Taylor & Francis
Taylor & Francis Group

Big Earth Data

ISSN: 2096-4471 (Print) 2574-5417 (Online) Journal homepage: www.tandfonline.com/journals/tbed20

A city-level dataset of population subcenters in Chinese cities for urban polycentric detection (2001–2021)

Xinyue Gu, Mengni Yu & Xintao Liu

To cite this article: Xinyue Gu, Mengni Yu & Xintao Liu (2025) A city-level dataset of population subcenters in Chinese cities for urban polycentric detection (2001–2021), Big Earth Data, 9:4, 1210-1225, DOI: 10.1080/20964471.2025.2560164

To link to this article: https://doi.org/10.1080/20964471.2025.2560164

Published online: 17 Sep 2025.

BIG EARTH DATA
2025, VOL. 9, NO. 4, 1210–1225
https://doi.org/10.1080/20964471.2025.2560164

DATA NOTE

A city-level dataset of population subcenters in Chinese cities for urban polycentric detection (2001–2021)

Xinyue Gu [a], Mengni Yu [b] and Xintao Liu [a]

[a]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China; [b]Thrust of Urban Governance and Design, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

ABSTRACT

Urban areas across the globe are experiencing a shift towards polycentric development, characterized by the emergence of multiple subcenters within cities that can respectively function as economic, social, and residential hubs. In response to this trend, we generate a city-level dataset of population subcenters covering 336 cities in China to serve dynamic urban polycentric detection from 2001 to 2021 though analyzing Landscan data. Our dataset has been validated by diverse socio-economic factors, demonstrating that it can provide a relatively accurate depiction of urban structural changes. It comprehensively captures the evolution of urban polycentric structures within China's rapidly transforming cities, offering detailed insights into the formation and dynamics of population subcenters over two decades. These findings can facilitate policymakers with evidence-based tools to optimize infrastructure and services distributions, thereby fostering efficient urban environments. Moreover, the dataset supports advanced spatiotemporal analysis and modeling, which are essential for understanding urban sustainable development. The dataset is beneficial to explore patterns of urban growth, assessing policy impacts, and developing predictive models for urban structure evolution. All data, figures and relevant results are publicly available on Zenodo: https://doi.org/10.5281/zenodo.14279505.

ARTICLE HISTORY
Received 5 April 2025
Accepted 17 August 2025

KEYWORDS
Urban structure; population center; polycentric structure; urban planning

1. Introduction

The correlation among urban spatial structure, economic development, innovation output, policy performance, etc., has consistently been a focus of exploration in urban planning (Li & Du, 2022; Li & Liu, 2018; Wang et al., 2021). From a regional scale perspective, urban spatial structures encompass two fundamental models in terms of development: monocentric and polycentric. The monocentric model emphasizes the concentrated development of a single primary urban pole, including the sprawl into suburban areas. In contrast, the polycentric model stresses on the development of

CONTACT Xintao Liu xintao.liu@polyu.edu.hk The Hong Kong Polytechnic University, Blok Z, Hung Hom, Kowloon, Hong Kong 999077, China

multiple urban poles, promoting social, economic, and environmental objectives with nearly equal equity levels (Bridge & Watson, 2008; Liu & Wang, 2016). Monitoring the polycentric structure of cities not only addresses changes in the external environment but also facilitates harmonious internal development, providing crucial data support and structural references to achieve a high-quality, sustainable urban development.

Early research, influenced by "central place theory", focused on exploring the scale performance of monocentric urban development structures (Mills, 1981). As urbanization has intensified and city sizes have rapidly expanded, a polycentric spatial strategy has become a favored development concept among scholars (Volgmann & Münter, 2022). Relevant studies have proven that polycentrism impacts various urban performances, including economic growth (Zhang et al., 2017), transportation efficiency (Wang & Debbage, 2021), urban resilience (Jia et al., 2020), and sustainable urban development (Yin et al., 2023). The scope of research has also extended further to internal urban scales (e.g., Central Business Districts), inter-urban scales (e.g., metropolitan regions), and trans-regional scales (e.g., continental "development poles" identified in the European Union's territorial development policies) (Halbert et al., 2006).

Meanwhile, major cities worldwide have actively invested in the development of new districts, promoting a spatial structural transformation from monocentric to polycentric urban systems. Examples include the Los Angeles Metropolitan Area in the United States (Lee, 2007), Greater London in the United Kingdom (Chiaradia et al., 2012), and South Korea's Busan-Ulsan-Gyeongnam (BUG) megacity project (Baek & Joo, 2022). Specifically, China stands out as the most typical representative with the largest number of such initiatives, providing researchers with an extensive field of study (Harrison et al., 2023; Huang et al., 2017; Yang et al., 2018). Given that China is the largest developing country, monitoring and measuring its urban spatial structures has become a critical research topic, which can provide significant reference for cities in the Global South.

This shift towards polycentric development reflects a global trend where cities aim to decentralize growth to alleviate congestion, improve efficiency, and promote balanced regional development (Crevoisier & Rime, 2021). In the context of China, this involves not only the construction of entirely new urban areas but also the revitalization and densification of existing secondary centers within larger metropolitan regions. After achieving a significant leap in urbanization rates, China has actively invested in the development of polycentric cities in recent years. This effort spans from the continental scale of the Belt and Road Initiative to the regional scale of urban agglomerations strategies (Gu et al., 2024), down to the establishment of national-level new districts, high-tech industrial development zones, and secondary urban centers, all showcasing pronounced multi-polar development characteristics (Yu et al., 2023). These policies have, on one hand, promoted urban economic development and spatial expansion; and on the other hand, led to numerous practical issues, such as uncontrolled urban sprawl, imbalanced spatial structures, and difficulties in unit collaboration (Fang, 2015; Jiang & Wei, 2024).

As a result, the question of whether urban development should be monocentric or polycentric—a topic being enthusiastically debated within the field of urban planning—has sparked intense discussions in China (Li & Liu, 2018; Liu et al., 2016; Wang & Niu, 2023). The debate is not only about the efficiency and effectiveness of different urban forms but also touches upon broader issues of sustainability, equity, and resilience in the face of rapid urbanization. The lessons learned from these developments can offer valuable insights into sustainable urban planning and management practices for other rapidly urbanizing regions around the world.

Therefore, conducting research on urban spatial structure with China as the backdrop can bring positive impacts from two perspectives. First, at the theoretical level, China's unique developmental context can provide a new perspective on the debate between "monocentric" and "polycentric" models. Compared to the market-oriented polycentric development adopted in American cities, the polycentric nature of Chinese cities manifests as "planned polycentrism", meaning that through local policy guidance, urban spatial structure planning is carried out to achieve polycentric development objectives (Li & Du, 2022). Its distinctive policies and abundant construction achievements may yield research conclusions different from those under the American context. Meanwhile, from a practical standpoint, foundational research on spatial structure holds significant practical value for addressing existing problems in Chinese cities and promoting the sustained and healthy development of urban economies. Urban policy formulation requires spatial structure as a prerequisite. Policymakers integrate and allocate resources, optimize land use layouts, promote industrial agglomeration and collaborative development, improve the distribution of public services and infrastructure, enhance residents' quality of life, and effectively control the disorderly expansion of cities, thereby boosting urban competitiveness (Wang et al., 2019).

Unfortunately, research on urban spatial structure in China is often plagued by the lack of a consistent definition and reliable datasets, which hinders effective investigation. Despite the wealth of studies on polycentricity, a few of them have comprehensively evaluated and reviewed this concept, leading to significant variations in the foundational statistics for polycentricity.

Since polycentricity can have different meanings at various geographic scales (e.g., within cities, between cities, and at the regional scale) and also presents ambiguous concepts from different analytical perspectives (e.g., morphological and functional diversity), research conclusions are scattered and even contradictory. For example, regarding whether polycentricity promotes urban economic development, Hua and Sun (2015) and Li and Liu (2018) affirm the positive implications of the polycentric development model based on its ability to alleviate diseconomies of agglomeration. Conversely, Li et al. (2019) and Li and Liu (2018) argue from the perspective of agglomeration externalities that the polycentric development model might hinder spatial agglomeration of elements within cities, thereby limiting improvements in urban economic efficiency.

Although research studies related to urban spatial structure require multi-year data across multiple cities, few of them provide multi-year statistics for urban structures across all of China. Li & Du (2022) calculated the polycentricity of 267 cities in China from 2006 to 2016, but their data lack recent statistics and do not cover the entire country. Li and Liu (2018) computed the polycentricity of 337 prefecture-level cities in China for the year 2014, missing a longitudinal comparison over multiple years. Chen et al. (2021) tabulated the degree of polycentricity for 23 provinces in China from 1997 to 2013, covering the whole area but with room for improvement in spatial precision of the data.

Based on these above backgrounds, this study defines the concept of polycentricity in urban areas and proceeds to analyze the spatial structure across all regions of China, quantifying the polycentric nature of urban forms. Specifically, the study adopts a morphological definition of urban centers based on "the dispersed distribution of population and land use" (Rauhut, 2017). Unlike functional diversity that exhibits characteristics of "decentralized concentration," morphological diversity manifests as "concentrated decentralization," featuring less influence from internal and external connections and being more suitable for unified studies over multiple years within a broad geographical scope (Burger & Meijers, 2012; Wang & Niu, 2023; Yu et al., 2022). Second, we are able to ensure accuracy and currency of when using the Landscan global population raster dataset as the data source (Li & Du, 2022; Li & Liu, 2018; Liu & Wang, 2016) while conducting research on all prefecture-level cities across China.

We provide urban spatial structure data for of China; these cover 336 prefecture-level administrative regions, comprising four municipalities directly under the Central Government, two special administrative regions of Hong Kong and Macao, as well as the Taiwan Province from 2001 to 2021, which has been shared in the public data repository Zenodo (Gu et al., 2025). This dataset offers extensive reusability and significant contributions to urban studies in China.

First, it serves as a foundational resource for multidisciplinary research, supporting an in-depth exploration into changes in urban form and their impacts in fields, such as urban planning, economics, and sociology. It also provides scientific evidence for government decision-making, facilitating optimizing the layout of public service facilities and assessing the effectiveness of urban development policies. Second, this dataset plays a critical role in environmental protection and sustainable development by helping identify relationships between high-density development zones and nature reserves. This thus promotes the achievement of relevant goals outlined in the United Nations' 2030 Agenda for Sustainable Development, particularly those concerning sustainable cities and communities (SDG 11). Overall, this dataset lays a solid foundation for analyzing the benefits of China's urban spatial structure, exploring the relationship between spatial structure and urban development, and informing the formulation of urban development policies.

2. Methods

2.1. Research area and data collection

The population data is derived from the 1-km resolution population spatial distribution raster data of the LandScan dataset for years between 2001 and 2021. The urban vector maps originate from the 1:1,000,000 vector database provided by the National Geographic Information Resource Catalog Service System website, which delineates urban boundaries and defines the scope of cities. This study is conducted at the prefecture-level city unit, encompassing a total of 336 prefecture-level cities. Generally speaking, as the main body of urban administrative units in China, prefecture-level cities consist of several or dozens of districts, county-level cities, and counties, covering areas of thousands of square kilometers with populations in the millions. In scale, they are equivalent to metropolitan areas in the United States or provincial/regional administrative units in Western Europe, fully demonstrating the agglomeration and dispersion of urban activities (He et al., 2019). Therefore, the samples of this study can characterize the entirety of China.

2.2. Identification of urban population subcenters

The scale of a city is determined by its total population, while the structure of a city is defined by the spatial distribution of its population. In China in particular, urban development is strongly influenced by planning policies, with infrastructure and public service allocations often based on per capita quotas. As a result, population centers in Chinese cities typically correspond to functional centers (Yue et al., 2019). Based on this, this study defines polycentric cities as those characterized by multiple spatially distinct and functionally significant population centers, while monocentric cities are defined as those dominated by a single center that concentrates the majority of the city's population and urban functions. Methods for identifying city centers, include minimum (absolute) density-based (Krugman, 1991), relative density-based (Giudici et al., 2024), nonparametric methods (Lee, 2007), among others. Due to the large number of cities involved in this study and the high variability in population distribution patterns across different regions, methods that rely on predefined density thresholds—such as minimum density-based or relative density-based approaches—are prone to arbitrariness and inconsistency. Therefore, we employ nonparametric methods based on spatial autocorrelation to calculate city centers, which does not require such assumptions.

Nonparametric methods for identifying urban centers include the percolation method (Cao et al., 2020) and clustering algorithms, such as DBSCAN (Tu et al., 2022) and KDE-based techniques (Kucukpehlivan et al., 2023). However, the percolation method requires determining an optimal threshold, while clustering algorithms often rely on repeated calibration of key parameters—parameters that may not be adaptable across multiple cities and years.

In contrast, the Local Moran's I (LMI) index offers a statistically grounded approach to identifying clusters of population density. It is particularly suitable for large-scale, consistent, and repeatable comparative analyses. As a measure of spatial autocorrelation, LMI is robust to data heterogeneity and captures local patterns more effectively. Originally proposed by Anselin (1995), LMI decomposes the global Moran's I into contributions from individual spatial units, enabling the detection of localized clusters and spatial association patterns. Thanks to its ability to identify statistically significant clusters by comparing each location to its neighbors—without the need for predefined thresholds—LMI is adaptable to different spatial scales and city types. Consequently, it has been widely used in urban studies to identify urban centers (Li & Liu, 2018; Liu & Wang, 2016; Tepanosyan et al., 2019; Wang et al., 2023). In this study, we adopt the Local Moran's I index to detect population clusters and extract urban subcenters, as illustrated in Figure 1.

STEP I: Extract potential population centers. In this study, the Local Moran's I index (LMI) was used to identify population centers for each city. Its calculation formula is as follows (Anselin, 1995; Zhang et al., 2008):

$$I_i = \frac{z_i - \bar{z}}{\sigma^2} \sum\nolimits_{j=1, j\neq i}^{n} \left[w_{ij}\left(z_j - \bar{z}\right)\right] \tag{1}$$

Where z_i is the value of variable z is the value of variable i; $\bar{z}$ is the mean of z over all n locations; z_j is the value of variable z at location j, where $j \neq i$; σ^2 is the variance of z; w_{ij} is the spatial weight, defined as the inverse of the distance d_{ij} between locations i and j.

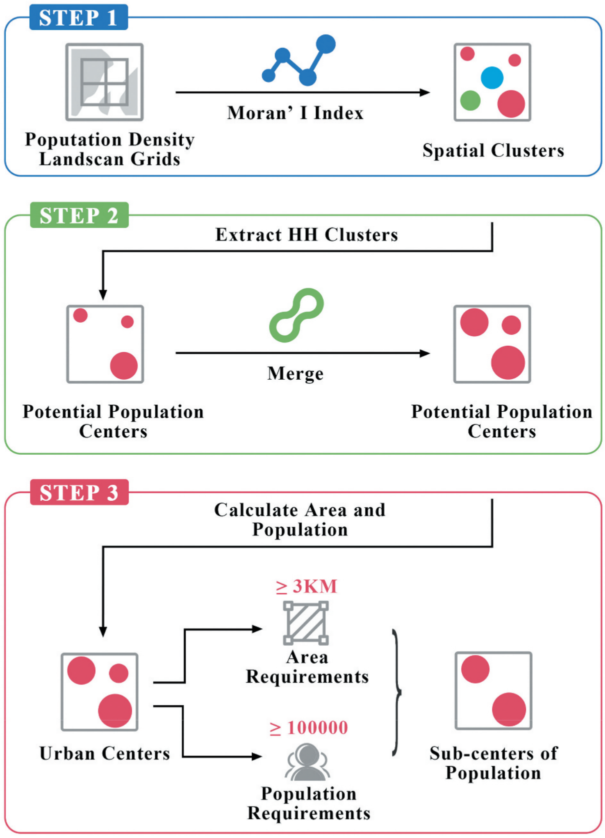

Figure 1. Framework diagram of identification of urban population subcenters.

This index measures the degree of spatial autocorrelation by evaluating each unit in relation to its neighboring units. Statistically significant Local Moran's I values are classified into four types: HH (High-High), LL (Low-Low), HL (High-Low), and LH (Low-High). Among these, HH indicates areas with high population density surrounded by similarly high-density areas, which are identified as potential population subcenters.

STEP II: Assign values to sub-centers of population. Considering that a population center should be a contiguous area, adjacent HH grids (potential population centers) are merged into centers. Then, the area of each center is calculated using ArcGIS, and LandScan population raster data is connected to obtain the area and resident numbers of each potential center.

STEP III: Determine urban centers. Following the method of Liu and Wang (2016), it is believed that a population center should contain at least three grids (3 km) and have more than 100,000 residents. Centers that do not meet these criteria, having smaller areas and fewer populations, are eliminated. Among all the identified population centers, the one with the largest total population is defined as the main center, while the remaining centers are classified as sub-centers. Based on this classification, we construct a relative polycentricity index to capture the distribution pattern of urban population centers.

Calculation of Urban Polycentricity Index. In this study, we use the proportion of population in each sub-center ($POP_{subcenter}$) to all population centres in a city to measure the degree of urban polycentricity (Wang et al., 2021), and the equation is as follows:

$$Poly = \frac{pop_{subcenter}}{pop_{subcenter} + pop_{maincenter}} \tag{2}$$

This indicator reflects the population share of sub-centers relative to the main center, thereby indicating their relative importance. A value close to 0 suggests a highly monocentric structure, where the main center holds a dominant position. In contrast, a value approaching 1 indicates a polycentric structure, where sub-centers are comparable in population size to the main center. This index enables comparative analysis of urban polycentricity across different cities and over time.

3. Data records

The comprehensive dataset constructed for this study is available for download on the public data repository Zenodo (Gu et al., 2025). The dataset includes 21 shapefiles and two aggregated CSV files, covering population distribution centers and polycentricity indices of 336 prefecture-level administrative regions in China (including Hong Kong, Macao, and Taiwan). The time span of the data ranges from 2001 to 2021.

- Shapefiles: The spatial patterns of population subcenters for each city in the whole China, from 2001 to 2021. Each shapefile includes two attributes: Area (the area of the corresponding subcenter, km^2) and Sum (the total population of the corresponding subcenter).
- CSV files: The number of centers and the urban polycentricity index for each city in the whole China, from 2001 to 2021.

Following the implementation of the small-town household registration management system reform in 2001, restrictions on hukou (household registration) in county-level cities and small towns were completely lifted in China, significantly increasing urban-rural mobility (Peng, 2023). This reform has further influenced urban population distribution, which is why this study selects the year 2001 as the starting point for analysis, with a 20-year time frame to explain the characteristics of population distribution in China's prefecture-level cities.

This research chose three significant urban agglomerations as sample areas to showcase the basic situation and trends of the data. They are the Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Pearl River Delta, as illustrated in Figure 2. These regions were selected due to their deep historical foundations, economic advancement, population density, and strategic importance, making them well-suited to represent the evolution of urban development in China. The Beijing-Tianjin-Hebei region, located in northern China, plays a key role in national political and industrial development. The Yangtze River Delta in eastern China, centered around Shanghai, is the country's most important economic engine. The Pearl River Delta in southern China, anchored by cities such as Guangzhou and Shenzhen, is known for its role in China's export-oriented economic reform and rapid urban expansion.

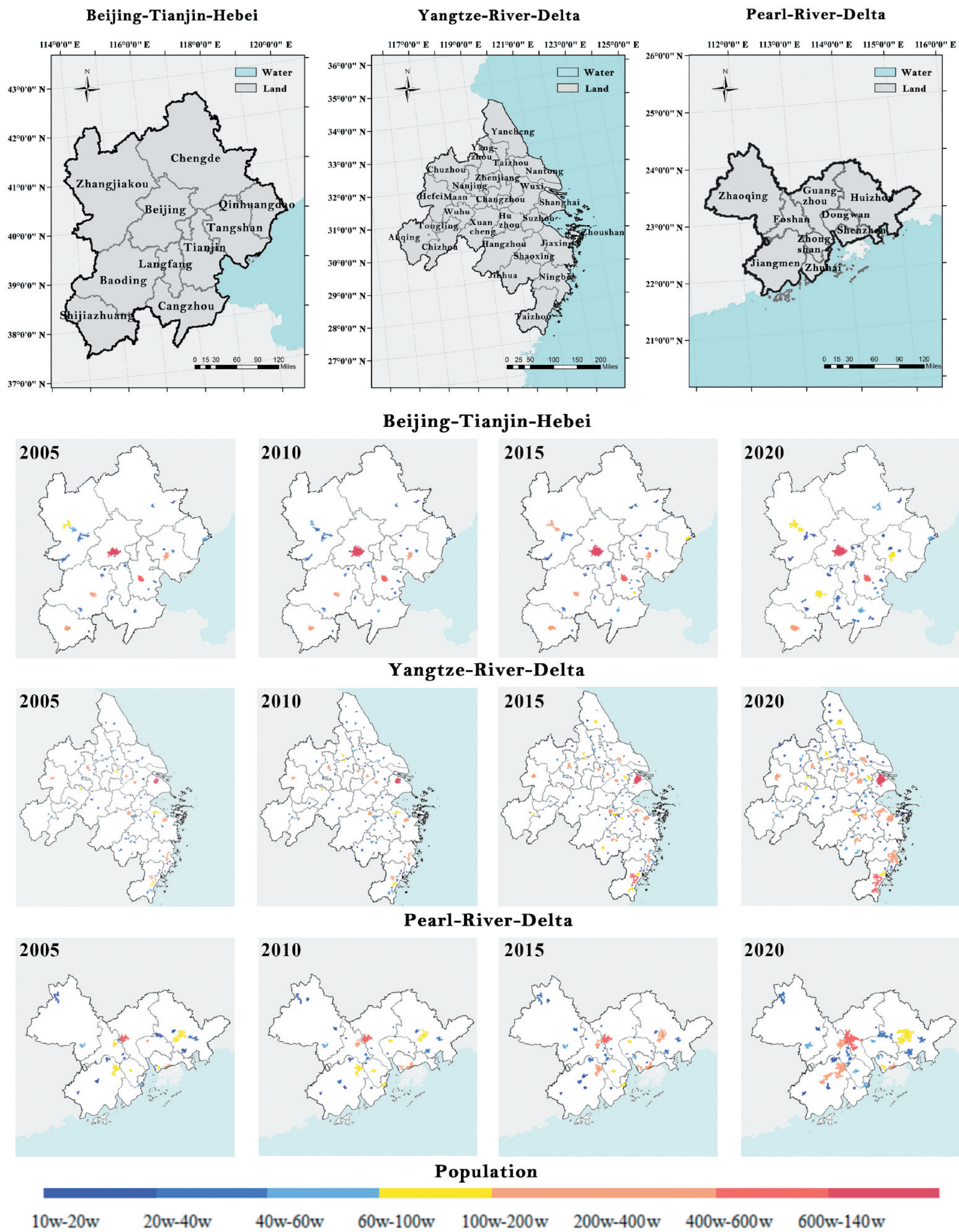

Figure 2. Spatial-temporal patterns of the urban structure in the cities of Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Pearl River Delta.

Urban agglomerations consist of several closely linked cities or metropolitan areas within a relatively small region, often characterized by economic integration and interconnected infrastructure. According to our dataset, it can be observed that the data generally conforms to the development patterns of cities, with the urban structure in these three urban agglomerations evolving from monocentric to polycentric.

4. Technical validation

Based on previous studies, we employed Pearson correlation analysis to establish the relationship between polycentricity index and urban scale, built environment, resource endowment, economic construction, and population as well as social development (Liu & Wang, 2016; Sat, 2018; Thiel et al., 2019). This approach was used to validate the calculated polycentricity index. During the calculation process, the data of each indicator were transformed using logarithms to eliminate the impact caused by inconsistencies in measurement units. The data source we refer here is the "China Urban Statistical Yearbook 2002–2022". Due to missing indicator data across different years, we selected valid data indicators to test some prefecture-level cities. Table 1 shows the validation indicators.

Through Pearson correlation analysis of the multicentricity indices for each year and the indicators listed in Table 1, we obtained the results shown in Figure 3. Despite some indicators being insignificant in certain years, at least three indicators per year were significant (P-value less than 0.1, meaning the correlation is significant). Besides, all P-values were less than 0.5, indicating the reliability and accuracy of our multicentricity index (Biau et al., 2010; Taylor & Bates, 2013).

To further validate the accuracy of the identified urban subcenters, we supplemented the statistical correlation analysis with direct spatial comparison and localized interpretation.

First, we compared our results with the Global Human Settlement Layer—Urban Centre Database (GHS-UCDB), which provides harmonized global data on urban center characteristics in both geospatial vector and tabular formats, using urban centers as reporting units (Melchiorri et al., 2024). Taking 2020 as an example, we selected four representative cities—Chongqing, Dalian, Nanning, and Nanchang—and overlaid our identified subcenters with the urban extents provided by the GHS-UCDB, as shown in Figure 4. These cities were chosen to reflect diverse geographic locations and urban development types in China: Chongqing, located in southwest

Table 1. Description of the control variables.

Category	Implications
Urban Scale	Built-Up Area
	Year-End Permanent resident population
Built Environment	Year-end Actual Cultivated Land Area
	Residential Land Area
	Green Space Area
Local Resource	Household Gas Consumption
	Total Water Resources
Economic Development	Total Retail Sales of Consumer Goods
	Actual Utilized Foreign Investment of the Year
	Average Wage of On-the-Job Employees
Social System	Number of Hospitals
	Number of Higher Education Institutions
	Total Postal Services Volume
	Total Telecom Services Volume
	Total Passenger Traffic

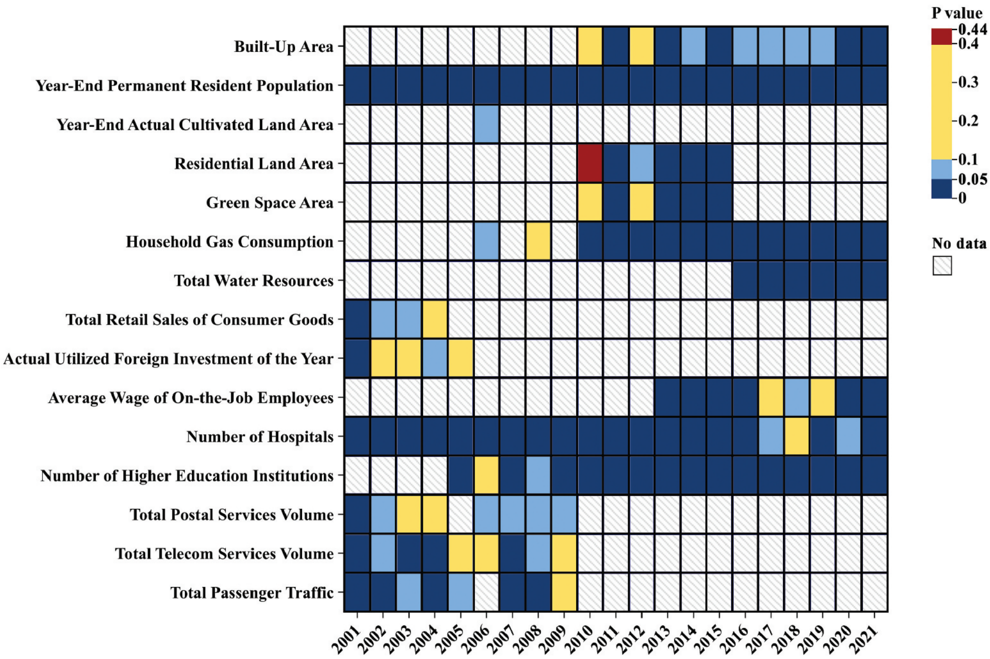

Figure 3. Pearson correlation heatmap between polycentricity and indicators.

China, is a mountainous inland megacity known for its polycentric and topography-constrained development; Dalian, in the northeast coastal region, represents a port city with an axial and dispersed spatial structure shaped by transportation corridors; Nanning (southern China) and Nanchang (central China) are fast-growing provincial capitals characterized by compact urban forms and high population density in their core areas. The results demonstrate a high degree of spatial consistency, particularly in the rapidly urbanizing suburban areas. Some discrepancies in newly emerging centers are likely due to the GHS-UCDB being released in 2024, whereas our analysis is based on population data from 2020. These variations are consistent with expected urban growth trajectories.

Second, we conducted visual validation using satellite imagery and recent urban planning documents. For example, in Chongqing, the identified centers covered the nine traditional central districts, while new centers emerged in suburban county-level cities such as Rongchang, Yongchuan, and Changshou, aligning well with known functional hubs and built-up land use patterns. This validation is also consistent with the findings of Zhang et al. (2022). In Dalian, the spatial distribution of subcenters aligns with the city's planned cluster-based development strategy, with a core in the central district and extensions southward toward the Lushunkou New Urban Area and northward toward clusters such as Wafangdian and Zhuanghe, forming a compact, corridor-based urban pattern along the Shenda and Danda axes.

The validations conducted through comparison with the GHS-UCDB and visual analysis of functional urban land use confirm that the identified subcenters in our dataset are both statistically sound and spatially aligned with the actual urban functional structure.

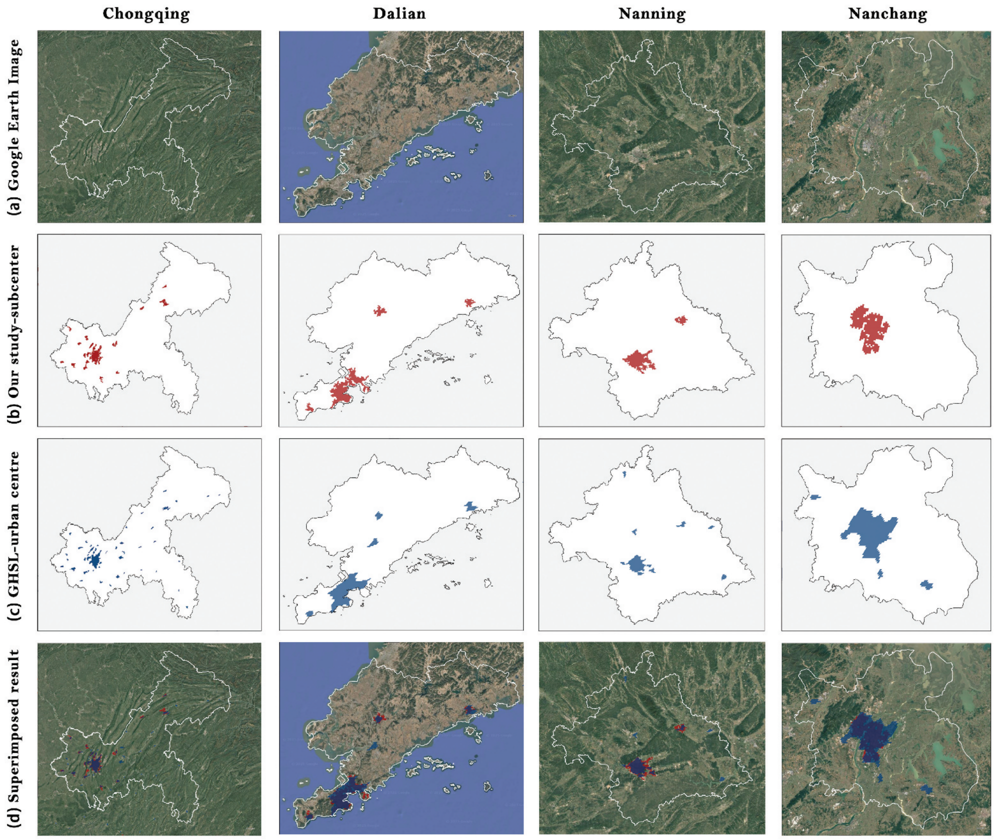

Figure 4. Comparison of urban centers between our results and GHS-UCDB R2024. (a) Google Earth imagery (2025). (b) Subcenters identified in our study (2020). (c) GHSL-defined urban centers (2024). (d) Superimposed comparison.

In addition to validation, we compared our dataset with existing urban center datasets, including both global resources like the GHS-UCDB and China-specific studies (Li & Liu, 2018; Liu & Wang, 2016; Wang et al., 2019). Our dataset offers significant improvements in both functional relevance and temporal resolution. First, unlike the GHS-UCDB, which provides static global urban center data, our dataset focuses specifically on Chinese cities, capturing annual population dynamics at a 1 km resolution. This enables more precise and functionally meaningful identification of subcenters that reflect real-world urban activity patterns. Second, while existing China-focused studies have assessed centrality for selected cities, they are limited in terms of temporal coverage, spatial scope, and methodological consistency. In contrast, our dataset systematically captures the evolution of urban centers across all Chinese prefecture-level cities over a 21-year period (2001–2021), offering a unified foundation for longitudinal analysis. These improvements make our dataset particularly valuable for studying urban development trajectories, conducting policy analysis, and supporting urban modeling, especially in rapidly growing and increasingly polycentric cities.

5. Usage notes

This study provides a comprehensive dataset on the population distribution centers of 336 cities in China from 2001 to 2021. The dataset, which are replicable, auditable, and scalable based on the methods and data adopted in this research, has been uploaded to the public data repository Zenodo for unrestricted access. This dataset lays a solid foundation for analyzing the benefits of urban spatial structures in China, exploring the relationship between spatial structure and urban development, and formulating urban development policies. The advantages of this dataset are as follows.

First, the dataset presented in this study encapsulates the fundamental characteristics of urban spatial structures in China. Compared to existing data, our dataset allows for a comprehensive observation of the spatiotemporal patterns of urban spatial structures in China, aiding in-depth analysis of the efficiency issues of urban spatial structures, thereby facilitating to determine whether urban development structures should be single-centered or multi-centered.

Second, the proposed dataset is compatible with other social data and is suitable for cross-sectional, time-series, and panel data studies. All these facilitate empirical research on external factors, such as the built environment, social systems, economic development, and technological innovation related to urban spatial structures and urban development. For instance, this dataset can evaluate the economic benefits of different spatial structures in cities and the impact of urban structures on urban innovative development.

Third, this dataset includes vector data of sub-centers across all regions of China, laying the groundwork for urban and regional development policies. For example, it enables the formulation of master plans according to city sub-centers, effective allocation of resource facilities, avoidance of resource waste or regional inequality, and forming urban development strategies to organize urban structures and effectively control the disorderly expansion of cities.

While the dataset generated in this study inevitably comes with certain limitations, further enhancement of this work can be conducted in the following aspects in the future. First, our current assessment of urban polycentricity is based solely on population data and does not account for the spatial distance between main and secondary centers (Li & Liu, 2018). To address this limitation, we plan to develop a more comprehensive polycentricity index that incorporates distance-weighted measures. Specifically, the future index will integrate: (1) the number of sub-centers, (2) the population share of sub-centers relative to the total urban population, and (3) spatial dispersion measures that reflect the distance between the centers.

Moreover, due to limitations in data availability—specifically, the lack of high-resolution, city-level socio-economic indicators with consistent temporal coverage from 2001 to 2021—we were unable to implement GWR analysis at this stage. Nonetheless, we recognize the value of spatial statistical models such as GWR for exploring the spatial associations between urban polycentricity and socio-economic factors. In future work, we plan to incorporate such models to better capture spatial heterogeneity and to gain deeper insights into the dynamic evolution of urban form and structure.

Lastly, the local Moran's I index was used to extract sub-centers from HH clusters in spatial clusters of LandScan data in this study. In some cities, due to the dispersed nature of urban populations prior to 2010, the density of LandScan cells was insufficient to yield statistically significant HH clusters, resulting in missing sub-center data for those years. Therefore, our future research will refine the computational precision and enrich the early-stage computational data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is funded by The Hong Kong Polytechnic University (UGC) [1-WZ43].

Notes on contributors

Xinyue Gu is a PhD candidate affiliated with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, China. Her research interests are data science, sustainable development, and human-environment interaction. She specializes in quantifying urban polycentricity using complex network methods, with applications in urban planning and policy evaluation.

Mengni Yu is a master student affiliated with the Thrust of Urban Governance and Design, The Hong Kong University of Science and Technology (Guangzhou), China. Her research interests are Urban morphology, regional planning, and spatial governance. Her work explores the socio-economic drivers of polycentric urban development, particularly in Chinese megacities.

Xintao Liu is an Associate Professor in the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, China. His research interests include geospatial artificial intelligence, urban mobility, and network modeling. His research is centered around GIScience, transportation geography, and complex networks. He specializes in integrating GeoAI into modeling human behavior and analyzing human-environment interaction in cyber and physical spaces.

ORCID

Xinyue Gu http://orcid.org/0000-0002-8166-6618
Mengni Yu http://orcid.org/0009-0001-9429-1296
Xintao Liu http://orcid.org/0000-0002-7323-9878

Data availability statement

The specific version of ArcGIS Pro was used to process and generate this dataset, as well as the visualization of the figures in this paper. All data, figures and relevant results are publicly available on Zenodo entitled "Chinese Urban Polycentricity Dynamics: A City-Level Dataset of Population Subcenters (2001–2021)": https://doi.org/10.5281/zenodo.14279505.

References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Baek, Y., & Joo, H. (2022). A study on the spatial structure of the Bu-Ul-Gyeong megacity using the city network paradigm. *Sustainability*, *14*(23). https://doi.org/10.3390/su142315845

Biau, D. J., Jolles, B. M., & Porcher, R. (2010). P value and the theory of hypothesis testing: An explanation for new researchers. *Clinical Orthopaedics and Related Research*, *468*(3), 885. https://doi.org/10.1007/s11999-009-1164-4

Bridge, G., & Watson, S. (2008). *A companion to the city*. John Wiley & Sons.

Burger, M., & Meijers, E. (2012). Form follows function? Linking morphological and functional polycentricity. *Urban Studies*, *49*(5), 1127–1149. https://doi.org/10.1177/0042098011407095

Cao, W., Dong, L., Wu, L., & Liu, Y. (2020). Quantifying urban areas with multi-source data based on percolation theory. *Remote Sensing of Environment*, *241*, 111730. https://doi.org/10.1016/j.rse.2020.111730

Chen, X., Zhang, S., & Ruan, S. (2021). Polycentric structure and carbon dioxide emissions: Empirical analysis from provincial data in China. *Journal of Cleaner Production*, *278*, 123411. https://doi.org/10.1016/j.jclepro.2020.123411

Chiaradia, A., Hillier, B., Schwander, C., & Wedderburn, M. (2012). Compositional and urban form effects on centres in Greater London. *Proceedings of the Institution of Civil Engineers - Urban Design and Planning*, *165*(1), 21–42. https://doi.org/10.1680/udap.2012.165.1.21

Crevoisier, O., & Rime, D. (2021). Anchoring urban development: Globalisation, attractiveness and complexity. *Urban Studies*, *58*(1), 36–52. https://doi.org/10.1177/0042098019889310

Fang, C. (2015). Important progress and future direction of studies on China's urban agglomerations. *Journal of Geographical Sciences*, *25*(8), 1003–1024. https://doi.org/10.1007/s11442-015-1216-5

Giudici, L. G., L, J., & Colombo, C. (2024). Density-based in-orbit collision risk model valid for any impact geometry. *Acta Astronautica*, *219*, 785–803. https://doi.org/10.1016/j.actaastro.2024.03.067

Gu, X., Tang, X., Chen, T., & Liu, X. (2024). Predicting the network shift of large urban agglomerations in China using the deep-learning gravity model: A perspective of population migration. *Cities*, *145*, 104680. https://doi.org/10.1016/j.cities.2023.104680

Gu, X., Yu, M., & Liu, X. (2025, January 1). *Chinese urban polycentricity dynamics: A city-level dataset of population subcenter (2001–2021)*. Zenodo. https://zenodo.org/records/14279505

Halbert, L., Convery, F. J., & Thierstein, A. (2006). Reflections on the polycentric metropolis. *Built Environment*, *32*(2), 110–113. https://doi.org/10.2148/benv.32.2.110

Harrison, J., Hoyler, M., Derudder, B., Liu, X., & Meijers, E. (2023). Governing polycentric urban regions. *Territory, Politics, Governance*, *11*(2), 213–221. https://doi.org/10.1080/21622671.2022.2083011

He, Q., Zeng, C., Xie, P., Tan, S., & Wu, J. (2019). Comparison of urban growth patterns and changes between three urban agglomerations in China and three metropolises in the USA from 1995 to 2015. *Sustainable Cities and Society*, *50*, 101649. https://doi.org/10.1016/j.scs.2019.101649

Hua, J., & Sun, B. (2015). Measuring the economic performance of polycentric spatial structure in Chinese metropolitan areas. *Urban Issues*, *9*, 68–73. https://doi.org/10.13239/j.bjsshkxy.cswt.150910

Huang, D., Liu, Z., Zhao, X., & Zhao, P. (2017). Emerging polycentric megacity in China: An examination of employment subcenters and their influence on population distribution in Beijing. *Cities*, *69* (September), 36–45. https://doi.org/10.1016/j.cities.2017.05.013

Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jia, J., & Christakis, N. A. (2020). Population flow drives spatial-temporal distribution of COVID-19 in China. *Nature*, *582*(7812), 389–394. https://doi.org/10.1038/s41586-020-2284-y

Jiang, N., & Wei, J. (2024). How does regional integration policy affect urban resilience? Evidence from urban agglomeration in China. *Environmental Impact Assessment Review*, *104*, 107298. https://doi.org/10.1016/j.eiar.2023.107298

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), 483–499. https://doi.org/10.1086/261763

Kucukpehlivan, T., Cetin, M., Aksoy, T., Kurkcuoglu, M. A. S., Cabuk, S. N., Pekkan, O. I., Cabuk, A., & Cabuk, A. (2023). Determination of the impacts of urban-planning of the urban land area using GIS hotspot analysis. *Computers and Electronics in Agriculture*, *210*, 107935. https://doi.org/10.1016/j.compag.2023.107935

Lee, B. (2007). "Edge" or "edgeless" cities? Urban spatial structure in US metropolitan areas, 1980 to 2000. *Journal of Regional Science*, *47*(3), 479–515. https://doi.org/10.1111/j.1467-9787.2007.00517.x

Li, W., Sun, B., & Zhang, T. (2019). Spatial structure and labour productivity: Evidence from prefectures in China. *Urban Studies*, *56*(8), 1516–1532. https://doi.org/10.1177/0042098018770077

Li, Y., & Du, R. (2022). Polycentric urban structure and innovation: Evidence from a panel of Chinese cities. *Regional Studies*, *56*(1), 113–127. https://doi.org/10.1080/00343404.2021.1886274

Li, Y., & Liu, X. (2018). How did urban polycentricity and dispersion affect economic productivity? A case study of 306 Chinese cities. *Landscape and Urban Planning*, *173*, 51–59. https://doi.org/10.1016/j.landurbplan.2018.01.007

Liu, X., Derudder, B., & Wu, K. (2016). Measuring polycentric urban development in China: An intercity transportation network perspective. *Regional Studies*, *50*(8), 1302–1315. https://doi.org/10.1080/00343404.2015.1004535

Liu, X., & Wang, M. (2016). How polycentric is urban China and why? A case study of 318 cities. *Landscape and Urban Planning*, *151*, 10–20. https://doi.org/10.1016/j.landurbplan.2016.03.007

Melchiorri, M., Mari Rivero, I., Florio, P., Schiavina, M., Krasnodebska, K., Politis, P., & Sulis, P. (2024). *Stats in the city-the GHSL urban centre database 2025–Public release GHS-UCDB* (R2024 ed.). Publications Office of the European. https://doi.org/10.2760/5259274

Mills, D. E. (1981). Growth, speculation and sprawl in a monocentric city. *Journal of Urban Economics*, *10*(2), 201–226. https://doi.org/10.1016/0094-1190(81)90015-2

Peng, X. (2023). The household registration system reform and social integration of the migrant population. In C. W. Su & Y. Xie (Eds.), *Household registration system reform in China's megacities* (pp. 135–175). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-5292-2_6

Rauhut, D. (2017). Polycentricity-one concept or many? *European Planning Studies*, *25*(2), 332–348. https://doi.org/10.1080/09654313.2016.1276157

Sat, N. A. (2018). Polycentricity in a developing world: A micro-regional analysis for morphological polycentricity in Turkey. *Geoscape*, *12*(2), 64–75. https://doi.org/10.2478/geosc-2018-0007

Taylor, J. A., & Bates, T. R. (2013). A discussion on the significance associated with Pearson's correlation in precision agriculture studies. *Precision Agriculture*, *14*(5), 558–564. https://doi.org/10.1007/s11119-013-9314-9

Tepanosyan, G., Sahakyan, L., Zhang, C., & Saghatelyan, A. (2019). The application of local Moran's I to identify spatial clusters and hot spots of Pb, Mo and Ti in urban soils of Yerevan. *Applied Geochemistry*, *104*, 116–123. https://doi.org/10.1016/j.apgeochem.2019.03.022

Thiel, A., Garrick, D. E., & Blomquist, W. A. (Eds.). (2019). *Governing complexity: Analyzing and applying polycentricity*. Cambridge University Press.

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), 9. https://doi.org/10.3390/ijerph19095153

Volgmann, K., & Münter, A. (2022). Understanding metropolitan growth in German polycentric urban regions. *Regional Studies*, *56*(1), 99–112. https://doi.org/10.1080/00343404.2020.1807491

Wang, M., & Debbage, N. (2021). Urban morphology and traffic congestion: Longitudinal evidence from US cities. *Computers, Environment and Urban Systems*, *89*, 101676. https://doi.org/10.1016/j.compenvurbsys.2021.101676

Wang, M., Derudder, B., & Liu, X. (2019). Polycentric urban development and economic productivity in China: A multiscalar analysis. *Environment & Planning A: Economy & Space*, *51*(8), 1622–1643. https://doi.org/10.1177/0308518X19866836

Wang, P., Liu, X., & Li, Y. (2021). Spatial structure, city scale and innovation performance of Chinese cities. *China Industrial Economy*, *5*, 114–132. https://doi.org/10.19581/j.cnki.ciejournal.2021.05.017

Wang, Y., Lv, W., Wang, M., Chen, X., & Li, Y. (2023). Application of improved Moran's I in the evaluation of urban spatial development. *Spatial Statistics*, *54*, 100736. https://doi.org/10.1016/j.spasta.2023.100736

Wang, Y., & Niu, X. (2023). Polycentricity measurement of China's urban agglomerations considering internal and external connections. *Frontiers of Urban and Rural Planning*, *1*(1), 19. https://doi.org/10.1007/s44243-023-00023-w

Yang, H., Dobruszkes, F., Wang, J., Dijst, M., & Witte, P. (2018). Comparing China's urban systems in high-speed railway and airline networks. *Journal of Transport Geography*, *68*, 233–244. https://doi.org/10.1016/j.jtrangeo.2018.03.015

Yin, H., Xiao, R., Fei, X., Zhang, Z., Gao, Z., Wan, Y., Guo, Y., Jiang, X., Cao, W., & Guo, Y. (2023). Analyzing "economy-society-environment" sustainability from the perspective of urban spatial structure: A case study of the Yangtze River delta urban agglomeration. *Sustainable Cities and Society*, *96*, 104691. https://doi.org/10.1016/j.scs.2023.104691

Yu, C., Long, H., Zhang, X., Tu, C., Tan, Y., Zhou, Y., & Zang, C. (2023). Regional integration and city-level energy efficiency: Evidence from China. *Sustainable Cities and Society*, *88*, 104285. https://doi.org/10.1016/j.scs.2022.104285

Yu, H., Yang, J., Li, T., Jin, Y., & Sun, D. (2022). Morphological and functional polycentric structure assessment of megacity: An integrated approach with spatial distribution and interaction. *Sustainable Cities and Society*, *80*, 103800. https://doi.org/10.1016/j.scs.2022.103800

Yue, W., Wang, T., Liu, Y., Zhang, Q., & Ye, X. (2019). Mismatch of morphological and functional polycentricity in Chinese cities: An evidence from land development and functional linkage. *Land Use Policy*, *88*, 104176. https://doi.org/10.1016/j.landusepol.2019.104176

Zhang, C., Luo, L., Xu, W., & Ledwith, V. (2008). Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. *Science of the Total Environment*, *398*(1), 212–221. https://doi.org/10.1016/j.scitotenv.2008.03.011

Zhang, T., Sun, B., & Li, W. (2017). The economic performance of urban structure: From the perspective of polycentricity and monocentricity. *Cities*, *68*, 18–24. https://doi.org/10.1016/j.cities.2017.05.002

Zhang, Y., Li, Y., Chen, Y., Liu, S., & Yang, Q. (2022). Spatiotemporal heterogeneity of urban land expansion and urban population growth under new urbanization: A case study of Chongqing. *International Journal of Environmental Research and Public Health*, *19*(13), 13. https://doi.org/10.3390/ijerph19137792