A New Method for Building-Level Population Estimation by Integrating LiDAR, Nighttime Light, and POI Data

Hongxing Chen,1,2 Bin Wu,1,2 Bailang Yu1,2, Zuoqi Chen,3,4 Qiuosheng Wu5, Ting Lian,6 Congxiao Wang,1,2 Qiaoxuan Li,1,2 and Jianping Wu1,2

1Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China
2School of Geographic Sciences, East China Normal University, Shanghai 200241, China
3Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou University, Fuzhou 350108, China
4The Academy of Digital China, Fuzhou University, Fuzhou 350108, China
5Department of Geography, University of Tennessee, Knoxville, TN 37996, USA
6Department of Geography, The University of Hong Kong, Hong Kong 999077, China

Correspondence should be addressed to Bailang Yu; blyu@geo.ecnu.edu.cn

Received 30 December 2020; Accepted 18 April 2021; Published 6 May 2021

Building-level population data are of vital importance in disaster management, homeland security, and public health. Remotely sensed data, especially LiDAR data, which allow measures of three-dimensional morphological information, have been shown to be useful for fine-scale population estimations. However, studies using LiDAR data for population estimation have noted a nonstationary relationship between LiDAR-derived morphological indicators and populations due to the unbalanced characteristic of population distribution. In this article, we proposed a framework to estimate population at the building level by integrating POI data, nighttime light (NTL) data, and LiDAR data. Building objects were first derived using LiDAR data and aerial photographs. Then, three categories of building-level features, including geometric features, nighttime light intensity features, and POI features, were, respectively, extracted from LiDAR data, Luojia1-01 NTL data, and POI data. Finally, a well-trained random forest model was built to estimate the population of each individual building. Huangpu District in Shanghai, China, was chosen to validate the proposed method. A comparison between the estimation result and reference data shows that the proposed method achieved a good accuracy with $R^2 = 0.65$ at the building level and $R^2 = 0.79$ at the community level. The NTL radiance intensity was found to have a positive relationship with population in residential areas, while a negative relationship was found in office and commercial areas. Our study has shown that by integrating both the three-dimensional morphological information derived from LiDAR data and the human activity information extracted from POI and NTL data, the accuracy of building-level population estimation can be improved.

1. Introduction

Fine-scale population data are essential in addressing various critical social, political, and environmental issues, such as epidemic control [1] and natural disaster relief [2]. Currently, most of the available population data are concentrated on administrative units, including countries, provinces, census tracts, and blocks. The usefulness of these data is limited in urban planning, disaster management, fertility policy, climate policy, and public health [3–7] due to the low spatial resolution. Meanwhile, the increasing availability and quality of socioeconomic data and three-dimensional morphological data have the potential to improve the fine-scale population estimations, especially at the building level. Therefore, there is a critical need to develop efficient methods for building-level population estimation by integrating context information from various data sources with different structures [8, 9].

One traditional way to acquire population distribution data is through population census [10, 11]. However, population census is labor intensive and time consuming and thus constrains population investigations to a coarse temporal scale [12]. Studies have thus turned to remote-sensing techniques which conduct simultaneous observation over large
areas in a cost-effective way to estimate population [13, 14]. Early studies used aerial photographs to extract urban extents and built relationships with populations [15]. Since then, similar studies have been conducted by using different remote-sensing data, such as Landsat [16, 17] and other satellite imagery [18]. Many well-known gridded population databases, such as the LandScan [19], the Gridded Population of the World (GPW) [20], the Global Human Settlement (GHS) [21], and the WorldPop [22], have been successively established and are being given much attention by researchers. However, population estimates from satellite image-based methods remain constrained at grid level due to the weak relationship between indicators and populations and a lack of 3D information of structures [12, 23, 24].

To overcome this limitation, volumetric approaches based on LiDAR data have been proposed for fine-scale population estimation [23, 25–32]. LiDAR is an active form of remote sensing that allows us to collect height information from a point cloud [33–35]. The LiDAR-derived height and volumetric information has been found to be helpful in estimating population at census block and subblock levels [12, 23, 28, 32]. For example, Xie et al. [23] applied urban morphological metrics derived from LiDAR data to estimate block-level population in Indianapolis, USA. Similar methods can be also found in [26, 28–30]. However, a potentially nonstationary relationship between population and building volume has been noted in studies due to the uncertainty and inhomogeneity of human activities [12, 23, 25].

Recently, nighttime light (NTL) data and social-sensing data, which were considered a good proxy for human activities [24, 36–40], are becoming very popular and have been widely applied in the study of population estimation. For NTL data, studies [18, 41–48] have demonstrated that a positive correlation between NTL intensity and population was found and validated at national and city scales, while social-sensing data, such as social media, point-of-interest (POI), cell phone location data, public transportation card records, and GPS trajectories, have great potential to improve the accuracy of the fine-scale population estimation [24, 37, 40]. Although many researchers have made exhaustive studies on population estimation using multisensor remote-sensing and social-sensing data, few studies [49, 50] considered fine-scale population mapping at the building level. The main reasons mainly include two aspects. First, satellite imagery datasets have coarse spatial resolutions which make fine-scale population estimation difficult using these data. Take NTL datasets as an example, the two commonly used NTL datasets, namely, the DMSP-OLS and Suomi National Polar-Orbiting Partnership-Visible Infrared Imaging Radiometer Suite (NPP-VIIRS) data, have a spatial resolution of 30 arc-seconds and 15 arc-seconds, which are insufficient for estimating the building-level population. Second, combining three-dimensional (3D) building morphology information extracted from LiDAR data with satellite imagery and social-sensing data is difficult.

The newly launched Lujia1-01 satellite provides a new data source of NTL at 130 m resolution [51, 52]. The latest study [53] demonstrated that compared with the NPP-VIIRS, Lujia 1-01 has fewer “blooming” phenomena and appears to be more suitable for providing local variations of socioeconomic and human activities. The main objective of this study was to propose a novel framework for estimating building-level populations by incorporating Lujia1-01 NTL, LiDAR, and POI data. First, we extracted building footprints and their corresponding features using airborne LiDAR data and near-infrared remote-sensing data. Then, we applied a random forest algorithm to estimate the population at the building level based on the multidimensional features derived from Lujia1-01 NTL, POI, and LiDAR data. Next, the accuracy of the resulting population distribution was assessed by using the population database of the Shanghai Public Security Bureau. Finally, we analyzed the impact and contribution of each feature on the population estimation results.

2. Study Area and Data

2.1. Study Area. We selected the northern part of the Huangpu District in Shanghai as our study area (Figure 1). Shanghai is the financial and economic center of China and is located at the mouth of the Yangtze River. The Huangpu District, located on the western bank of the Huangpu River, is one of the most famous financial and trade centers in Shanghai. Although the administrative area of the Huangpu District is only 20 km², the permanent resident population in the area reached 654,800 in 2017. There are various types of buildings in the Huangpu District, such as skyscrapers and different types of residential buildings. There are mainly three types of residential buildings, including townhouses, walk-up buildings, and elevator buildings. Based on the residential structures, these three types of buildings can also be referred to as low-rise, mid-rise, and high-rise. Townhouses which were originally built for single families are usually two to four stories in height. Both elevator and walk-up buildings were built for multifamily purposes. Walk-up buildings usually range up to six stories high and do not feature elevators, while elevator buildings are usually 7 to 30 stories high. The various building types make the selected area a great place for validating our proposed framework.

2.2. Data

2.2.1. Lujia1-01 NTL Data. The Lujia1-01 satellite is equipped with a 4-megapixel Complementary Metal Oxide Semiconductor (CMOS) sensor and is composed of 2,048 × 2,048 unique detectors to record the weak nightlight on Earth’s surface [54]. Lujia1-01 NTL data were calibrated through an on-orbit radiation calibration method which uses daytime data to estimate the relative calibration coefficients of each Lujia1-01 nighttime detector. The 130 m resolution Lujia1-01 NTL data were downloaded from the Hubei Data and Application Center at http://www.hbeos.org.cn. As shown in Figure 2, the cloud-free Lujia1-01 NTL data on 15 July 2018 were selected in our study.

2.2.2. LiDAR Data and Near-Infrared Aerial Photographs. High-resolution airborne LiDAR data and near-infrared aerial photographs were used for deriving the building boundaries. The airborne LiDAR data (Figure 3) were
collected by an airborne system of Optech ALTM3100 in 2006. The LiDAR point measurements, which record a series of \((x, y, z)\) lists, were provided by the Geomatics Center at Shanghai Municipal Institute of Surveying and Mapping. The LiDAR data used in our study consisted of 42,547,794 points and the point density was about four points per \(m^2\).
We also obtained the 1 m spatial resolution aerial photographs with green, red, and near-infrared spectral bands in 2006 provided by the Shanghai Bureau of Surveying and Mapping. The normalized difference vegetation index (NDVI) was first computed as the difference between near-infrared and red reflectance divided by their sum. Then, the pixels with their NDVI values greater than zero were marked as vegetation, and the rest of the pixels were marked as non-vegetation [55, 56]. The distribution of vegetation is shown in Figure 4.

2.2.3. POI Data. The POI data in our study were downloaded from the Resource and Environment Data Cloud Platform (http://www.resdc.cn), a data center established by the
The extraction of building objects consists of three steps: normalized digital surface model (nDSM) generation, building boundary extraction, and building boundary regularization.

The method of Huang et al. [55] was employed to generate the nDSM from LiDAR point clouds. Firstly, a 1 m resolution digital surface model (DSM) was interpolated from the original airborne LiDAR point clouds by using the linear Triangulated Irregular Network (TIN) interpolation method. Second, the ground points were extracted from the LiDAR point clouds by using a progressive morphological filter [58], and a Digital Terrain Model (DTM) grid was then interpolated from the ground points. Finally, we created the nDSM (as shown in Figure 4) by subtracting the DTM from the DSM grid.

The next step was to extract buildings from the nDSM. First, to eliminate the influence of vegetation, vegetation pixels were masked from the original nDSM. After removing vegetation, a base building height threshold value of 3.5 m was selected to remove small nonbuilding objects from the nDSM [59]. The building objects were subsequently identified through a recursive connected component identification and indexing algorithm and dilation and erosion operations to eliminate spurious objects and smooth object boundaries [55, 56]. Finally, by tracking the boundaries of each building, the vector polygons of buildings were created, and the small vector polygons with an area of less than 100 square meters were removed.

The boundaries of the buildings extracted by the above method are jagged, which is inconsistent with the actual situation of the boundaries of buildings. Therefore, the building boundaries were further regularized to acquire straight boundary lines. The polyline compression algorithm proposed by Gribov [60] was applied to regularize buildings.

In our work, the accuracy assessment is evaluated by contrasting the result of the manual delineation method from that of the proposed automated method. Because the manual method for reference building boundaries is extremely time consuming, we randomly selected five 250 m by 250 m areas (blue boxes in Figure 4) as the validation areas for accuracy assessment. Principal criterion for the validation area selection is simple in the sense that the areas should reasonably cover different residential building types. The building reference data of the five validation areas were obtained via manual interpretation of aerial images from 2006 with 1 m spatial resolution. Detection Rate (DR) and Matched Overlay (MO) [61] were used as two accuracy evaluation indicators, wherein DR indicates the percentage of correctly detected

Table 1: The description of POI data categories.

| POI categories        | Number of records | Reclassified POI categories |
|-----------------------|-------------------|-----------------------------|
| Residential community | 1,114             | Residential                 |
| Hotel                 | 297               |                             |
| Catering              | 1,892             |                             |
| Shopping              | 69                |                             |
| Landmarks             | 29                | Service                     |
| Financial             | 562               |                             |
| Business              | 526               |                             |
| Retail                | 4,614             |                             |
| Entertainment         | 688               |                             |
| Industry              | 6                 | Industry and commerce       |
| Corporations          | 5,207             |                             |
| Public facilities     | 539               |                             |
| Media                 | 82                |                             |
| Education             | 460               |                             |
| Life service          | 315               | Institutional               |
| Healthcare            | 397               |                             |
| Sports                | 47                |                             |
| Government            | 776               |                             |
| Traffic facilities    | 1,270             |                             |
| Automobile services   | 327               | Transportation              |

Institute of Geographical Sciences and Natural Resources of the Chinese Academy of Sciences. According to the metadata, the POI data which was collected in 2016 contain 19,217 POI records. A previous study [57] has indicated that POIs represent a much finer-grained picture of land use at an aggregated level and are good proxies for multiple land uses. Although the POI categories are not the same as conventional land use types, they follow the land use codes and can reflect land use types. As a result, POI data has the possibility to describe land use. As shown in Table 1, the POI records were reclassified into five categories based on the Chinese land use classification criterion (GB/T21010-2007), including residential, service, industry and commerce, institutional, and transportation.

3. Methods

Figure 5 shows the workflow of our method. In general, we took three steps to estimate the building-level population distribution. Firstly, the LiDAR data and aerial photographs were used to extract individual building objects. Then, POI information, human activity information, and three-dimensional morphological information were extracted from POI data, NTL data, and LiDAR data, respectively. Finally, we built a random forest model to estimate the population for each individual building and compared the estimations with the building-level census data of 2016 which were extracted from the population database released by the Shanghai Public Security Bureau.
buildings to the total of reference buildings, and MO indicates that the total area of overlapping buildings partly accounts for the total area of the reference buildings.

3.2. Feature Extraction. Based on the Luojia1-01 NTL data, POI data, and LiDAR data, 33 features were preliminarily selected, including six NTL intensity features, two geometric features, and 25 POI features.

3.2.1. NTL Intensity Features. Chen et al. [62] established five statistical indicators of NTL intensity to reflect the socioeconomic situation, including the minimum intensity, maximum intensity, total NTL intensity, averaged intensity, and standard deviation of intensity. All the five NTL intensity features which reflect light intensity characteristics are proven to be good indicators for human activities. Previous studies [63, 64] have also demonstrated the ability of these features to estimate population. Besides, the slope of the NTL intensity surface which reveals the spatial variation of the human activity intensity gradient from the center to the periphery was also calculated. The slope of NTL intensity was computed as follows:

$$\text{Slope}_i = \frac{\text{Rise}_i}{\text{Run}_i} \times 100\%,$$

where Rise$_i$ denotes the difference between the NTL intensity value of the $i$th pixel and the largest NTL intensity value among the eight adjacent pixels of the $i$th pixel. Run$_i$ denotes the horizontal distance from the center of the $i$th pixel to the center of the pixel (among eight adjacent pixels) which has the largest NTL intensity value. The definitions of the above six NTL intensity features from Luojia1-01 data are detailed in Table 2.

3.2.2. Geometric Features of Buildings. The area and volume of buildings have a strong correlation with the number of people living in the buildings [49, 65]. The geometric attributes depict the geographical position and size of building boundary, and vertical measures of the buildings [33]. Here, we calculated the area and volume based on building data and LiDAR-derived nDSM data.

3.2.3. POI Features. Different categories of POI represent different human activities; therefore, they have different levels of correlation with population density [4, 66]. To facilitate the construction of POI features with buildings as statistical units, we convert POI records into kernel density data. The Kernel Density Estimation (KDE) was applied to generate five kernel density layers based on five reclassified POI categories [67]. Then, the mean, standard deviation, maximum, minimum, and sum of the five layers of kernel density were calculated for each individual building. As listed in Table 2, a total of 25 POI features were extracted. The letters a, b, c, d, and e in Table 2 represent the kernel density for the five reclassified POI categories listed in Table 1.

3.3. Population Estimation

3.3.1. Random Forest Algorithm. Random forest (RF) is a bagging ensemble learning algorithm introduced by Breiman [68]. RF models generate an internal unbiased estimate of the generalization error by means of a forest building process,
and each tree is constructed from a random subset of the training dataset. The prediction accuracy of RF models is evaluated by the Out-Of-Bag (OOB) accuracy, which is the prediction accuracy of the left-out sample average of all trees [69]. To assess the explanatory power of each feature, feature importance is introduced. Generally, the Gini index or the OOB error rate is used as evaluation indices for feature importance. The Gini index is defined as the total reduction in node impurity averaged over all trees [64] and can be used as a general indicator of feature correlation. The higher the value of the Gini index, the more important the feature. The multifeature nonlinear regression problem can be well solved by the RF regression model. Therefore, the RF regression model is widely used in the study of population estimation [24, 37, 40].

### 3.3.2. Population Estimation Using RF Model.

The entire dataset released by the Shanghai Public Security Bureau was randomly divided into two datasets: 40% as the training set and 60% as the testing set. The training set was used for RF model development and for tuning model parameter validation. The optimized model was then applied to the testing set for accuracy assessment.

We used the backward elimination method [70] to select features that provide the best predictive ability of the RF model. Firstly, we trained the RF model with all the features and removed the feature with the lowest Gini importance. Secondly, if the OOB error of the model increased, the feature is added back to the model. The process was repeated until no further improvements were observed. Finally, the remaining features were used to train the RF model. When training the model, several parameters need to be determined by the grid search method [71]. Table 3 shows the parameters that we optimized in the RF model. In order to investigate the robustness of the RF model, we also used different randomly chosen training samples to repeat the experiment 20 times. A mean $R^2$ of 0.71 was achieved with a low standard deviation of 0.014, which suggests the effectiveness and robustness of the random forest model. Finally, the spatial distribution of building population in the study area was generated based on the RF regression model.

### 3.4. Accuracy Assessment.

The reference population dataset provided by the Shanghai Public Security Bureau records the permanent population of each building. Therefore, we used this dataset as the reference data to verify the accuracy of the population estimation results of the RF model. To evaluate the accuracy of the RF model, the correlation between the estimated population and the reference population was calculated by using a linear regression method. We first calculated the coefficient of determination ($R^2$) and root mean squared error (RMSE) between the reference data and the estimated data. Then, we calculated the residuals between the reference data and the estimated data to further examine the errors in the population estimation results of the RF model.

### Table 2: Independent features derived from multisource data.

| Feature name | Description | Data source |
|--------------|-------------|-------------|
| Mean_NTL    | The mean NTL radiation value of all grids in a single building. |        |
| Sum_NTL     | The total NTL radiation value of all grids in a single building. |        |
| STD_NTL     | The standard deviation of NTL radiation value of all grids in a single building. |        |
| Mean_NTL_slope | The mean slope of NTL radiation value of all grids in a single building. |        |
| Min_NTL     | The minimum radiation value of NTL of all grids in a single building. |        |
| Max_NTL     | The maximum radiation value of NTL of all grids in a single building. |        |
| Area_building | The area of an individual building footprint. | LiDAR |
| Volume_building | The volume of an individual building. |        |
| Min_POI_a/b/c/d/e | The minimum kernel density of all grids in a single building. | POI    |
| Max_POI_a/b/c/d/e | The maximum kernel density of all grids in a single building. |        |
| Mean_POI_a/b/c/d/e | The mean kernel density of all grids in a single building. |        |
| Sum_POI_a/b/c/d/e | The total kernel density of all grids in a single building. |        |
| STD_POI_a/b/c/d/e | The standard deviation of kernel density of all grids in a single building. |        |

### Table 3: The description and value of random forest model parameters.

| Parameter name | Description | Optimal value |
|----------------|-------------|---------------|
| n_estimators   | Number of trees | 120           |
| Max_features   | Maximum number of features that can be selected at each node | 20            |
| Max_depth      | Maximum depth of the tree | 16            |
| Min_samples_split | Minimum number of samples required to split an internal node | 20            |
| Min_samples_leaf | Minimum number of samples required to be at a leaf node | 1             |
3.5. Feature Contribution Analysis. The Gini index was calculated to explain the overall contribution of different features to the estimated population results. As the feature importance is often insufficient to fully explain the relationship between input features and predicted values [72], Palczewska et al. [73] proposed a method for calculating feature contributions to explain the random forest model. Feature contributions are equal to the sum of the change probability of being in class over all nodes [73], along the path from the root node to the terminal node in a decision tree of the random forest model. The feature contributions are calculated separately for each instance to provide detailed information of the relationships between features and the predicted values.

4. Results

4.1. Results of Building Object Extraction. The result of regularized building boundaries is shown in Figure 6. A total of 6,194 buildings have been extracted. These buildings have an average size of 602 m², with the smallest area of 100 m² and the largest area of 10,048 m². By overlying the nDSM data, we found that the lowest building height is 3.65 m, the highest building height is 460 m, and the average height is 6.7 m. Compared with the reference building data, the overall accuracy of DR and MO of building objects extracted was 92.58% and 94.28%, respectively. The results indicated a high accuracy of the building detection method in our study.

4.2. Population Estimation Results. Figure 7 shows the selected features ordered by the Gini index. 20 features were identified as core features by the backward elimination method. These 20 core features include 13 POI features, two geometric features, and five NTL intensity features. Volume_building, which has an importance weight of 20.7%, was found to be the most important feature. Among the three types of feature sets, the POI feature sets were found to be more important than the other two feature sets in terms of both the number and rankings of features, with a total importance weight of 53.3%. The residential POI feature which has an importance weight of 17.4% was identified as the most important POI feature. It also shows the importance of building-related features in population estimation as the study area contains a large number of residential buildings.

The 20 selected features were then applied into the RF regression model to estimate the population for the entire study area. The spatial distribution of the estimated population is shown in Figure 8(a). The building with the largest population has 729 people. 38.11% of the buildings have a population fewer than 100 people. These buildings are mainly located in commercial and office areas (Figure 8(b)) which consist of nonresidential buildings and some small residential areas (Figure 8(c)) where the buildings are small in size. Only 33 buildings were found to have a population larger than 500. The 33 buildings were mainly concentrated southeast of the study area, suggesting a high-density residential area. The southern part of the study area is occupied by many old residential areas, most of which have a population of fewer than 300 people for each building. The
Figure 7: The Gini index of selected features.

Figure 8: The spatial distribution of population estimation results (a). (b, c) Two examples are shown.
commercial buildings with a small population (most of them have a population of fewer than 100) are mainly located in the northern part of the study area.

4.3. Accuracy Analysis. Figure 9 illustrates the relationship between the reference population and the estimated population. The $R^2$ between estimated populations and actual populations is 0.65, and it is significant at the 0.01 level, which suggests that the estimated population and the reference population were well correlated. Figure 9 also suggests that our method tends to slightly underestimate. This can also be found from the spatial distribution of the estimated population and the actual population. By comparing the reference population (Figure 10(a)) and estimated population (Figure 8), the range of the actual population (from 0 to 1,170) was larger than that of the estimated population (from 0 to 729), indicating that the results estimated by the RF model was underestimated. Two outliers (see Figures 9(a) and 9(b)) which the RF model failed to predict were selected to conduct an in-depth analysis. Figure 9(a) is the Shanghai Shimao Plaza (depicted in a red box) which is a commercial building with few people living in there. However, it has a very large building volume and thus turned out to be an overestimation. Figure 9(b) is a walk-up building (depicted in a red box) which is located in an old and densely populated residential district. Besides, the building is also located in the catchment area for a school (depicted in a blue box). However, it turned out to be an underestimation as it has a small building volume.

The residuals of each building between the estimated population and the actual population were calculated to further assess the accuracy of the estimated results. Figure 10(b) shows the residuals between the estimated population and the actual population. There are 3,101 buildings, which account for 83.45% of the testing set, and these have a residual range from -50 to 50 (depicted in yellow in Figure 10(b)). This indicates that the overall difference between the estimated population and the actual population is very small.

A simple accuracy comparison analysis was also conducted. In the existing literature related to building-level population estimation, we found that the study area in the research of Yao et al. [50] is very similar to ours. In Yao et al.’s study, they selected five central urban areas (i.e., Yuexiu, Liwan, Tianhe, Haizhu, and Baiyun) in Guangzhou, China as their study area. These areas are all highly developed urban centers with relatively stable development, which are very similar to our study areas. Both two study areas have similar building types, including skyscrapers and different types of residential buildings (i.e., townhouses, walk-up buildings, and elevator buildings). Unfortunately, as the data used in Yao et al.’s study is inaccessible in our study area, we were unable to repeat the experiment described in their research for direct comparison. Therefore, only the accuracy result was selected for comparison. It also should be noted that, due to the lack of building-level validation data, Yao et al.’s study conducted the accuracy validation of population estimation results at the community level. Hence, we performed the accuracy comparison analysis at the community level (a total of 131 communities in our study area). The spatial distribution of population estimation results at the community level is shown in Figure 11.

As shown in Figure 12, the $R^2$ of the proposed method at the community level is 0.7864. This is slightly better than or comparable to the result reported in Yao et al.’s study (i.e., an $R^2$ of 0.7422 at the community level).

5. Discussion

5.1. Feature Contribution of the RF Model. Figure 13 shows the relationship between the feature contributions and three
Figure 10: The comparison between the actual and estimated population in the study area. (a) Actual population. (b) Residual between the actual and estimated population.

Figure 11: The spatial distribution of population estimation results at the community level.
of the most important features (i.e., \textit{Volume\_building}, \textit{Sum\_POI\_a}, and \textit{Sum\_POI\_d}) and a nighttime light intensity feature of \textit{Sum\_NTL}. The scatter plot shows the features which have a strong relationship with population, such as \textit{Volume\_building} and \textit{Sum\_POI\_a} (residential). As shown in Figure 13(a), with the continuous increase of building volume, the feature contribution value gradually rises. When the building volume is greater than 25,000 m$^3$, the feature contributions stop increasing. This is because when the building volume is larger than 25,000 m$^3$, such buildings belong to nonresidential buildings, such as stadiums and office buildings. As shown in Figure 13(b), with the gradual increase of the feature \textit{Sum\_POI\_a} (residential) value, the feature contributions also rise gradually, and most of the feature contributions are distributed between -50 and 50. In Figure 13(c), the feature contribution increases overall with the increase of the institutional POI kernel density. This suggests that more people live in areas with high \textit{Sum\_POI\_d} (institutional) values; these areas are always characterized by dense business activities and well-established public infrastructure. From the absolute value of the feature contribution, the absolute value of the contribution of the building volume can reach more than 400, far higher than other features, which is consistent with the order of feature importance (Figure 7).

5.2. Population Estimation Ability of NTL Data. As shown in Figure 13(d), with an increase in the sum of NTL radiance intensity, the feature contributions show a decreasing trend. This is not consistent with some existing studies [18, 40, 74], as they have proven that there is a significant positive...
correlation between the sum of NTL radiance intensity and population at large scales. To further analyze the feature contributions of the sum of NTL radiance intensity, a spatial visualization was conducted on the feature contributions of the sum of NTL radiance intensity (Figure 14(a)), and images at night were also combined for observation. We found that the feature contributions of the sum of NTL radiance intensity were mostly negative and small in the office areas (Figure 14(b)) and commercial areas (Figure 14(c)), while in the residential areas (Figure 13(d) and Figure 14(e)), the feature contributions were mostly positive. This is because the office and commercial areas have large NTL radiance intensity at night, while residential areas have a small NTL radiance intensity at night. Most of the top important features show positive relations with population estimation. For example, Figure 13(a) has shown that building volume has a very strong and positive relationship with the population, which means that the larger the building volume, the larger the population. Usually, office and commercial buildings have very large building volumes, which means a large population will be predicted by the RF model if only the building volume was considered. The feature contribution of our RF model found that the large NTL radiance intensity has a negative relationship for population estimation. As the office and commercial buildings have very large NTL radiance intensity, therefore, the large NTL radiance intensity has weakened the overestimation of population in office and commercial areas which indicates a positive feature contribution. Previous studies have demonstrated the positive relationship between the NTL radiance intensity and population at large scales, such as at the country level and the city level, but without studying the role of NTL radiance intensity in population estimation at a fine scale, such as at the building level. Our study shows that for building-level population estimation, a positive correlation was found between NTL radiance intensity and population in residential areas, while a negative correlation was found in commercial areas.

5.3. Merit and Limitations. To enable richer insights regarding questions related to building-level population estimation, novel solutions should be able to integrate data coming from many different data sources. This study provided a new perspective and made a pioneering effort for building-level population estimation by using both the three-dimensional morphological information and the human activity information. We used LiDAR data to extract building footprints and derived areal and volumetric properties which are closely related to population counts. Given that there is a nonstationary relationship between LiDAR-derived morphological indicators and populations, we integrated data from POI and NTL data to differentiate the human activities. By integrating LiDAR, nighttime light, and POI data, our method can reflect building-level populations in a more precise way than what is possible for each individual data source.

It should be noted that the inconsistency of the data sets might have influenced the accuracy of population estimation in this study. However, the Huangpu District has had a stable development in the past 20 years, and the three-dimensional morphology (especially buildings) and cityscape changed very little. So, the inconsistent data acquisition dates may alert little impact on our results as the changes of the major
data for estimating population are relatively insignificant. Furthermore, due to the difficulty in obtaining fine LiDAR point clouds for the entire Shanghai City, our research area is limited to the Huangpu District, Shanghai. If the LiDAR data for entire city are available, the method can be applied to large-scale building-level population investigations comprising both urban and suburban environments. Due to the dynamic nature of the building-level population, continued evaluation and validation of the proposed method for different geographic regions and multiple temporal scales (e.g., daytime and nighttime) should be conducted to better establish the effectiveness of the results. Finally, as the Luoja 1-01 satellite is no longer providing new data since June 2019, the Luoja1-01 NTL data can be replaced by other high-resolution NTL data such as JLI-3B and EROS-B when applying our method to other areas.

6. Conclusions

Building-level population estimation in urban areas is particularly difficult due to the unbalanced distribution of human activities and the lack of 3D building information. In this study, we proposed a framework for population estimation at the building level by integrating both the three-dimensional morphological information derived from LiDAR data and the human activity information extracted from POI and NTL data. Through a rigorous feature selection process, 20 of the initial 33 features were finally selected for the well-trained RF model. The trained RF model was applied to map the spatial distribution of the population at the building level. Subsequently, the accuracy of the model estimation results was evaluated and analyzed, and the model was explained in detail from two aspects of feature importance and feature contributions.

The $R^2$ between the reference and the estimated population was 0.65, indicating a satisfying overall accuracy. The analysis of the residuals shows that the RF model tends to overestimate for the buildings with a low population and underestimate for the buildings with a high population. Building volume was the most important feature to estimate population with an importance of 20.7%. The NTL radiance intensity was found to be less important than two POI-related features (i.e., Sum_POI_a (residential) and Sum_POI_d (institutional)). From the perspective of feature contributions, we found that when using the random forest model to estimate the population for each individual building, NTL radiance intensity has a positive effect on population estimation in residential areas, while it shows a significant negative effect in office and commercial areas.

Due to the difficulty of data collection, this study only focuses on the building-level population estimation for a single time phase. If long time series NTL, LiDAR, and social-sensing data from other cities are available, our method can be further validated for long-term fine-scale population estimations. It should also be noted that the Luoja1-01 NTL data currently only cover some parts of the world, so population estimation using other high-resolution NTL data (such as JLI-3B and EROS-B) as alternative data sources requires further research. Finally, how to combine three-dimensional morphological data (such as LiDAR point clouds) with other geospatial big data to improve the fine-scale population estimation is also a research area that needs to be further explored.

Data Availability

The Luoja1-01 NTL data and codes that support the findings of this study are available at the figshare repository (doi:10.6084/m9.figshare.13465742). The original POI data, LiDAR data, population data, and aerial photographs cannot be shared publicly due to restrictions.

Conflicts of Interest

No potential conflict of interest was reported by the authors.

Authors’ Contributions

Hongxing Chen and Bin Wu contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant numbers 41871331, 41801343, and 42001357).

References

[1] E. M. Weber, V. Y. Seaman, R. N. Stewart et al., “Census-independent population mapping in northern Nigeria,” Remote Sensing of Environment, vol. 204, pp. 786–798, 2018.
[2] Y. Zhou, M. Ma, K. Shi, and Z. Peng, “Estimating and interpreting fine-scale gridded population using random forest regression and multisource data,” ISPRS International Journal of Geo-Information, vol. 9, no. 6, p. 369, 2020.
[3] T. Ahola, K. Virrantaus, J. M. Krisp, and G. J. Hunter, “A spatio-temporal population model to support risk assessment and damage analysis for decision-making,” International Journal of Geographical Information Science, vol. 21, no. 8, pp. 953–953, 2007.
[4] M. Bakillah, S. Liang, A. Mobasher, J. Jokar Arsanjani, and A. Zipf, “Fine-resolution population mapping using OpenStreetMap points-of-interest,” International Journal of Geographical Information Science, vol. 28, no. 9, pp. 1940–1963, 2014.
[5] T. C. Brown, V. Mahat, and J. A. Ramirez, “Adaptation to future water shortages in the United States caused by population growth and climate change,” Earth’s Future, vol. 7, no. 3, pp. 219–234, 2019.
[6] J. Huang, D. Qin, T. Jiang et al., “Effect of fertility policy changes on the population structure and economy of China: from the perspective of the shared socioeconomic pathways,” Earth’s Future, vol. 7, no. 3, pp. 250–265, 2019.
[7] P. Deville, C. Linard, S. Martin et al., “Dynamic population mapping using mobile phone data,” Proceedings of the National Academy of Sciences of the United States of America, vol. 111, no. 45, pp. 15888–15893, 2014.
[8] D. Martin, S. Cockings, and S. Leung, “Developing a flexible framework for spatiotemporal population modeling,” Annals
of the Association of American Geographers, vol. 105, no. 4, pp. 754–772, 2015.

[9] E. K. Noji, “Estimating population size in emergencies,” Bulletin of the World Health Organization, vol. 83, no. 3, pp. 164–164, 2005.

[10] Y. Chen, F. Guo, J. Wang, W. Cai, C. Wang, and K. Wang, “Provincial and gridded population projection for China under shared socioeconomic pathways from 2010 to 2100,” Scientific Data, vol. 7, no. 1, p. 83, 2020.

[11] M. E. Hauer, “Population projections for U.S. counties by age, sex, and race controlled to shared socioeconomic pathway,” Scientific Data, vol. 6, no. 1, p. 190005, 2019.

[12] Y. Zhao, G. A. Ovando-Montejo, A. E. Frazier, A. J. Mathews, K. C. Flynn, and E. A. Ellis, “Estimating work and home population using LiDAR-derived building volumes,” International Journal of Remote Sensing, vol. 38, no. 4, pp. 1180–1196, 2017.

[13] S. Froliking, J. Qiu, S. Boles et al., “Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China,” Global Biogeochemical Cycles, vol. 16, no. 4, pp. 38.1–38.10, 2002.

[14] G. Li and Q. Weng, “Measuring the quality of life in city of Indianapolis by integration of remote sensing and census data,” International Journal of Remote Sensing, vol. 28, no. 2, pp. 249–267, 2007.

[15] W. R. Tobler, “Satellite confirmation of settlement size coefficients,” Area, vol. 1, no. 3, pp. 30–34, 1969.

[16] J. Mennis, “Generating surface models of population using dasymetric mapping,” The Professional Geographer, vol. 55, no. 1, pp. 31–42, 2003.

[17] M. Monmonier and G. A. Schnell, “Land use and land cover data and the mapping of population density,” International Yearbook of Cartography, vol. 24, pp. 115–121, 1984.

[18] P. Sutton, D. Roberts, C. Elvidge, and K. Baugh, “Census from heaven: an estimate of the global human population using night-time satellite imagery,” International Journal of Remote Sensing, vol. 22, no. 16, pp. 3061–3076, 2001.

[19] B. Bhaduri, E. Bright, P. Coleman, and M. L. Urban, “Land-Scan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics,” Geoforum, vol. 69, no. 1-2, pp. 103–117, 2007.

[20] W. Tobler, U. Deichmann, J. Gottsegen, and K. Maloy, “World population in a grid of spherical quadrilaterals,” International Journal of Population Geography, vol. 3, no. 3, pp. 203–225, 1997.

[21] M. Melchiorri, A. J. Florczyk, S. Freire, M. Schiavina, M. Pesaresi, and T. Kemper, “Unveiling 25 years of planetary urbanization with remote sensing: perspectives from the global human settlement layer,” Remote Sensing, vol. 10, no. 5, p. 768, 2018.

[22] C. T. Lloyd, A. Sorichetta, and A. J. Tatem, “High resolution global gridded data for use in population studies,” Scientific Data, vol. 4, no. 1, p. 170001, 2017.

[23] Yanhua Xie, A. Weng, and Qihao Weng, “Population estimation of urban residential communities using remotely sensed morphologic data,” IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 5, pp. 1111–1115, 2015.

[24] T. Ye, N. Zhao, X. Yang et al., “Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model,” Science of the Total Environment, vol. 658, pp. 936–946, 2019.

[25] P. Dong, S. Ramesh, and A. Nepali, “Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data,” International Journal of Remote Sensing, vol. 31, no. 21, pp. 5571–5586, 2010.

[26] Z. Lu, J. Im, and L. Quackenbush, “A volumetric approach to population estimation using LiDAR remote sensing,” Photogrammetric Engineering & Remote Sensing, vol. 77, no. 11, pp. 1145–1156, 2011.

[27] Z. Lu, J. Im, L. Quackenbush, and K. Halligan, “Population estimation based on multi-sensor data fusion,” International Journal of Remote Sensing, vol. 31, no. 21, pp. 5587–5604, 2010.

[28] F. Qiu, H. Sridharan, and Y. Chun, “Spatial autoregressive model for population estimation at the census block level using LiDAR-derived building volume information,” Cartography and Geographic Information Science, vol. 37, no. 3, pp. 239–257, 2010.

[29] L. Tomáš, L. Fonseca, C. Almeida, F. Leonardi, and M. Pereira, “Urban population estimation based on residential buildings volume using IKONOS-2 images and lidar data,” International Journal of Remote Sensing, vol. 37, no. sup1, pp. 1–28, 2016.

[30] S. Wang, Y. Tian, Y. Zhou, W. Liu, and C. Lin, “Fine-scale population estimation by 3D reconstruction of urban residential buildings,” Sensors, vol. 16, no. 10, p. 1755, 2016.

[31] S. Wu, X. Qiu, and L. Wang, “Population estimation methods in GIS and remote sensing: a review,” GIScience & Remote Sensing, vol. 42, no. 1, pp. 80–96, 2005.

[32] S. Wu, L. Wang, and X. Qiu, “Incorporating GIS building data and census housing statistics for sub-block-level population estimation,” The Professional Geographer, vol. 60, no. 1, pp. 121–135, 2008.

[33] B. Wu, B. Yu, Q. Wu et al., “An extended minimum spanning tree method for characterizing local urban patterns,” International Journal of Geographical Information Science, vol. 32, no. 3, pp. 450–475, 2018.

[34] B. Wu, B. Yu, Q. Wu, Y. Huang, Z. Chen, and J. Wu, “Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests,” International Journal of Applied Earth Observation and Geoinformation, vol. 52, pp. 82–94, 2016.

[35] B. Wu, B. Yu, W. Yue et al., “A voxel-based method for automated identification and morphological parameters estimation of individual street trees from mobile laser scanning data,” Remote Sensing, vol. 5, no. 2, pp. 584–611, 2013.

[36] C. D. Elvidge, K. Baugh, M. Zhizhin, F. C. Hsu, and T. Ghosh, “VIIRS night-time lights,” International Journal of Remote Sensing, vol. 38, no. 21, pp. 5860–5879, 2017.

[37] K. Li, Y. Chen, and Y. Li, “The random forest-based method of fine-resolution population spatialization by using the international space station night time photography and social sensing data,” Remote Sensing, vol. 10, no. 10, p. 1650, 2018.

[38] X. Li, H. Xu, X. Chen, and C. Li, “Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of China,” Remote Sensing, vol. 5, no. 6, pp. 3057–3081, 2013.

[39] K. Shi, B. Yu, Y. Huang et al., “Evaluating the ability of NPP-VIIRS night-time light data to estimate the gross domestic product and the electric power consumption of China at multiple scales: a comparison with DMSP-OLS data,” Remote Sensing, vol. 6, no. 2, pp. 1705–1724, 2014.

[40] B. Yu, T. Lian, Y. Huang et al., “Integration of nighttime light remote sensing images and taxi GPS tracking data for
population surface enhancement,” *International Journal of Geographical Information Science*, vol. 33, no. 4, pp. 687–706, 2019.

[41] M. Amaral, A. M. V. Monteiro, G. Camara, and J. A. Quintanilha, “DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon,” *International Journal of Remote Sensing*, vol. 27, no. 5, pp. 855–870, 2006.

[42] H. Bagan and Y. Yamagata, “Analysis of urban growth and estimating population density using satellite images of night-time lights and land-use and population data,” *GisScience & Remote Sensing*, vol. 52, no. 6, pp. 765–780, 2015.

[43] C. D. Elvidge, K. E. Baugh, J. B. Dietz, T. Bland, P. C. Sutton, and H. W. Kroehl, “Radiance calibration of DMSP-OLS low-light imaging data of human settlements,” *Remote Sensing of Environment*, H. W. Kroehl, Ed., vol. 68, no. 1, pp. 77–88, 1999.

[44] M. L. Imhoff, W. T. Lawrence, D. C. Stutzer, and C. D. Elvidge, “A technique for using composite DMSP/OLS ‘City Lights’ satellite data to map urban area,” *Remote Sensing of Environment*, vol. 61, no. 3, pp. 361–370, 1997.

[45] F. R. Stevens, A. E. Gaughan, C. Linard, and A. J. Tatem, “Dis-aggregating census data for population mapping using random forests with remotely-sensed and ancillary data,” *PLoS One*, vol. 10, no. 2, p. e0107042, 2015.

[46] P. Sutton, “Modeling population density with night-time satellite imagery and GIS,” *Computers, Environment & Urban Systems*, vol. 21, no. 3–4, pp. 227–244, 1997.

[47] P. C. Sutton, “A scale-adjusted measure of ‘Urban sprawl’ using nighttime satellite imagery,” *Remote Sensing of Environment*, vol. 86, no. 3, pp. 353–369, 2003.

[48] C. Zeng, Y. Zhou, S. Wang, F. Yan, and Q. Zhao, “Population spatialization in China based on night-time imagery and land use data,” *International Journal of Remote Sensing*, vol. 32, no. 24, pp. 9599–9620, 2011.

[49] S. Ural, E. Hussain, and J. Shan, “Building population mapping with aerial imagery and GIS data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 6, pp. 841–852, 2011.

[50] Y. Yao, X. Liu, X. Li et al., “Mapping fine-scale population distributions at the building level by integrating multisource geospatial big data,” *International Journal of Geographical Information Systems*, vol. 31, no. 6, pp. 1–25, 2017.

[51] C. Li, L. Zou, Y. Wu, and H. Xu, “Potentiality of using Luojia1-01 night-time light imagery to estimate urban community housing price—a case study in Wuhan, China,” *Sensors*, vol. 19, no. 14, p. 3169, 2019.

[52] X. Li, X. Li, D. Li, X. He, and M. Jendryke, “A preliminary investigation of Luojia-1 night-time light imagery,” *Remote Sensing Letters*, vol. 10, no. 6, pp. 526–535, 2019.

[53] C. Wang, Z. Chen, C. Yang et al., “Analyzing parcel-level relationships between Luojia 1-01 nighttime light intensity and artificial surface features across Shanghai, China: a comparison with NPP-VIIRS data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 85, p. 101989, 2020.

[54] G. Zhang, L. Li, Y. H. Jiang, X. Shen, and D. R. Li, “On-orbit relative radiometric calibration of the night-time sensor of the Luojia1-01 satellite,” *Sensors*, vol. 18, no. 12, p. 4225, 2018.

[55] Y. Huang, B. Yu, J. Zhou et al., “Toward automatic estimation of urban green volume using airborne LiDAR data and high resolution remote sensing images,” *Frontiers of Earth Science*, vol. 7, no. 1, pp. 43–54, 2013.

[56] S. Yu, B. Yu, W. Song et al., “View-based greenery: a three-dimensional assessment of city buildings’ green visibility using floor green view index,” *Landscape and Urban Planning*, vol. 152, pp. 13–26, 2016.

[57] E. Louw and F. Bruinsma, “From mixed to multiple land use,” *Journal of Housing and the Built Environment*, vol. 21, no. 1, pp. 1–13, 2006.

[58] K. Zhang, C. Shu-Ching, D. Whitman, M.-L. Shyu, J. Yan, and C. Zhang, “A progressive morphological filter for removing nonground measurements from airborne LIDAR data,” *IEEE Transactions on Geoscience & Remote Sensing*, vol. 41, no. 4, pp. 872–882, 2003.

[59] B. Yu, H. Liu, J. Wu, Y. Hu, and L. Zhang, “Automated derivation of urban building density information using airborne LiDAR data and object-based method,” *Landscape and Urban Planning*, vol. 98, no. 3–4, pp. 210–219, 2010.

[60] A. Gribov, “Searching for a compressed polyline with a minimal number of vertices,” in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, 2017.

[61] Wenbo Song and T. L. Haithcoat, “Development of comprehensive accuracy assessment indexes for building footprint extraction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 402–404, 2005.

[62] Z. Chen, B. Yu, W. Song et al., “A new approach for detecting urban centers and their spatial structure with nighttime light remote sensing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 11, pp. 6305–6319, 2017.

[63] C. P. Lo, “Modeling the population of China using DMSP operational linescan system nighttime data,” *Photogrammetric Engineering & Remote Sensing*, vol. 67, no. 9, pp. 1037–1047, 2001.

[64] X. Zhao, B. Yu, Y. Liu et al., “Estimation of poverty using random forest regression with multi-source data: a case study in Bangladesh,” *Remote Sensing*, vol. 11, no. 4, p. 375, 2019.

[65] K. K. Lwin and Y. Murayama, “Estimation of building population from LiDAR derived digital volume model,” in *Spatial Analysis and Modeling in Geographical Transformation Process*, Y. Murayama and R. B. Thapa, Eds., pp. 87–98, Springer, Dordrecht, 2011.

[66] J. Cai, B. Huang, and Y. Song, “Using multi-source geospatial big data to identify the structure of polycentric cities,” *Remote Sensing of Environment*, vol. 202, pp. 210–221, 2017.

[67] Y. Zhang, Q. Li, H. Huang, W. Wu, X. Du, and H. Wang, “The combined use of remote sensing and social sensing data in fine-grained urban land use mapping: a case study in Beijing, China,” *Remote Sensing*, vol. 9, no. 9, p. 865, 2017.

[68] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[69] C. Strobl, J. Malley, and G. Tutz, “An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests,” *Psychological Methods*, vol. 14, no. 4, pp. 323–348, 2009.

[70] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[71] P. M. Lerman, “Fitting segmented regression models by grid search,” *Applied Statistics*, vol. 29, no. 1, pp. 77–84, 1980.

[72] C. Strobl, A. L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, “Conditional variable importance for random forests,” *BMC Bioinformatics*, vol. 9, no. 1, pp. 307–307, 2008.
[73] A. Palczewska, J. Palczewski, R. M. Robinson, and D. Neagu, “Interpreting random forest models using a feature contribution method,” in 2013 IEEE 14th International Conference on Information Reuse & Integration (IRI), pp. 112–119, San Francisco, CA, USA, 2013.

[74] B. R. Tripathy, V. Tiwari, V. Pandey et al., “Estimation of urban population dynamics using DMSP-OLS night-time lights time series sensors data,” IEEE Sensors Journal, vol. 17, no. 4, pp. 1013–1020, 2017.