.

Cluster 5: The volume of such communities is 22.2%. The main resident characteristics of this cluster are low- and middle-income residents with higher education (Component IV), primary and secondary industry employees (Component VII), and house price. The range of settlements is concentrated in the central business district (CBD). The median housing price in the communities is around ¥23k–24k per square meter, which is the highest level in the city.

Cluster 6: The volume of such communities accounts for 24.1%, which is the largest community share volume. The main explanatory factors are the middle-aged and elderly local residents (Component V) and the non-local resident young business service workers (Component III). These people live in communities where the housing prices are at relatively low levels. As can be seen in Figure 1, they live in areas far from the inner city. In addition to some local residents who are native to these townships, non-local workers also choose to live in these areas for the lower housing prices and rents.

Cluster 7: The volume of this community is 7.5%. The main characteristics of this group are primary and secondary industry employees (Component VII), non-local resident young business service workers (Component III), long-distance commuting workers (Component II), and families with school-age children (Component VI). These people live far away from the inner city in areas where the housing prices are at low levels. The income level of this cluster is not high and the commuting distances are usually long.

Cluster 8: The main characteristics of this small group are low- and middle-income residents with higher education (Components IV), and upper middle-income class

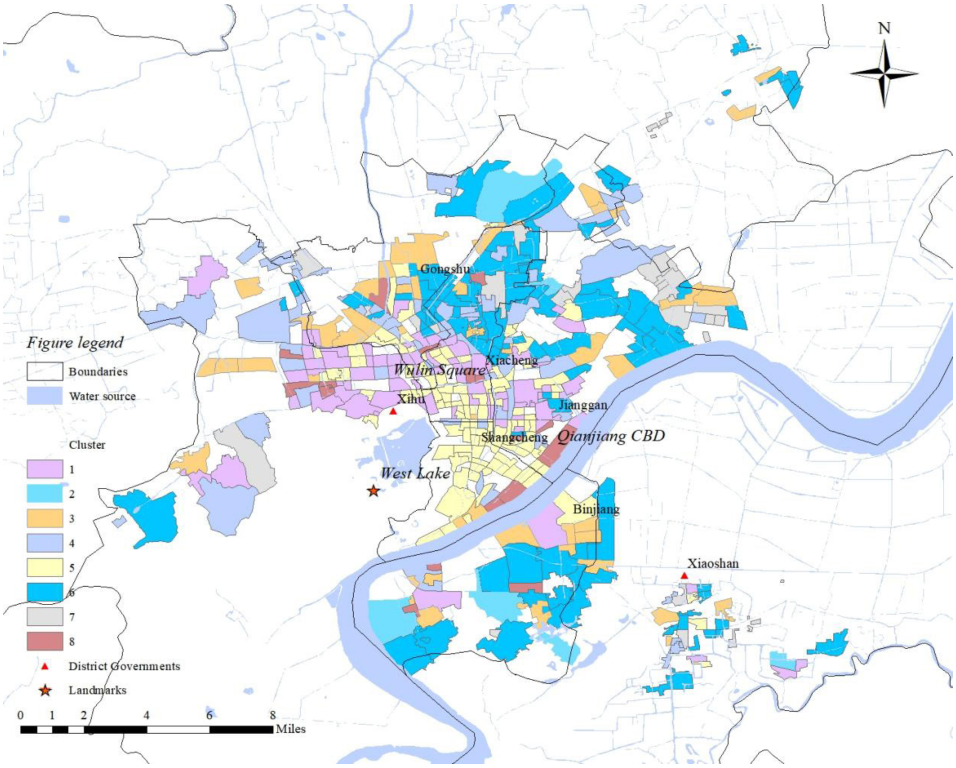

Figure 1. Colored map based on administrative division sub-clustering (The base map is from Baidu Maps)

Table 4. Cross-tabulation of house price clustering and social area clustering

Clusters	Communities	Housing price					Median house price (¥/m^2)
		Low	Middle-low	Middle	Upper middle	High	
CL 1	74	0	15	43	13	3	22500
CL 2	24	15	9	0	0	0	10500
CL 3	50	6	35	9	0	0	16964
CL 4	72	12	38	21	1	0	16133
CL 5	118	2	29	63	22	2	23388
CL 6	128	46	74	7	0	0	13500
CL 7	40	18	21	1	0	0	12000
CL 8	21	3	6	10	2	0	21453

(Components I). The discriminant coefficient of 0.096 for the house price factor showed its characteristic correlation in this cluster. This type of population is distributed in a small area and the house prices are high. According to Figure 1, this group of people mainly live in high-end communities, close to the new CBD, the traditional CBD, and the upcoming third CBD located in the western part of the city.

As mentioned earlier in this Section, the sociodemographic characteristics of some clusters are highly consistent with the average house price level in their neighborhood. In order to further assess the correlation between residential spatial differentiation and housing prices, we clustered the house prices of communities into five categories from low to high, and conducted spatial coupling analysis with social areas. The results in Table 4 clearly show the consistency of these two elements. In the three social areas with high average house prices, the majority of communities have a moderate house price, and only a limited number of communities are in the mid-high and high price range. A similar situation was observed in the social areas with low home prices, where the majority of communities are in the mid-low range. This suggests that there is no significant polarization of house price levels across social areas.

The community clustering was able to identify most of the social characteristics obtained from the sub-district clustering in Zhang et al. (2022). Some small differences may be present because sub-district-level clustering does not allow for a detailed composition of smaller social units, so it would be unable to identify minority factors. The results obtained from the community-level data would also allow for more diverse and superior findings from the social context analysis.

4. Diversity and residential segregation in districts

The functional division of the districts in Hangzhou is quite evident, and the trend of dual-center development is significant (Yue et al., 2010). The Shangcheng District is the emerging central business district; the Xihu District is famous for its quality educational resources, while the Gongshu District is the most polarized in terms of residential differentiation. Two city centers are attractive to residents, while the suburbs include areas of integration with urban villages, and industrial and technology parks.

The diversity index was calculated using Equation (1). The higher the diversity in the performance of a district, the higher the value of the index of diversity (Wu et al., 2014; Mengdi et al., 2018; Křížková & Šimon, 2022). The results in Table 5 show that the area with the lowest level of diversity is Shangcheng District, where the new CBD is located, and the house prices are the highest in the city. The second-tier districts with low diversity are Xihu and Xiacheng Districts, which have higher numbers of quality schools and prosperous business areas, respectively. These districts attract families whose children are students, as well as some people with high standards of living. Although the overall accessibility to the CBD from Xiacheng District and Gongshu District is good, the uneven development between the north and south, as well as the accessibility to the CBD, have led to a more pronounced residential diversity in both districts. A similar situation arises in Jianggan District, where community diversity is weaker along the border with Shangcheng District. The other areas are in the suburbs of Hangzhou, which have mixtures of local residents and migrant workers. The regional diversity index is generally high in these places and shows a radial trend of greater diversity in areas further away from the CBD. The uneven distribution of social resources over time may exacerbate residential segregation and class solidification in different parts of the city.

Based on the results of the cluster analysis, the proportion of communities in various clusters within each district could be calculated. Finally, the cumulative percentage of districts that exceed a certain threshold can be obtained. The threshold was set as {5%, 10%, 20%, ... 90%, 100%}; the results are shown in Table 6 and Figure 2.

Table 5. Diversity index of administrative divisions

Zonal diversity index										
District	CL 1	CL 2	CL 3	CL 4	CL 5	CL 6	CL 7	CL 8	H	H/Hmax
Shangcheng	0	0	2	4	29	0	0	2	0.747	0.071
Xiacheng	16	0	5	10	27	13	2	3	1.667	0.159
Xihu	29	0	7	12	13	3	1	4	1.578	0.151
Gongshu	8	2	12	16	12	16	4	2	8.937	0.854
Jinaggan	20	1	7	17	23	23	11	3	6.448	0.616
Binjiang	2	4	12	1	1	15	1	3	10.471	1
Xiaoshan	10	1	5	5	3	11	7	0	5.424	0.518

Table 6. Cumulative percentage of districts for clusters

Cumulative percentage								
	CL 1	CL 2	CL 3	CL 4	CL 5	CL 6	CL 7	CL 8
5%	85.71%	14.29%	100.00%	85.71%	85.71%	71.43%	42.86%	42.86%
10%	71.43%	14.29%	57.14%	85.71%	71.43%	71.43%	28.57%	0.00%
20%	42.86%	0.00%	14.29%	14.29%	42.86%	57.14%	0.00%	0.00%
30%	14.29%	0.00%	14.29%	0.00%	28.57%	14.29%	0.00%	0.00%
40%	14.29%	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	0.00%
50%	0.00%	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	0.00%
60%	0.00%	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	0.00%
70%	0.00%	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	0.00%
80%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
90%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
100%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

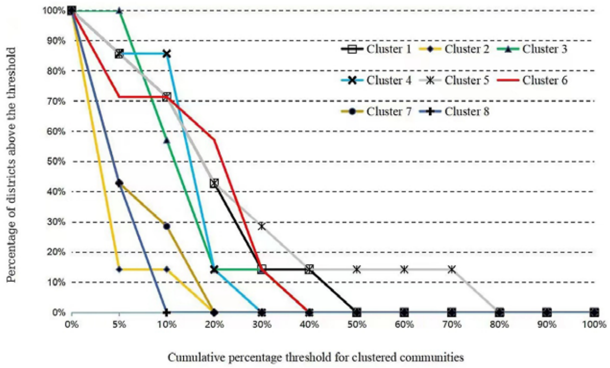

Figure 2. Dissimilarity of the clusters in districts

The residential dissimilarity of various groups was found to basically follow a bifurcated trend. Cluster 1, Cluster 3, Cluster 4, Cluster 5, and Cluster 6 have a relatively high dissimilarity, while Cluster 2, Cluster 7, and Cluster 8 are at a low level. Cluster 5 is the most prominent group. The proportion of districts with more than 30% of this Cluster is 28.7%, and the percentage of communities covered by this group in Shangcheng district is more than 70%. The percentage of districts exceeding the 5% threshold is 85.7%, indicating that Cluster 5 is widely distributed throughout the city. Cluster 1 has a very similar distribution trend to Cluster 5, as the watershed between them is the threshold of 20%. Cluster 6 (Non-local young business service workers, as well as middle-aged and elderly local population) performs most prominently at the 20% level. Its gentle curve shows that it has a relatively high proportion in the majority of districts. Some similarity was also observed between the distribution of Cluster 3 (Families with school-age children) and Cluster 4 (Local residents engaged in primary and secondary industries), whose curves plummet after exceeding the 5% and 10% thresholds, respectively.

The remaining three clusters, i.e., Clusters 2, 7, and 8 are less distributed across districts. Their curves show a plummeting trend, with the watershed at the 5% and 10% threshold indicating that these are minority groups in the city. Neither the low-income group (Cluster 2) nor the high-income affluent group (Cluster 8) has developed a prominent concentration.

Overall, the results of this analysis are consistent with the distribution in Figure 1, and the distribution of each Cluster in the various districts is generally consistent with the pattern of public facility supply. Hangzhou has had a significant improvement in its living standards, business atmosphere, educational resources, and entrepreneurial opportunities in recent years. The overall residential distribution of different social groups was found to be harmonious. No residential imbalance was found at the income level, while a relatively high spatial differentiation was found among the highly educated, families with school-age children, and workers in primary and secondary industries. Residential heterogeneity is mainly influenced by the distribution of industries and educational facilities.

5. Discussion

The diversity indexes for the districts have shown certain degrees of social class solidification in the city. Some districts have uneven development, resulting in significant differences in the populations clustered in different parts of these areas. Affected by housing prices and amenities, residential segregation has been characterized by the highest concentration of residential groups in city centers. The closer the area to the periphery, the more diverse the residents' distribution. For example, residential segregation is more pronounced in places close to the CBDs and high-quality education districts. Because of the concentration of the primary and secondary industries, residential diversity is more pronounced in suburban areas.

Better educational and social resources attract people to live in nearby communities, but the uneven distribution of resources in a city would lead to serious social polarization. If residential segregation continues, then it would lead to deep-rooted class divisions among the residents and create a detriment to further social development. Some past policies, such as the school district allocation policy of "Nearby enrollment" and "Zero school choice", have not produced actual equity but may have even exacerbated income divisions and inequality in educational opportunities (Wen et al., 2017). Peng et al. (2021) stated that policies should discourage the gentrification of low-priced houses by ensuring the rights of low-income groups to have their children attend high-quality primary schools. So, cities must have more reasonable resource allocation policies, such as improving the quality of life and the residential facilities. Moreover, more reasonable allocations of educational resources and the construction of commercial complexes would alleviate the pressure on the selection of residential locations.

The housing price distribution has a strong spatial correlation with the pattern of residential differentiation, reflecting the interaction between housing location choice and house prices. People seeking higher standards of living and families with school-age children should have more options for home ownership and no longer be limited to the CBDs or the vicinities of prestigious schools. Such measures would not only promote the healthy development of intra-city housing prices and the diversification of industries in various areas but also more reasonable housing prices to reduce residential differentiation and consequent social conflicts. The change in Hangzhou's urban development strategy in the last two years has shown that the city's administration has been making the necessary effort. In 2021, a major adjustment to the administrative divisions merged the Shangcheng and Jianggan Districts, as well as the Xiacheng and Gongshu Districts. This zoning adjustment has merged administrative districts with strong and weak residential segregation, thus leveling community distribution and rationalizing the distribution of social resources.

Conclusion

Using socio-demographic variables and variables related to housing conditions, this study used PCA and cluster analysis to classify residential communities in Hangzhou into eight categories. There is a clear spatial differentiation of living space among different social groups. The cluster containing the largest number of communities, dominated by local middle-aged, elderly, and young commercial service workers, resides mainly in the northeastern and southeastern peripheral areas of the city, where the level of housing prices is low. Highly educated people and those employed in primary and secondary industries live in new areas of the city that have expanded in the last two decades, between the traditional downtown and the new CBD. The upper middle-income class occupies both CBD areas, along with the local middle-aged and elderly residents. Families with school-age children are mainly clustered in the older urban areas, where various facilities are convenient, and in the new, western part of the city, where educational resources are higher. The residence pattern showed that the trend of "double-center" development was obvious. Business areas and quality school districts still play significant roles in influencing the housing preferences of residents. The combination of urban and rural areas on the outskirts of the city has resulted in concentrations of primary industries, as well as industrial and technological parks, and presents a distinctly mixed housing situation. This phenomenon is caused by the different functional zoning of the city, unbalanced regional development, uneven distribution of resources, and differences in housing prices. Nowadays, with the influx of non-local immigrants and the rise of new classes in the city, the pattern of spatial differentiation of urban living is gradually becoming more complex.

This study argued that future urban development and housing market-related policies need to focus more on the following aspects. First, future policies should strengthen public facility investments in the outskirts of the city, as well as enhance its educational, employment, and commercial resources, thus improving the city's residential heterogeneity. Second, the housing policy should focus on alleviating the imbalance of development inside a city and provide residents with more diversified housing purchase choices. Third, future development policies should focus more on providing quality housing for all districts of the city. Finally, zoning adjustment, such as the merging of administrative districts with strong and weak residential segregation, is also a viable optimization solution.

Acknowledgements

The authors thank two anonymous referees for helpful comments and feedback.

Funding

This research was supported by National Natural Science Foundation of China [No. 72174178].

Author contributions

Ling Zhang conceived the study and were responsible for the design and development of the data analysis. Baihong Wu were responsible for data collection and analysis and wrote the first draft of the article.

Disclosure statement

Authors do not have any competing financial, professional, or personal interests from other parties.

References

Allen, R., Burgess, S., Davidson, R., & Windmeijer, F. (2015). More reliable inference for the dissimilarity index of segregation. *The Econometrics Journal*, *18*(1), 40–66. https://doi.org/10.1111/ectj.12039

Badcock, B. (1997). Restructuring and spatial polarization in cities. *Progress in Human Geography*, *21*(2), 251–262. https://doi.org/10.1191/030913297670500369

Baer, W. C., & Williamson, C. B. (1988). The filtering of households and housing units. *Journal of Planning Literature*, *3*, 127–152. https://doi.org/10.1177/088541228800300201

Benefield, J. D. (2009). Neighbourhood amenity packages, property price, and marketing time. *Property Management*, *27*, 348–370. https://doi.org/10.1108/02637470910998483

Bhat, C. R., & Guo, J. (2004). A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B: Methodological*, *38*(2), 147–168. https://doi.org/10.1016/S0191-2615(03)00005-5

Chhetri, P., Han, J. H., & Corcoran, J. (2009). Modelling spatial fragmentation of the Brisbane housing market. *Urban Policy and Research*, *27*(1), 73–89. https://doi.org/10.1080/08111140802468291

Cox, T., & Hurtubia, R. (2021). Subdividing the sprawl: endogenous segmentation of housing submarkets in expansion areas of Santiago, Chile. *Environment and Planning B: Urban Analytics and City Science*, *48*(7), 1770–1786. https://doi.org/10.1177/2399808320947728

Dolnicar, S., & Grün, B. (2008). Challenging "Factor–Cluster segmentation". *Journal of Travel Research*, *47*(1), 63–71. https://doi.org/10.1177/0047287508318910

Feng, R., & Han, R. (2021). Study on the spatial differentiation pattern of housing prices in Chengdu-Chongqing City cluster. *Open Access Library Journal*, *8*(2), 1–11. https://doi.org/10.4236/oalib.1107118

Fikire, A. H. (2021). Determinants of urban housing choice in Debre Berhan town, North Shewa zone, Amhara region, Ethiopia. *Cogent Economics & Finance*, *9*(1), 1885196. https://doi.org/10.1080/23322039.2021.1885196

Forrest, R. (2012). *Marginalization in urban China: comparative perspectives*. https://doi.org/10.1080/02673037.2012.617929

Frankenberg, E. (2013). The role of residential segregation in contemporary school segregation. *Education and Urban Society*, *45*(5), 548–570. https://doi.org/10.1177/0013124513486288

Harris, R., & Owen, D. (2018). Implementing a multilevel index of dissimilarity in R with a case study of the changing scales of residential ethnic segregation in England and Wales. *Environment and Planning B: Urban Analytics and City Science*, *45*(6), 1003–1021. https://doi.org/10.1177/2399808317748328

Johnston, R., Forrest, J., & Poulsen, M. (2001). Sydney's ethnic geography: new approaches to analysing patterns of residential concentration. *Australian Geographer*, *32*(2), 149–162. https://doi.org/10.1080/00049180123731

Knox, P. (1987). *Urban social geography: an introduction* (2nd ed.). Longman.

Křížková, I., & Šimon, M. (2022). Measuring residential segregation of non-European migrants using the individualised neighbourhood method: how does Czechia fit to the European landscape? *Applied Geography*, *144*, 102730. https://doi.org/10.1016/j.apgeog.2022.102730

Li, S. M., Hou, Q., Chen, S., & Zhou, C. (2010). Work, home, and market: the social transformation of housing space in Guangzhou, China. *Urban Geography*, *31*(4), 434–452. https://doi.org/10.2747/0272-3638.31.4.434

Li, Z., & Wu, F. (2008). Tenure-based residential segregation in post-reform Chinese cities: a case study of Shanghai. *Transactions of the Institute of British Geographers*, *33*(3), 404–419. https://doi.org/10.1111/j.1475-5661.2008.00304.x

Liu, L., Huang, Y., & Zhang, W. (2018). Residential segregation and perceptions of social integration in Shanghai, China. *Urban Studies*, *55*(7), 1484–1503. https://doi.org/10.1177/0042098016689012

Liu, Y., Dijst, M., & Geertman, S. (2014). Residential segregation and well-being inequality between local and migrant elderly in Shanghai. *Habitat International*, *42*, 175–185. https://doi.org/10.1016/j.habitatint.2013.12.005

Massey, D. S. (2012). Reflections on the dimensions of segregation. *Social Forces*, *91*(1), 39–43. https://doi.org/10.1093/sf/sos118

Mengdi, W., Zhifeng, L., & Yang, X. (2018). Study on spatial-temporal evolution of residential differentiation in Shanghai from an employee perspective. *China City Planning Review*, *27*(3), 6–15.

Modai-Snir, T., & Plaut, P. (2019). The analysis of residential sorting trends: measuring disparities in socio-spatial mobility. *Urban Studies*, *56*(2), 288–300. https://doi.org/10.1177/0042098018798759

Murdie, R. A. (1969). *Factorial ecology of metropolitan Toronto 1951–1961* (Research Paper No. 116). Department of Geography, University of Chicago.

Peng, Y., Tian, C., & Wen, H. (2021). How does school district adjustment affect housing prices: an empirical investigation from Hangzhou, China. *China Economic Review*, *69*, 101683. https://doi.org/10.1016/j.chieco.2021.101683

Preston, V., & Ray, B. (2020). Placing the second generation: a case study of Toronto. *The Canadian Geographer / Le Géographe canadien*, *64*(2), 215–231. https://doi.org/10.1111/cag.12597

Qian, Z. (2015). Hangzhou. *Cities*, *48*, 42–54. https://doi.org/10.1016/j.cities.2015.06.004

Reed, R. (2013). The contribution of social area analysis: modelling house price variations at the neighbourhood level in Australia. *International Journal of Housing Markets and Analysis*, *6*(4), 455–472. https://doi.org/10.1108/IJHMA-02-2013-0011

Rees, P. H. (1970). Concepts of social space: toward an urban social geography. *Geographical Perspectives on Urban Systems*, *6*(1), 306–394.

Reich, R. B. (1991). *The work of nations*. Alfred A. Knopf.

Sarstedt, M., & Mooi, E. (2019). *A concise guide to market research: the process, data, and methods using IBM SPSS statistics* (3rd ed.). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56707-4

Shaw, F. (2002). Is the ageing population the problem it is made out to be? *Insight*, *4*, 4–11. https://doi.org/10.1108/14636680210438990

Shevky, E., & Bell, W. (1955). *Social area analysis; theory, illustrative application and computational procedures*. Stanford University Press.

Shevky, E., & Williams, M. (1949). *The social areas of Los Angeles, analysis and typology*. Pub. for the John Randolph Haynes and Dora Haynes Foundation by the Univ. of California Press.

Shoemaker, S. (1994). Segmenting the U.S. travel market according to benefits realized. *Journal of Travel Research*, *32*(3), 8–21. https://doi.org/10.1177/004728759403200303

Šimon, M., Křížková, I., & Klsák, A. (2020). Immigrants in large Czech cities 2008–2015: the analysis of changing residential patterns using population grid data. *Geografie*, *125*(3), 343–374. https://doi.org/10.37040/geografie2020125030343

Smith, D. W., & Scarpaci, J. L. (2000). Urbanization in transitional societies: an overview of Vietnam and Hanoi. *Urban Geography*, *21*(8), 745–757. https://doi.org/10.2747/0272-3638.21.8.745

Soja, E. W. (1980). The socio-spatial dialectic. *Annals of the Association of American Geographers*, *70*(2), 207–225. https://doi.org/10.1111/j.1467-8306.1980.tb01308.x

Tammaru, T., Strömgren, M., Van Ham, M., & Danzer, A. M. (2016). Relations between residential and workplace segregation among newly arrived immigrant men and women. *Cities*, *59*, 131–138. https://doi.org/10.1016/j.cities.2016.02.004

Tomal, M. (2021). Housing market heterogeneity and cluster formation: evidence from Poland. *International Journal of Housing Markets and Analysis*, *14*(5), 1166–1185. https://doi.org/10.1108/IJHMA-09-2020-0114

Tu, X., Fu, C., Huang, A., Chen, H., & Ding, X. (2022). DBSCAN spatial clustering analysis of urban "Production–Living–Ecological" space based on POI data: a case study of central urban Wuhan, China. *International Journal of Environmental Research and Public Health*, *19*(9), 5153. https://doi.org/10.3390/ijerph19095153

Waddell, P., & Borning, A. (2004). A case study in digital government: developing and applying UrbanSim, a system for simulating urban land use, transportation, and environmental impacts. *Social Science Computer Review*, *22*(1), 37–51. https://doi.org/10.1177/0894439303259882

Wang, Y. P. (2017). Progress and problems of urban housing reform. *Social Policy Reform in China*, 177–190. https://doi.org/10.4324/9781315194561-14

Wei, Y. D., & Li, W. (2002). Reforms, globalization, and urban growth in China: the case of Hangzhou. *Eurasian Geography and Economics*, *43*(6), 459–475. https://doi.org/10.2747/1538-7216.43.6.459

Wen, H., Xiao, Y., & Zhang, L. (2017). School district, education quality, and housing price: evidence from a natural experiment in Hangzhou, China. *Cities*, *66*, 72–80. https://doi.org/10.1016/j.cities.2017.03.008

Williams, D. R., & Collins, C. (2001). Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Reports*, *116*(5), 404–416. https://doi.org/10.1016/S0033-3549(04)50068-7

Wu, Q. (2016). *A theoretical and empirical study on residential spatial differentiation in large cities*. Science Press.

Wu, Q., Cheng, J., & Young, C. (2017). Social differentiation and spatial mixture in a transitional city-Kunming in southwest China. *Habitat International*, *64*, 11–21. https://doi.org/10.1016/j.habitatint.2017.03.019

Wu, Q., Cheng, J., Chen, G., Hammel, D. J., & Wu, X. (2014). Socio-spatial differentiation and residential segregation in the Chinese city based on the 2000 community-level census data: a case study of the inner city of Nanjing. *Cities*, *39*, 109–119. https://doi.org/10.1016/j.cities.2014.02.011

Wu, Y., Luo, J., & Peng, Y. (2020). An optimization-based framework for housing subsidy policy in China: theory and practice of housing vouchers. *Land Use Policy*, *94*, 104526. https://doi.org/10.1016/j.landusepol.2020.104526

Xing, Y., Tarimo, C. S., Ren, W., & Zhang, L. (2022). The impact of health insurance policy on the fertility intention of rural floating population in China: empirical evidence from cross-sectional data. *International Journal of Environmental Research and Public Health*, *20*(1), 175. https://doi.org/10.3390/ijerph20010175

Yue, W., Liu, Y., & Fan, P. (2010). Polycentric urban development: the case of Hangzhou. *Environment and Planning A*, *42*(3), 563–577. https://doi.org/10.1068/a42116

Zhang, L., Zhu, L., Shi, D., & Hui, E. C. M. (2022). Urban residential space differentiation and the influence of accessibility in Hangzhou, China. *Habitat International*, *124*, 102556. https://doi.org/10.1016/j.habitatint.2022.102556

Zhao, P. (2013). The impact of urban sprawl on social segregation in Beijing and a limited role for spatial planning. *Tijdschrift Voor Economische En Sociale Geografie*, *104*(5), 571–587. https://doi.org/10.1111/tesg.12030

Appendix

Table A1. Component matrix after maximum variance rotation

	Component						
	1	2	3	4	5	6	7
Average price of community	0.105	0.128	–0.117	0.167	0.173	–0.463	–0.419
Housing space per capita	0.238	–0.19	0.197	–0.027	0.404	0.096	0.364
Number of cars per household	0.276	–0.21	0.191	0.294	0.174	0.569	0.339
Population ratio 6–19 years old	0.059	–0.082	–0.26	0.1	–0.116	0.719	–0.056
Population ratio 20–29 years old	–0.033	–0.1	0.647	0.074	–0.302	–0.288	0.327
Population ratio 30–59 years old	0.08	–0.059	0.396	0.022	0.151	0.729	0.01
Proportion of population with university or higher education	0.372	0.209	0.378	0.584	0.092	–0.067	–0.177
Proportion of respondents with monthly income between 5k–10k	0.327	0.091	0.167	0.798	–0.1	–0.112	–0.135
Proportion of respondents with monthly income between 10k–15k	0.787	0.011	0.051	0.242	–0.067	0.004	–0.084
Proportion of respondents with monthly income above 15k	0.86	0.033	–0.032	–0.027	–0.059	0.067	–0.026
Proportion of monthly household income between 10k–20k	–0.101	–0.03	–0.004	0.803	0.006	0.21	0.069
Proportion of monthly household income between 20k–30k	0.583	–0.005	–0.157	0.51	–0.084	0.062	–0.047
Proportion of monthly household income between 30k–40k	0.833	0.046	–0.084	–0.014	–0.099	0.022	0.009
The proportion of household registration in local	–0.218	–0.086	–0.28	–0.049	0.84	–0.029	–0.058
The proportion of household registration ratio	–0.148	–0.013	–0.304	–0.005	0.856	–0.026	–0.093
Proportion of population work as head of state institutions and enterprises	0.425	0.155	0.186	0.298	0.287	0.203	–0.096
Proportion of commercial and service workers	–0.176	–0.033	0.675	–0.028	–0.324	0.061	–0.046
Proportion of personnel in agricultural, forestry, animal husbandry and fishery	–0.049	0.072	–0.029	0.063	0.043	–0.012	0.689
Proportion of operators of production and transportation equipment	–0.09	–0.063	0.126	–0.148	–0.119	0.055	0.675
Proportion of other employees	–0.054	0.004	–0.835	–0.2	0.091	–0.219	–0.123
Average travel time to entertainment	–0.026	0.684	0	0.047	–0.109	0.108	–0.061
Average travel time for daily living	–0.05	0.666	–0.041	0.016	–0.001	–0.053	0.124
Average travel time to work	0.055	0.671	–0.12	0.059	0.013	–0.192	–0.319
Average travel time to shopping	0.147	0.631	0.08	–0.053	0.004	–0.051	0.158
Average travel time to home	0.052	0.838	–0.024	0.08	–0.007	–0.146	–0.169

Note: Spatial features of principal components; deeper colors indicate higher scores given by the community. The factor scores for each community indicates the performance of a weighted linear combination of the items in that principal component.

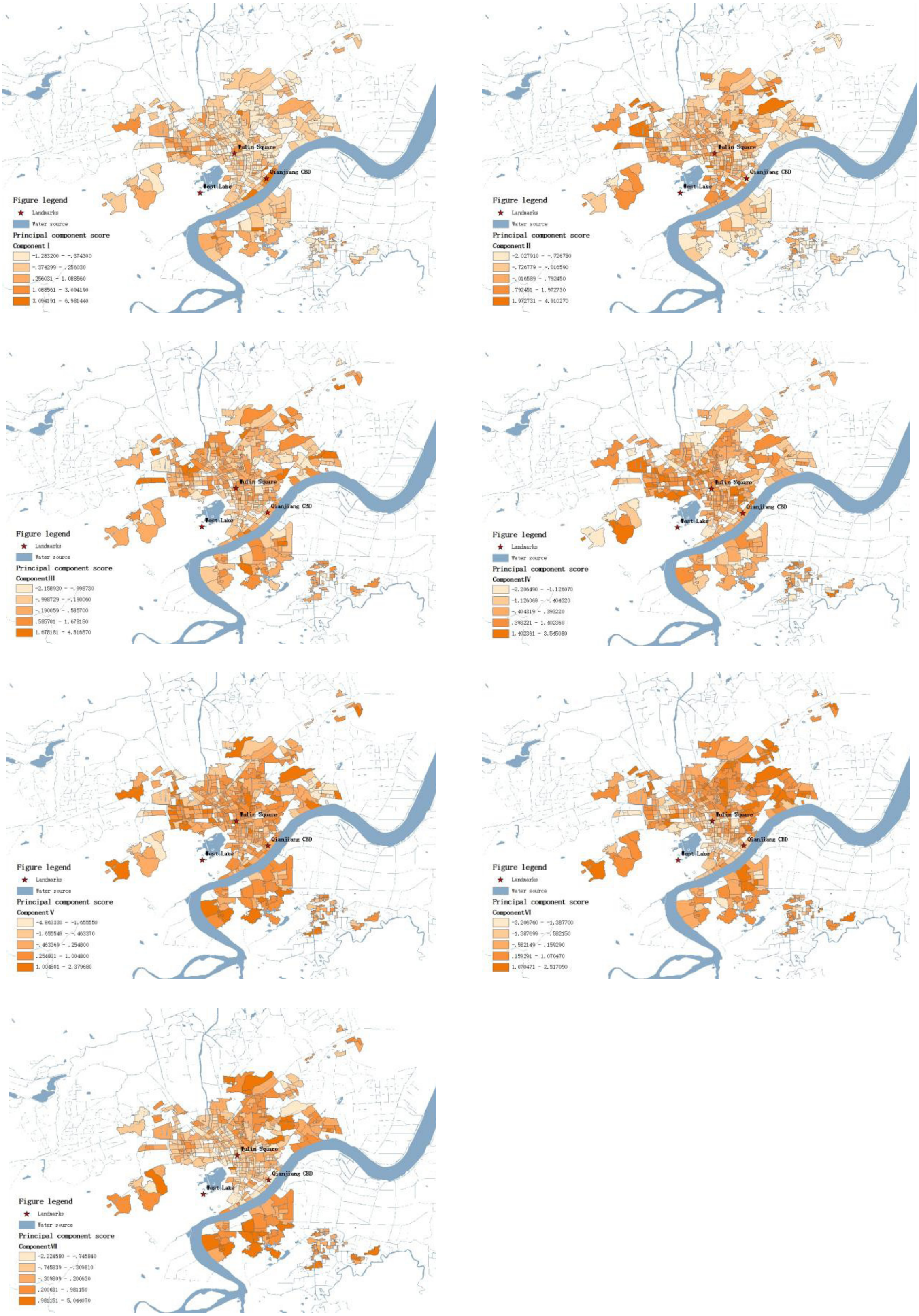

Figure A1. Colored map of factor scores