International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 14, Issue 7 [July 2025] PP: 61-69

Spatialization of Population Distribution in Zhengzhou City Based on Multi-Source Data Fusion

Author Xirui Wen[1] Yu Wang[1,*] Shengnan Zhao[1,2] Guoqing Fan[1] Chenqing Xu[1] Menghao Xuan[1]
[1]*School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, 454000, China*
[2]*Jiaozuo Municipal Natural Resources and Planning Bureau Shanyang Service Center, Jiaozuo, 454000, China*
Corresponding Author: Yu Wang, wangyuchn@whu.edu.cn

ABSTRACT: *High-precision population distribution data are crucial for urban management and the optimization of resource allocation. The widely used WorldPop population dataset is limited by its low update frequency and insufficient accuracy at fine scales within cities. This study integrates multi-source geospatial data, including nighttime light data(NTL), building volume density, mobile phone signaling data, points of interest (POI), and road networks. The entropy weight method and the analytic hierarchy process (AHP) are employed to calculate the weights of each indicator. These methods are used to spatially allocate the 2024 permanent population statistics of Zhengzhou City, creating a 50 m × 50 m gridded population product covering five districts. The accuracy of the product is validated at three spatial scales. First, a regression analysis is conducted between the population statistics of 15 schools in Zhengzhou City from their official websites in 2024 and our product, yielding an R^2 of 0.93 (compared to 0.85 for WorldPop). Second, the product is validated using the population statistics of 31 sub-districts in Zhengzhou City, with an R^2 of 0.84 (compared to 0.55 for WorldPop). Finally, when comparing the aggregated population of this product with the statistical population in each administrative area, it shows an R^2 of 0.96, higher than WorldPop's 0.66. The validation results demonstrate that the integration of multi-source data effectively addresses the bias issues associated with single data sources in representing population distribution, providing a reliable medium for high-precision population spatialization. This method not only offers a valuable reference for urban planning practices but also enhances urban governance and the optimization of public resource allocation.*

Keywords: *Multi-source data fusion; Zhengzhou City; Population spatialization; Entropy weight method; Analytic hierarchy process (AHP); Spatial analysis*

Date of Submission: 02-01-2025 Date of acceptance: 12-07-2025

I. INTRODUCTION

Under the background of new-type urbanization, the accuracy of population spatial distribution is crucial for urban development and the efficiency of national land space governance, including urban planning and infrastructure planning[1]. Traditional population census data, although considered the most authoritative data source, have significant limitations. The population statistics are confined to the administrative divisions of townships or sub-districts, providing only aggregated numerical results. Moreover, the census process is time-consuming, labor-intensive, and costly, making it difficult to meet the timely demands of academic research[2]. With the development of remote sensing and GIS, multi-source satellite data have been increasingly and extensively applied in urban planning and resource monitoring. Among them, NTL data, which can represent the gradient of regional economic activity and has been proven to be highly correlated with population[3], have been widely used in population spatialization studies by scholars such as Lai Xianlong[4] and Zhao Xin[5].

In this context, the WorldPop population dataset integrates multi-source geospatial data, including satellite imagery, road networks, and land use, and applies machine learning methods such as geographically weighted regression[6, 7], to transform national or provincial statistical data into high-resolution continuous surfaces. The LandScan population dataset utilizes data on nighttime lights, land use, topography, and transportation networks, primarily relying on the random forest algorithm to spatialize finer county-level statistical information[8, 9], resulting in a continuous gridded dataset that reflects the dynamics of daily population changes. However, these models have significant shortcomings, showing notable population allocation errors in areas such as roads, non-residential areas, and urban villages. On the one hand, non-residential light sources such as road lighting and billboards cause population to be incorrectly allocated due to the interference of NTL data. On the other hand, urban villages with high building density but weak nighttime light signals suffer from under-allocation of population due to the lack of effective information[10]. Meanwhile, existing models generally ignore the vertical population distribution characteristics within high-rise buildings. The combined effect of these two factors further exacerbates the population spatial allocation bias in high-rise building areas of urban villages and high-density urban centers[11, 12].

With the development of the Internet, the emergence of multi-source geospatial big data has provided a "three-dimensional" perception dimension for population spatialization: mobile phone signaling data, by providing high-precision, high-frequency, and full-space coverage of population dynamics, can capture the characteristics of population distribution and mobility[13-15]; Point of Interest (POI) data, containing rich spatial semantic information, can be used for urban functional zoning and identification[16, 17]; building height and three-dimensional shape data can interpret the vertical population carrying capacity[18]. Therefore, this study incorporates mobile phone signaling data to address the detection of nighttime population aggregation in areas without POI annotations; introduces building data with height to constrain the geographical scope of population distribution and quantify differences in vertical floor density; adds POI data to enhance the retention of detailed population distribution information; and supplements NTL data to indicate the intensity of economic activity and the extent of the built-up area. The weights of each data source are determined through an entropy weight-analytic hierarchy process(EWM-AHP) model to create a 50-m resolution permanent population spatial distribution dataset for the region in 2024. Finally, through multi-dimensional cross-validation based on the open-source WorldPop dataset, the seventh national population census data, and population data from closed areas such as universities, the accuracy and statistical reliability of the population spatial distribution results of the dataset are systematically evaluated.

This study integrates multi-source geospatial data to generate a 50-m resolution population spatialization dataset for Zhengzhou City in 2024. This achievement not only establishes a refined population framework for complex scenarios in megacities but also provides data-driven decision support for spatial governance, emergency resource allocation, and public safety early warning in high-density urban areas. Additionally, it offers a paradigm reference for other cities and related research, holding significant practical value for promoting urban management and sustainable development.

II. MATERIALS AND METHODS

STUDY AREA

Zhengzhou City, located in the Central Plains of China, serves as the capital of Henan Province, a national central city, an important node of the Belt and Road Initiative, and a national comprehensive transportation hub. It is a core area for political, economic, and cultural activities in Henan Province. This study focuses on five main urban districts of Zhengzhou City, namely Huiji District, Jinshui District, Erqi District, Zhongyuan District, and Guancheng Hui District (Figure 1). These districts are the main areas for commercial activities, economic cores, industrial parks, and population aggregation in Zhengzhou City. The highly concentrated population in these districts is mainly due to the excellent transportation accessibility, dense employment opportunities, and continuously growing economic vitality. Based on multi-source geospatial data, such as NTL data and building data, this study conducts population spatialization research in these five districts to explore the interactive mechanisms of multidimensional factors on population distribution patterns. The results will provide scientific decision-making support for optimizing urban functional zoning, improving the efficiency of public service facility allocation, and perfecting the transportation network layout.

DATA SOURCES AND PROCESSING

The data used in this study include the Zhengzhou City Statistical Yearbook data, boundary data, Sentinel-2A imagery maps, data related to the seventh national population census, WorldPop population grids, NTL data, building data (including building height)[19], mobile phone signaling data, POI data, and road network data, as shown in Table 1.

Based on the functional attributes of buildings, non-residential patches such as shopping malls, factories, and government office buildings were excluded, while residential buildings and university dormitories, which have a strong correlation with permanent population, were retained. To address the issue of missing building census data in urban villages, compensation rules were set according to the "Rural and Village Residential Design Code": unextracted residential buildings were uniformly estimated to have a floor height of 3 meters and an average of 3 floors to ensure the continuous expression of population carrying capacity in the urban-rural transition zones. The building volume density (unit: m^3/km^2) was calculated by combining the building footprint area and floor height data, and a 2024 building volume density grid dataset for the study area was constructed using

ArcGIS. To avoid the influence of non-residential light sources such as streetlights on population allocation in NTL data, the NTL data were masked with building volume greater than zero grid data to retain the corresponding areas.

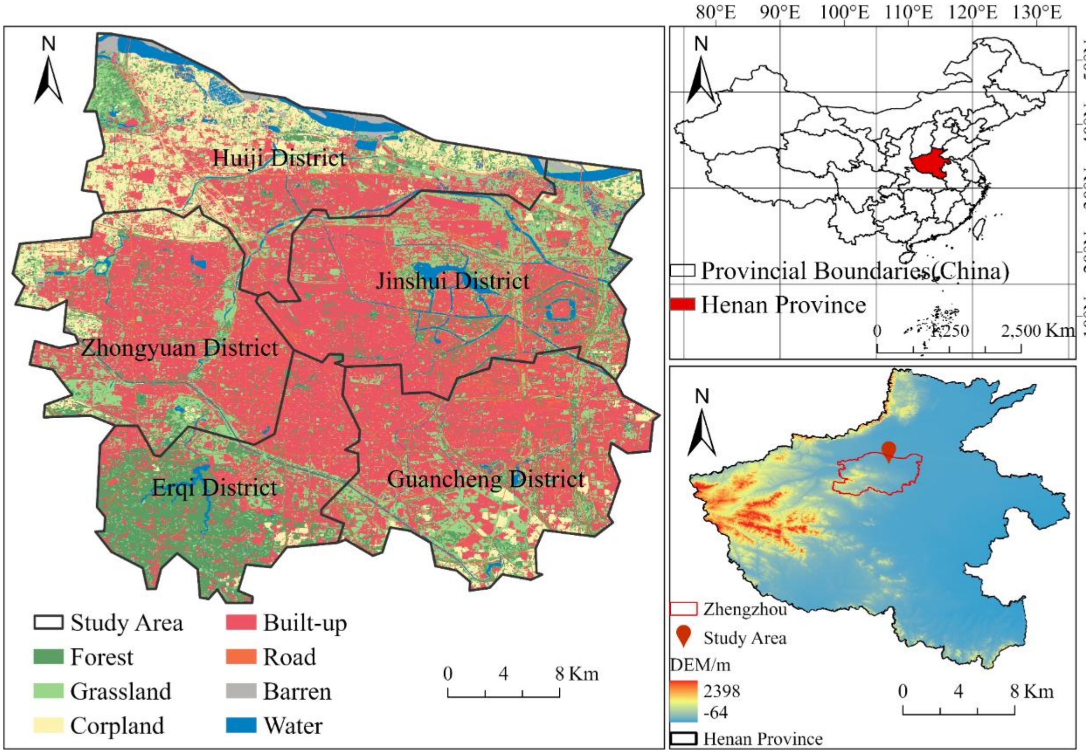


Fig. 1. Location map of the study area

Table 1. Data sources.

Data Name	Year	Data sources
NTL data	2024	SDGsat1 https://www.sdgsat.ac.cn
Mobile phone signaling data	2024	Unicom
POI data	2024	Open Steet Map（OSM） https://openstreetmap.maps.arcgis.com/home/index.html
Building data	2024	《Scientific Data》 https://figshare.com/articles/dataset
Boundary data	2024	OSM
Image of Sentinel-2A	2024	Google Earth Engine https://developers.google.cn/earth-engine
Population data by district	2024	Zhengzhou Statistical Yearbook 2024 https://tjj.zhengzhou.gov.cn
Worldpop	2024	Worldpop official website https://hub.worldpop.org/geodata/summary?id=50680
Seven popular data	2020	Zhengzhou Statistics Bureau
Seventh National Census Population Raster Products	2020	Professor Chen Yuehong's team https://doi.org/10.6084/m9.figshare.24916140.v1
University population	2024	Statistics from the school's official website
Road network data	2024	OSM

This study acquired POI data from the OSM platform. Based on the close correlation between POI types and population distribution, 12 key categories—including transportation facilities—were identified. Subsequently, seven experts with extensive backgrounds in urban planning and population geography were invited. These experts performed pairwise comparison scoring on these 12 key category indicators using the AHP, thereby deriving the weight coefficients for each category (Table 2). Finally, a comprehensive POI indicator was constructed through weighted summation (POI density of each category × corresponding weight). This comprehensive indicator, together with three other indicators, was integrated into a multi-source data weighting calculation framework. This integration significantly enhanced the accuracy of the data analysis[Error! Reference source not found.].

Table 2. Classification and Statistics of POIs.

POI type	Weight	POI type	Weight
Sports and fitness	0.065	Business Residences	0.106
Healthcare	0.079	Tourist attractions	0.056
Leisure and Entertainment	0.074	Science and Education Culture	0.082
Life services	0.088	Hotel accommodation	0.071
Transport infrastructure	0.095	Shopping consumption	0.099
Companies and Enterprises	0.089	Food and Dining	0.096

METHODS

This study adopts a coupled model of the entropy weight method and the AHP to integrate the advantages of both objective and subjective weighting. The entropy weight method, based on information entropy theory, objectively calculates weights using the inherent discrete degree of the indicator data[Error! Reference source not found.].However, it is sensitive to data distribution and overlooks domain knowledge. AHP, on the other hand, quantifies the relative importance of indicators based on expert experience but may introduce bias due to subjective preferences[Error! Reference source not found.-Error! Reference source not found.]; The combination of these two methods can complement their respective strengths. To eliminate the dimensional differences among multi-source indicators and avoid the interference of zero values in the entropy weight method calculation, the Min-Max normalization and translation method is applied to standardize the original data. First, the dimensionless transformation of four types of indicators—building data, NTL, mobile phone signaling data, and POI data—is conducted using the range method (Equation 1):

$$X_{ij}^{*} = \frac{X_{ij} - min(X_j)}{max(X_j) - min(X_j)} (i = 1,2, \dots, n; j = 1,2, \dots, 4) \tag{1}$$

In the formula, is the original observation value of the *j*-th indicator for the *i*-th spatial unit, and X_{ij}^{*} is the normalized result. To address the issue of zero values in the normalized data, a small translation value (ϵ=0.001) is applied to construct a non-zero dataset $X_{ij}' = X_{ij}^{*} + \in$, ensuring the mathematical stability of subsequent logarithmic operations to ensure the mathematical stability of subsequent logarithmic operations.

The entropy weight method calculation is then executed on the translated data in the ArcGIS platform through the spatial statistics module and the raster calculator, following these steps:

(1) Construction of the Indicator Proportion Matrix(P_{ij}): Calculate the distribution proportion of each spatial unit in a single indicator dimension:

$$P_{ij} = \frac{X_{ij}'}{\sum_{i=1}^{n} X_{ij}'} \tag{2}$$

(2) Quantification of Information Entropy(e_j): Assess the information discrete degree of indicators based on Shannon entropy theory:

$$e_j = -K \times \sum_{i=1}^{n} (P_{ij} \times \ln(P_{ij})), \quad K = \frac{1}{\ln(n)} \tag{3}$$

represents the proportion of the *i*-th sample value under the *j*-th indicator, is the translated value of the indicators, and *n* is the number of objects involved in the evaluation. The smaller the value, the higher the information utility of the indicator.

(3) Generation of Utility Weights(d_j) : Calculate the indicator weights through the inverse calculation of information entropy:

$$d_j = 1 - e_j, \quad w_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \tag{4}$$

(4) Spatial Weighted Fusion: Perform spatial convolution of the normalized original data with the weight matrix to generate a comprehensive population distribution potential score S_m:

$$S_m = \sum_{j=1}^{4} w_j X_{ij} \tag{5}$$

(5) Normalization of Population Allocation Coefficients: Convert the potential score into a proportional coefficient a:

$$a = \frac{S_m}{\sum_{i=1}^{n} S_m} \tag{6}$$

To reconcile the weight allocation bias between data-driven and expert experience, this study constructs a coupled subjective-objective weight calculation model using the EWM-AHP method. First, seven experts with extensive backgrounds in urban planning and population geography were selected to conduct pairwise importance evaluations of the four types of indicators (nighttime lights, buildings, POI, and mobile phone signaling data) based on Saaty's 1-9 scale criterion. Individual judgment matrices were generated, and the consistency ratio was validated to be less than 0.1 to ensure the logical consistency of group decision-making. Next, based on the equilibrium principle of game theory, a combined optimization objective function was constructed, and the optimal coupling coefficients α (EWM) = 0.63 and β (AHP) = 0.37 were obtained using the Lagrange multiplier method. This indicates that the weight allocation leans more towards objective data-driven approaches. To compensate for and correct the missing parts of the building data, the building weights were adjusted accordingly. The final weights were mobile phone signaling data (0.320), nighttime lights (0.285), buildings (0.199), and POI data (0.196). The four indicator data grids were allocated and summed according to the final weights to obtain the population spatialization grid.

III. RESULTS AND DISCUSSION

ZHENGZHOU 50M POPULATION GRID PRODUCTS

Based on the population grid analysis results (Figure 2), the population distribution in the main urban area of Zhengzhou shows significant spatial heterogeneity. The population is mainly concentrated in the southern part of Huiji District, the western part of Jinshui District, the northwestern part of Guancheng Hui District, the northeastern part of Erqi District, and the southeastern and central parts of Zhongyuan District. This spatial pattern is basically consistent with the trend reflected in the WorldPop dataset and also matches the district-level population statistical data, that is, Jinshui District and Erqi District have the highest population density, while Huiji District has the lowest population density. However, there are significant differences between this product and WorldPop in some areas (red-boxed areas in Figure 3). Specifically, in the densely packed urban villages of Guancheng Hui District and Huiji District, this study effectively compensates for the insufficient ability of traditional NTL data to identify informal residences by integrating POI and mobile phone signaling data, thereby increasing the population allocation intensity in these areas. In contrast, in the industrial-dominated Zhongyuan District, this product actively excludes the interference of road and non-residential light sources by strictly limiting the allocation scope of residential areas and using a more refined allocation method, thus avoiding large continuous areas of high population distribution. As a result, the estimation results are lower than those of WorldPop. The validation results show that these targeted data processing methods significantly improve the accuracy of the population distribution results of this product, making it closer to the actual statistical data than WorldPop.

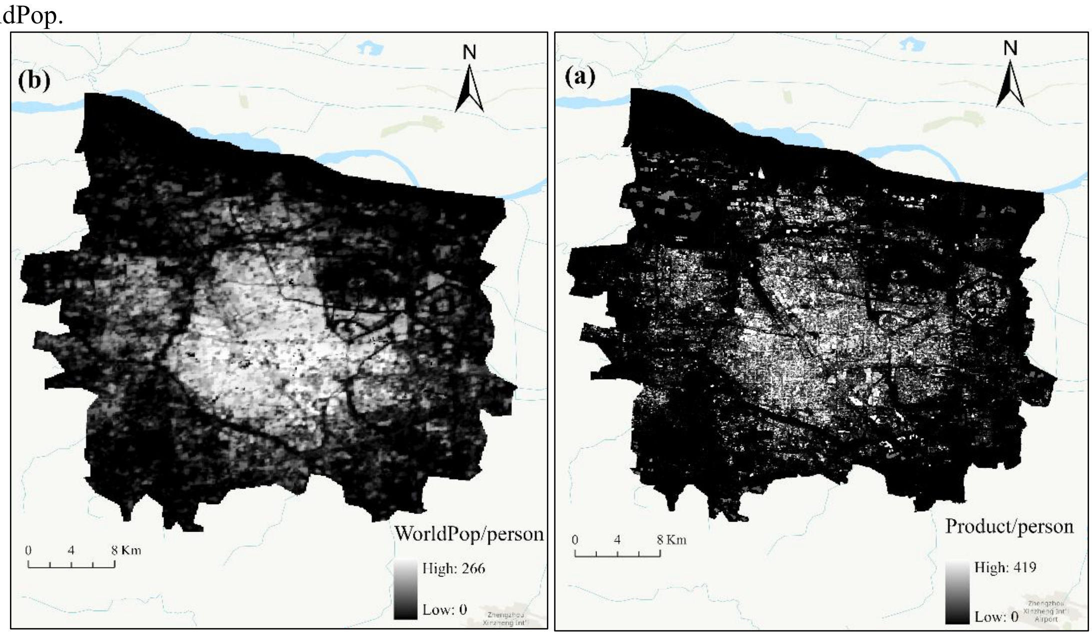


Fig. 2 Population raster plots of this study's product (a) and WorldPop (b)

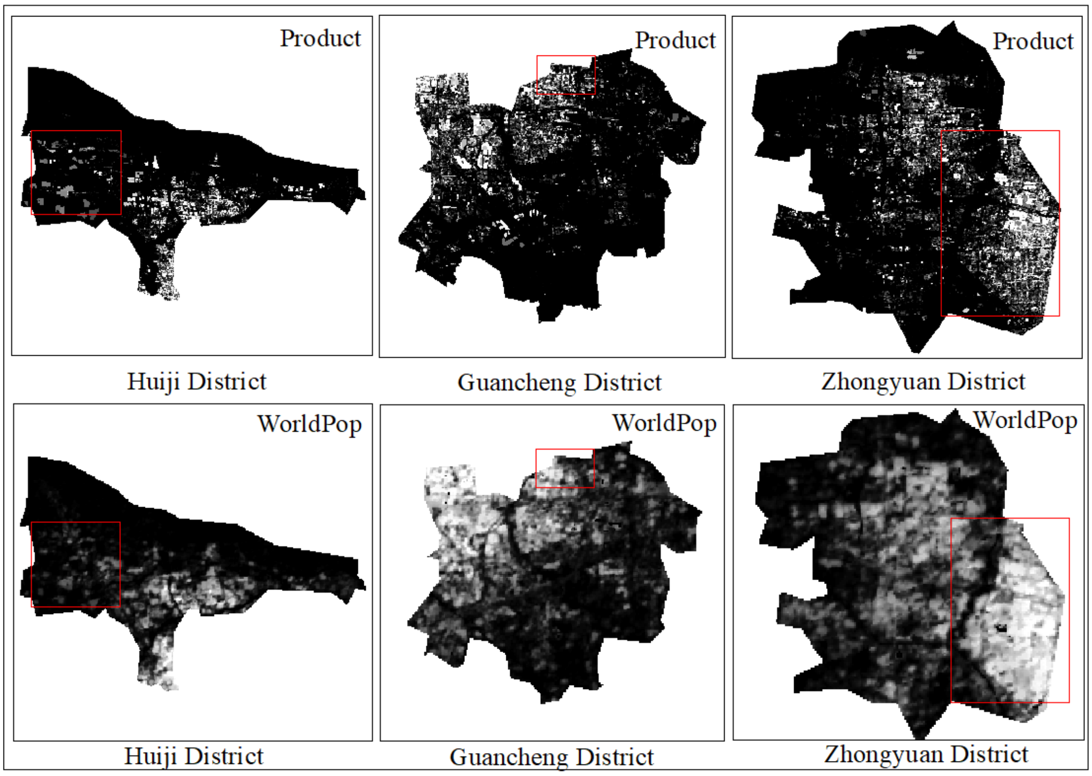


Fig. 3 Zoning Detail Comparison

ACCURACY ASSESSMENT

Based on the population data of each district in Zhengzhou City from the 2024 Zhengzhou City Statistical Yearbook, this study extracted the population distribution results within the study area and validated the accuracy of the population distribution against the WorldPop dataset. The comparative analysis indicates that our product significantly outperforms the WorldPop data in terms of spatial matching and statistical consistency: at the school scale, the goodness of fit (R^2) between our product and the statistical yearbook data reaches 0.93 (compared to 0.85 for WorldPop), as shown in Figure 4 (a) and (b). In the absence of sub-district data for Zhengzhou City in 2024, we based our curve fitting on the 2020 sub-district data, considering the increasing trend in total population from 2021 to 2024 as documented in the Zhengzhou City Statistical Yearbook, which reflects the regional GDP growth rate, changes in registered population, and the evolution of industrial employment structure. Although the total population shows an increasing trend, the spatial agglomeration pattern and overall change trend remain relatively stable. The R^2 value between our population distribution results and the statistical yearbook data at the sub-district scale reaches 0.84 (compared to 0.55 for WorldPop), as shown in Figure 4 (c) and (d). Following the development trend, the 50-m resolution results are expected to have a certain level of accuracy. At the district scale, the R^2 value reaches 0.96 (compared to 0.66 for WorldPop), as shown in Figure 4 (e) and (f).

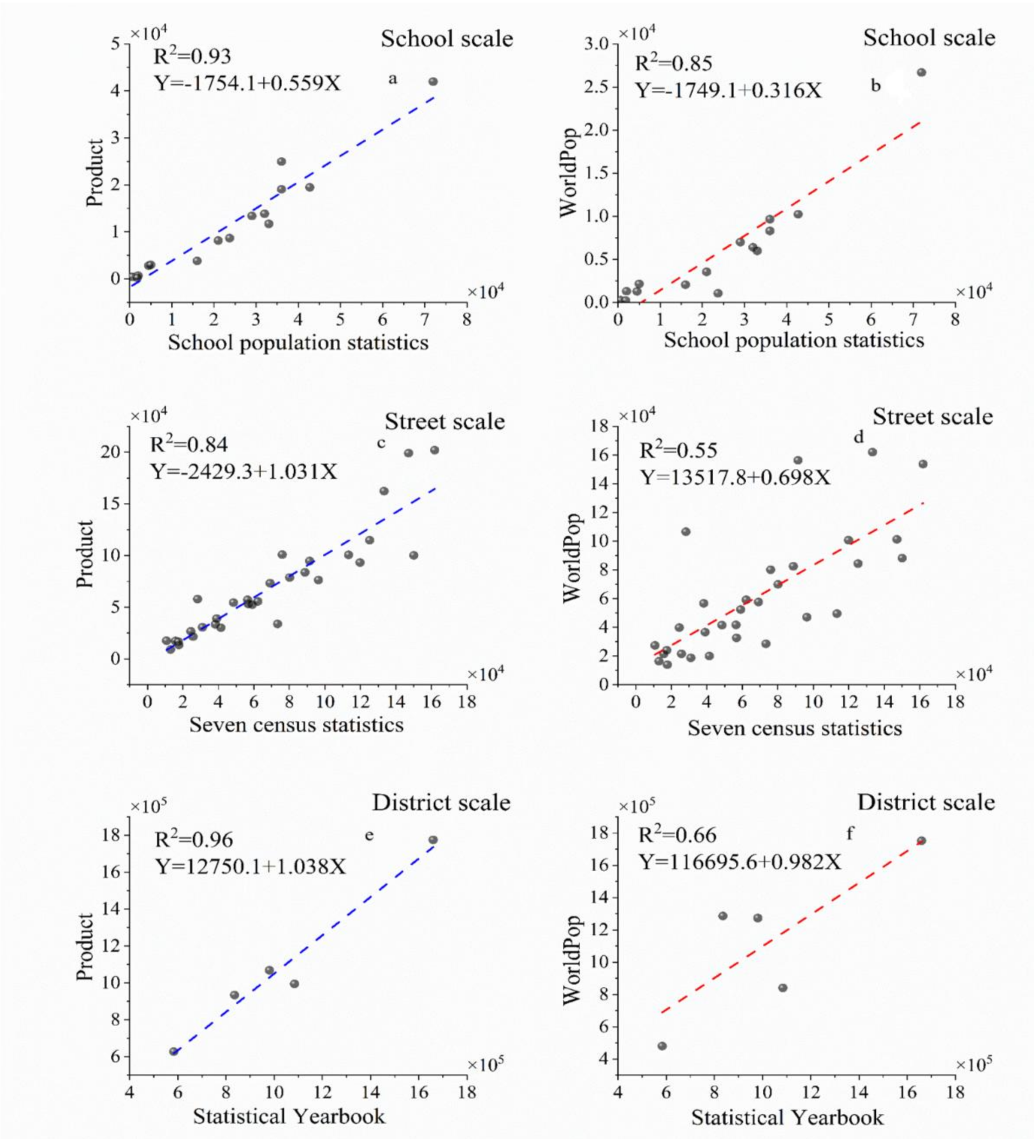
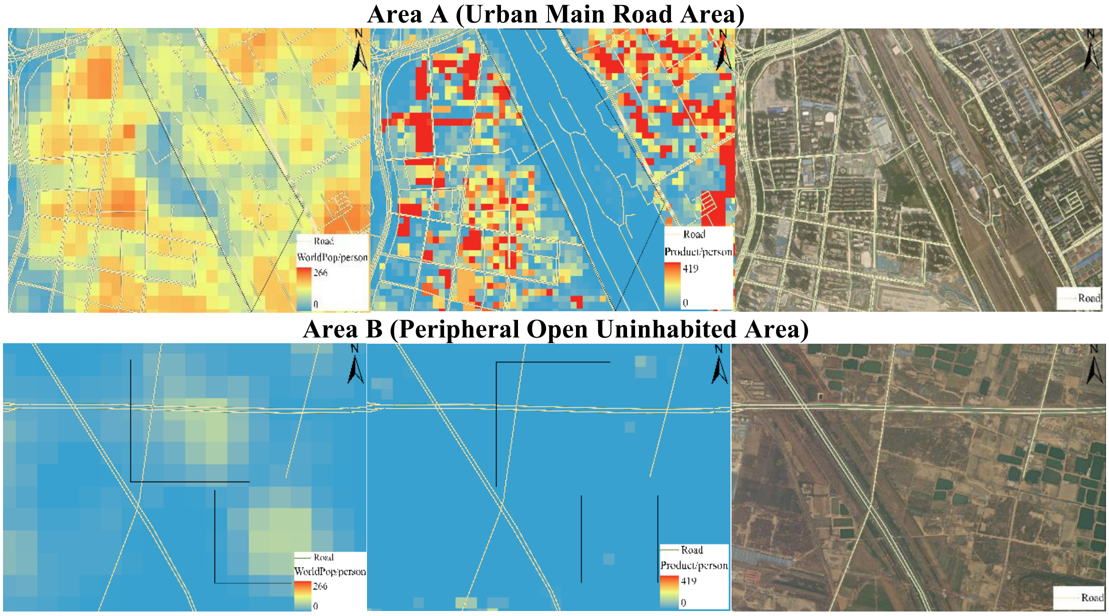


Fig. 4 Validation and fitting curves at different scales

VALIDATION OF MULTI-SOURCE DATA INTEGRATION

This study indicates that while NTL data can effectively identify urban population hotspots, it may lead to estimation biases in population distribution due to light pollution in non-residential areas. The integration of building vector data significantly reduces the oversaturation effect of NTL data. However, the current building data have insufficient spatial recognition accuracy in residential areas such as urban villages and shantytowns. Therefore, it is necessary to further integrate the distribution of livelihood-related facilities from POI data, such as "commercial residences"and "transportation facilities",the residential identification features from mobile phone signaling data, and the accessibility parameters from road network data.

In this study, two typical non-residential areas were selected for comparative analysis: urban main roads and the open, uninhabited peripheral areas of the city. The urban main roads are primarily dominated by transportation functions. Although there are no residential buildings, there are high-intensity road lighting sources. The WorldPop dataset commonly misallocates population in such areas, while our product effectively avoids the interference of non-residential light sources and does not allocate permanent population. Similarly, in the open, uninhabited peripheral areas far from the built-up areas, where there are no residential buildings, the WorldPop dataset still shows false population distribution, while our product identifies the lack of effective residential support elements in this area and does not allocate population. The comparative results of the above cases clearly demonstrate that, in terms of effectively excluding non-residential area interference and identifying true inhabited areas, our product has higher spatial allocation accuracy and reliability compared to the WorldPop dataset, as shown in Figure 5.



(a) Sentinel-2A image map (b) Product of this study (c) WorldPop dataset

Fig. 5 Partial comparison chart

SCIENTIFIC RATIONALE FOR SPATIAL SCALE SELECTION

The adoption of a 50m spatial resolution for small-scale population spatialization—rather than higher-resolution alternatives—is grounded in technical compatibility with multi-source data integration. Although 10m resolution enhances geometric precision of building contours, it introduces fragmented artifacts that disrupt the spatial continuity of residential units. Such discrete distributions impede population allocation algorithms from effectively identifying integrated residential functional units, causing deviations between weight assignments and actual settlement patterns while significantly increasing computational complexity in data fusion without substantive accuracy gains. Furthermore, 50m resolution optimizes spatial congruence with downscaled nighttime light (NTL) data. By applying residential cluster masking to NTL, interference from non-residential light sources is substantially suppressed, enhancing residential unit identification accuracy. Empirical validation confirms that 50m resolution achieves dual optimization: it finely delineates neighborhood-scale population distribution patterns while balancing algorithmic stability and computational efficiency in multi-source data fusion. This technical synergy ensures coordinated enhancement of building-patch geometric integrity and population spatial allocation precision, providing a robust spatial framework for microscale population modeling.

IV. CONCLUSION

This study developed a high-precision population spatialization model by integrating multi-source data, including nighttime light, mobile phone signaling, building height, and POI data. Mobile phone signaling data effectively detected nighttime population agglomerations in areas lacking POI annotations, supplemented population signals in light-deficient zones, and mitigated under-allocation issues. Building height data revealed the regulatory role of vertical plot ratio on population distribution, while a building mask of residential areas significantly reduced non-residential light interference in NTL data. NTL data and POI functional density jointly characterized population attraction disparities between economically active zones and life-service areas. Road network data quantified the constraining effect of transportation accessibility on population diffusion using centrality indices. To address the absence of street/community-level population statistics, an EWM-AHP (Entropy Weight Method-Analytic Hierarchy Process) coupled model was employed to calculate weights, successfully generating a 50m-resolution gridded population dataset for Zhengzhou City in 2024.

Accuracy validation demonstrated that this method significantly outperformed the WorldPop benchmark product across multiple spatial scales:At the district level, the fitting accuracy with street statistical data reached R^2= 0.96 (WorldPop: R^2= 0.66), with notable error reduction in Huiji District's low-density residential areas and Jinshui District's high-density commercial zones.For micro-scale validation, R^2reached 0.93 (WorldPop: 0.85) for 15 schools within Zhengzhou's five main administrative districts, and 0.84 (WorldPop: 0.55) for 31 street units, confirming its superior capability in capturing fine-grained population heterogeneity and gradient features around schools (density contrasts between school district housing clusters and suburban campuses).These advantages primarily stem from the 50m high-resolution data characteristics. Compared to WorldPop's 100m product, this dataset delivers more refined spatial structures and significantly enhances fine-grained population characterization in complex urban environments.

Acknowledgement: This research was supported by several grants: The Science and Technology Project of the Henan Province (Grant No. 232102320244 and No. 242102320198), Key Laboratory of Mine Spatio-Temporal Information and Ecological Restoration, MNR(Grant No. KLM202307), University Basal Research Fund of Henan Province (Grant No. NSFRF240605). The authors express their gratitude to the editors and the anonymous review.ers for their valuable comments and suggestions

REFRENCES

[1] AUBRECHT C, OZCEYLAN D, STEINNOCHER K, et al. Multi-level geospatial modeling of human exposure patterns and vulnerability indicators[J]. NATURAL HAZARDS, 2013,68(1): 147-163.
[2] X DONGSHENG, Y SONG, S P U SCHOOL OF CIVIL ENGINEERING AND ARCHITECTURE, et al. A review of population spatial distribution based on nighttime light data[J]. Remote Sensing for Land & Resources, 2019,31(03): 10-19. (in Chinese)
[3] STEVENS F R, GAUGHAN A E, LINARD C, et al. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data[J]. PLOS ONE, 2015,10(2): 22.
[4] L XIANLONG, H NIANLONG, H PENGJIN, et al. Population Spatialization in Hainan Province Based on NPP/VIIRS Nighttime Light Remote Sensing Data[J]. Bulletin of Science and Technology, 2021,37(12): 24-29. (in Chinese)
[5] Z XIN, S YINGQIANG, L YILUN, et al. Population Spatialization Based on Satellite Remote Sensing and POI Data:Guangzhou as an Example[J]. Tropical Geography, 2020,40(01): 101-109. (in Chinese)
[6] Q JIAHAO, L LIN, P DONGMEI, et al. The model of population spatial distribution considering spatial autocorrelation and spatial heterogeneity[J]. Science of Surveying and Mapping, 2022,47(07): 216-226. (in Chinese)
[7] W XI, N XIAOGANG, Z HANCHAO, et al. Population spatialization by integrating LJ1-01 nighttime light and WeChat positioning data——taking Beijing city as an example[J]. Science of Surveying and Mapping, 2022,47(02): 173-183. (in Chinese)
[8] Y C GUO, J C HUANG, H X LIN. 2020,35(01): 219-232. (in Chinese)
[9] J XUELI, X YONGLIANG, G HONGMEI, et al. Application of improved random forest model in population spatialization[J]. Bulletin of Surveying and Mapping, 2023(06): 155-160. (in Chinese)
[10] BHARTI N, TATEM A J, FERRARI M J, et al. Explaining Seasonal Fluctuations of Measles in Niger Using Nighttime Lights Imagery[J]. SCIENCE, 2011,334(6061): 1424-1427.
[11] GAUGHAN A E, STEVENS F R, HUANG Z, et al. Spatiotemporal patterns of population in mainland China , 1990 to 2010[J]. SCIENTIFIC DATA, 2016,3: 11.
[12] TATEM A J, QIU Y, SMITH D L, et al. The use of mobile phone data for the estimation of the travel patterns and imported Plasmodium falciparum rates among Zanzibar residents[J]. MALARIA JOURNAL, 2009,8: 12.
[13] Z WEIJING, W DE, X DONGCAN, et al. Dynamic characteristics of Shanghai's population distribution using cell phone signaling data[J]. Geographical Research, 2017,36(05): 972-984. (in Chinese)
[14] Z XIAOWEI, Z JIALEI. Population Structure Characteristics and Territorial Space Planning Response of Xi'an City:Based on the 7th National Population Census Data and Cellphone Signaling Data[J]. Planners, 2022,38(05): 47-54. (in Chinese)
[15] H QUANHUA, Z JIZHE, L DONGYANG. Spatial Characteristics of Rural Population Activities in Qin-Han New City Villages Based on Cellphone Signaling Analysis[J]. Modern Urban Research, 2021(02): 98-105. (in Chinese)
[16] L ZHENGLIAN, G ZHIPENG, W HUAYI, et al. Fine-Scale Population Spatialization by Synthesizing Building Data and POI Data[J]. Journal of Geomatics, 2021,46(05): 102-106. (in Chinese)
[17] L JINXIANG, T MING, GULIZIPA·MULATI, et al. Research on Population Spatialization Based on Buildings and POI Data[J]. Earthquake Research in China, 2023,39(04): 774-784. (in Chinese)
[18] Y XIN, X HAORAN, C CHUNSEN, et al. Improvements in population spatialization that take into account building information[J]. Science of Surveying and Mapping, 2025,50(04): 165-175. (in Chinese)
[19] ZHANG Y, ZHAO H, LONG Y. CMAB: A Multi-Attribute Building Dataset of China[J]. SCIENTIFIC DATA, 2025,12(1): 18.