

# Estimation of Population Distribution Using Google Building Footprints and Administrative Data: A Case Study of Hue City

DO Van Duong<sup>1</sup>, NGUYEN Quang Minh<sup>2,\*</sup>

<sup>1</sup> Hanoi University of Natural Resource and Environment, Hanoi, Vietnam

<sup>2</sup> Hanoi University of Mining and Geology, Hanoi, Vietnam

\* Corresponding author: [nguyenquangminh@humg.edu.vn](mailto:nguyenquangminh@humg.edu.vn)

**Abstract:** *Reliable estimates of population distribution are essential for effective urban planning, infrastructure development, and disaster management. In this study, we propose a spatial analysis approach to estimate the intra-urban population distribution of Hue City by combining Google Building Footprint data with official administrative population statistics. The high-resolution building outlines provided by Google serve as a proxy for the spatial extent of residential structures. By integrating these building footprints with ward-level census data and applying dasymetric mapping techniques, we redistribute population figures based on the density and distribution of built-up areas. Additional spatial layers, such as land use zoning and road networks, are used to refine the estimates and exclude non-residential zones. The resulting population density map reveals fine-scale patterns of settlement intensity across the city, highlighting both core urban centers and emerging peripheral developments. This method demonstrates the value of open geospatial data for enhancing traditional population models, especially in rapidly urbanizing contexts with limited access to detailed local datasets.*

**Keywords:** *open-data, population distribution estimation*

## 1. Introduction

Accurate and fine-scale population distribution data are critical for evidence-based decision-making in urban planning, infrastructure development, and disaster risk management. Traditional population statistics, typically collected through censuses and household surveys, are often aggregated at administrative units such as wards or districts. While these data are valuable for demographic analysis, their coarse spatial resolution limits their usefulness in applications that require fine-scale intra-urban population patterns. Such applications include optimizing urban infrastructure such as public transportation networks, water supply and sewage systems, and communication networks as well as locating emergency shelters and assessing population exposure to natural hazards (Mennis, 2009; Stevens *et al.*, 2015). In rapidly urbanizing cities, where land-use dynamics and informal settlements evolve swiftly, the development of urban infrastructure and disaster preparedness are crucial, making the demand for high-resolution population maps increasingly urgent.

Recent advancements in geospatial technologies, coupled with the increasing accessibility of open-source data, have provided significant opportunities to overcome the challenges associated with limited high-resolution population distribution information. Dasymetric mapping, which redistributes population counts from administrative units to finer spatial units based on ancillary data, has emerged as an effective technique for refining population estimates (Eicher and Brewer, 2001). In particular, building footprint data have proven to be strong proxies for residential population distribution, as they directly represent the spatial extent of built-up areas where people live (Tiecke *et al.*, 2017; Boo *et al.*, 2022). The growing availability of high-quality, open-access building footprint datasets, such as those provided by Google and Microsoft, offers significant potential for improving population distribution models, especially in developing regions where detailed local datasets are scarce (Huang *et al.*, 2021). These datasets provide consistent and standardized coverage across large areas, overcoming the limitations of traditional cadastral or survey data, which are often outdated, incomplete, or restricted by administrative policies (Fetai *et al.*, 2022). Furthermore, building footprints act as direct proxies for human habitation, enabling a more accurate redistribution of population counts compared to ancillary data such as land-use classifications or night-time lights, which frequently misrepresent population in commercial or non-residential zones (Ural, Hussain and Shan, 2011). Their high spatial resolution also allows the detection of fine-scale settlement patterns, including informal or peri-urban developments often missing from official maps (Ehrlich *et al.*, 2018; Corbane *et al.*, 2019). Importantly, in many developing regions, there is a lack of technical capacity, computational resources, and expertise to extract building footprints directly from high-resolution satellite or aerial imagery. Openly available datasets thus eliminate this barrier, providing ready-to-use information

that significantly enhances population modeling efforts (Leyk *et al.*, 2019).

In Vietnam, studies on high-resolution population mapping remain limited, despite the country's rapid urban expansion and increasing vulnerability to climate change and natural disasters. Hue City, a medium-sized urban center with complex land-use patterns and ongoing peri-urban development, presents an ideal case study for testing innovative spatial approaches to population estimation. By leveraging Google Building Footprint data and integrating them with official ward-level census statistics, this study applies dasymetric mapping techniques to produce a fine-scale population density map for Hue City. Additional spatial layers, such as land use zoning and road networks, are incorporated to exclude non-residential areas and improve the accuracy of the redistribution process.

This research contributes to the growing body of literature demonstrating the value of open geospatial data in enhancing traditional population models. Beyond providing a high-resolution view of population distribution in Hue, the proposed approach also offers a replicable framework for other rapidly urbanizing cities with limited access to detailed local datasets. The results aim to support urban planners, policymakers, and disaster management agencies in making more informed decisions based on reliable and spatially explicit population information.

## 2. Method

### 2.1. Data Sources

This study utilizes multiple spatial and statistical datasets to produce a fine-scale population density map for Hue City through dasymetric mapping techniques. The data includes Google building footprints data, administrative polygons with population information, and land-use zoning data.

#### 2.1.1. Google Building Footprints

The primary spatial dataset comprises high-resolution Google Open Buildings footprints, which provide vectorized building outlines automatically extracted from high-resolution satellite imagery using advanced machine learning models. The dataset is publicly available in CSV format and provides polygon geometries encoded in Well-Known Text (WKT), and can be accessed at <https://sites.research.google/gr/open-buildings/#open-buildings-download>. This dataset represents building outlines automatically extracted from high-resolution satellite imagery using deep learning models. These footprints serve as presentation for residential structures, as their detailed geometry accurately represents the spatial extent and distribution of built-up areas across the city. The dataset offers near-complete coverage of urban, suburban, rural and mountainous areas and includes attributes such as building areas which use square metre as unit value. The original building data was then converted to the ESRI shapefile format which can be used easily with open-source software such as Quantum GIS (QGIS). The Google Building footprints in shapefile format is presented in Fig. 1.

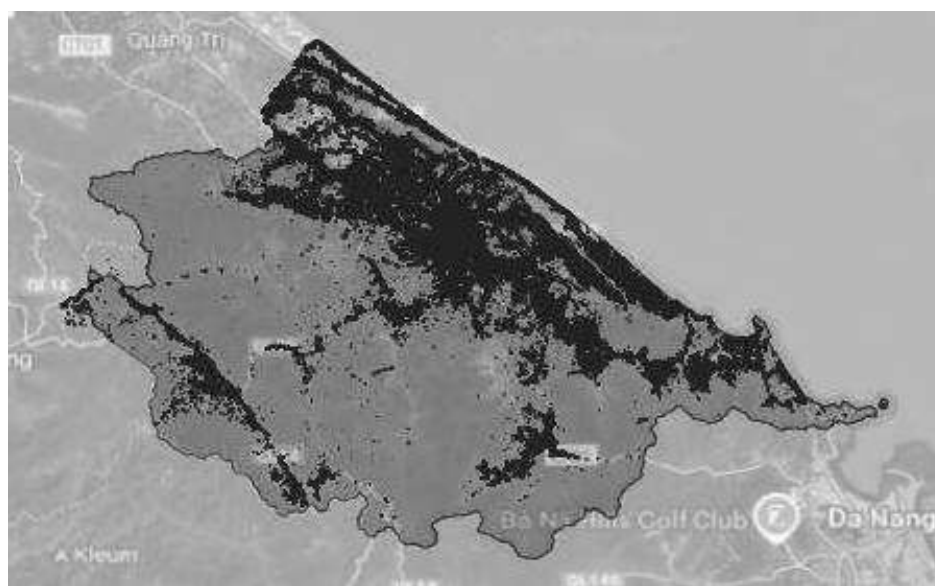


**Fig. 1.** The Google Building Footprints of City of Hue

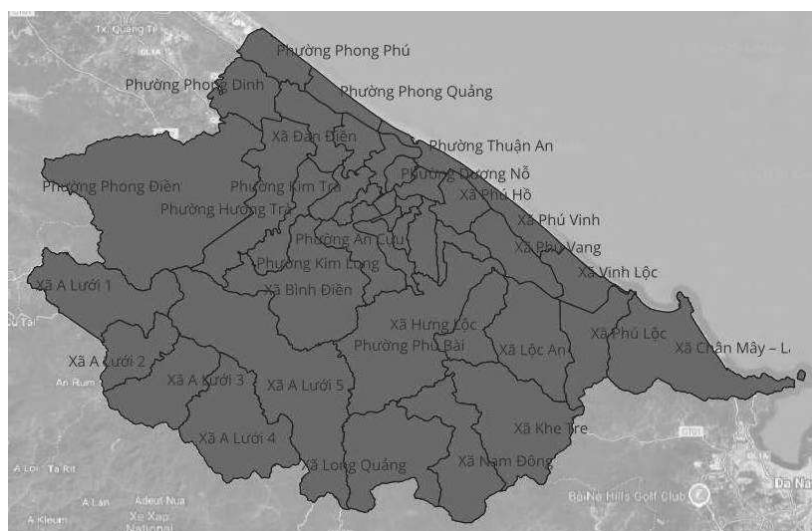


**Fig. 2.** Administrative Boundaries of communes and wards in Hue City

### 2.1.2. Administrative data

The administrative boundary data used in this study were updated according to the most recent official administrative restructuring issued by the National Assembly of Vietnam. The boundary and other information of Hue City are described in Resolution 1675/NQ-UBTVQH15 on the Arrangement of Commune-Level Administrative Units of Hue City in 2025 issued by the Standing Committee of the National Assembly of Vietnam. The administrative boundary can be found at URL address as <https://vnsdi.mae.gov.vn/bandohanhchinh/>. A provisional list with information on area and population of each communes and wards is issued by the Ministry of Home Affairs in 2025 in Official Document 2896/BNV-CQĐP. The map of communes and wards of Hue City is presented in Fig. 2.

### 2.1.3. Zoning data

The National Land-Use Data (Dữ liệu hiện trạng sử dụng đất quốc gia), provided by the General Department of Land Administration – Ministry of Natural Resources and Environment (MONRE), is the most authoritative dataset for land-use classification in Vietnam. At a scale of 1:25,000, it provides detailed, standardized land-use categories essential for distinguishing residential from non-residential areas in dasymetric population mapping (Fig. 3). The zoning codes and their using for population mapping are presented in Tab. 1.

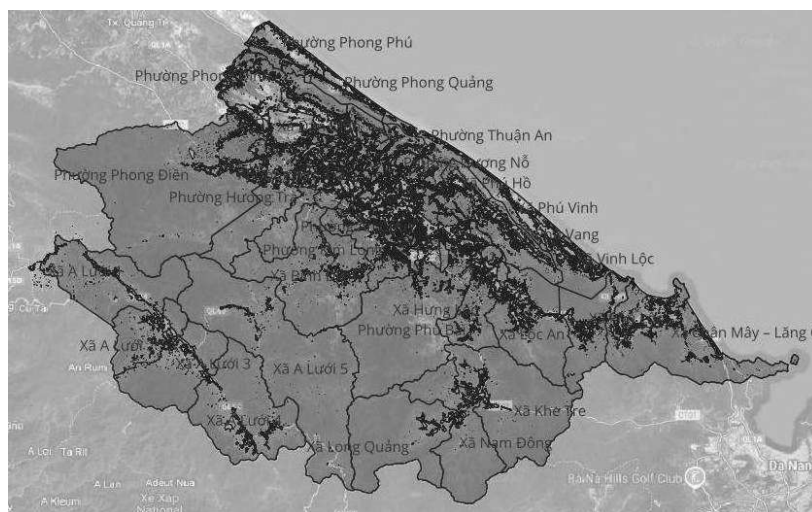


**Fig. 3.** Urban and Rural Residential Land Zones of Hue City

**Tab. 1.** Land-use zoning data for population mapping

| Code                 | Land-use Type (Vietnamese)                     | English Equivalent                     | Relevance for Population Mapping                                             |
|----------------------|------------------------------------------------|----------------------------------------|------------------------------------------------------------------------------|
| ODT                  | Đất ở tại đô thị                               | Urban residential land                 | Include (100%)                                                               |
| ONT                  | Đất ở tại nông thôn                            | Rural residential land                 | Include (100%)                                                               |
| TMD                  | Đất thương mại, dịch vụ                        | Commercial-service land                | Include (optionally, weighted ~50%) – mixed-use buildings (e.g., shophouses) |
| SKC                  | Đất cơ sở sản xuất, kinh doanh phi nông nghiệp | Industrial land                        | Exclude                                                                      |
| DTS                  | Đất công trình sự nghiệp                       | Institutional/public facilities        | Exclude                                                                      |
| DGD                  | Đất giáo dục và đào tạo                        | Education and training land            | Exclude                                                                      |
| CQP                  | Đất quốc phòng                                 | National defense land                  | Exclude                                                                      |
| NTS / NNT            | Đất nghĩa trang, nghĩa địa, nhà tang lễ        | Cemetery, burial, funeral service land | Exclude                                                                      |
| CST                  | Đất cây xanh, công viên                        | Parks and green space                  | Exclude                                                                      |
| SON / NTS (thủy sản) | Đất sông suối, hồ ao, thủy sản                 | Water bodies, aquaculture land         | Exclude                                                                      |

## 2.2. Method

The population redistribution followed an area-weighted dasymetric approach, which is widely used in population mapping (Eicher and Brewer, 2001; Patel *et al.*, 2017). The steps are as follows:

### 2.2.1. Residential Building Selection

The accuracy of dasymetric population redistribution greatly depends on correctly identifying residential buildings. In this study, land-use zoning data were overlaid with the Google Building Footprints to classify buildings as residential or non-residential prior to population redistribution. Only buildings located within residential land-use zones which includes rural land-use (Land-use code ONT) and urban land-use (Land-use code ODT) were retained. Non-residential buildings (industrial, commercial, public facilities) were excluded based on the land-use zoning layer.

The steps are as follows:

#### Step 1: Preparation of Zoning Data

The zoning layer was standardized to the same coordinate system as the building footprint data (EPSG:4326).

#### Step 2: Spatial Overlay Operation

A spatial intersection (overlay) was performed between the building footprint layer and the zoning layer. For each building polygon, the corresponding land-use type was assigned based on the polygon in which it falls using spatial selection function in QGIS.

Buildings overlapping multiple land-use zones were assigned to the zone covering the largest proportion of their footprint.

#### Step 3: Residential Building Filtering

Only buildings falling within residential zones (e.g., “urban residential,” “mixed-use residential,” “low-density housing”) were retained. Buildings located entirely within non-residential zones (e.g., industrial, institutional, parks, or water bodies) were excluded.

### 2.2.2. Weight Calculation (Building Footprint Area)

For each ward, the sum of all residential building footprint areas was first calculated to represent the total available residential floor space within that administrative unit. The footprint area of an individual building was then used as a proxy indicator of its capacity to accommodate people—the assumption being that larger buildings typically house more residents than smaller ones. By comparing the area of each

building to the total residential building area in the ward, a relative weight for population redistribution was established. In other words, buildings with larger footprints were assigned a proportionally higher share of the ward's total population, whereas smaller buildings received a smaller share. This assumption follows the principle of areal weighting in dasymetric mapping, where building size is used as a surrogate for residential capacity in the absence of detailed data on building height, number of floors, or occupancy rates.

### 2.2.3. Population Redistribution

Ward-level population counts were proportionally redistributed to buildings according to their footprint area using the following equation

$$P_b = P_w \frac{A_b}{\sum_{i=1}^n A_i} \quad (1),$$

where  $A_b$  is area of each building,  $\sum_{i=1}^n A_i$  is sum of areas of all building with a ward/commune,  $P_w$  is population of corresponding ward/commune and  $P_b$  is estimated number of people in each building based on the assumption that each person in a ward uses the same house area.

### 2.2.4. Population Density Estimation

After the population was redistributed to individual buildings, the building-level population estimates were aggregated into a uniform grid of 100 m × 100 m cells to create a continuous population density surface. This gridded representation is very necessary because building footprints vary in shape and size, direct comparisons can be difficult. Aggregating population estimates into standard grid cells allows population density to be expressed as people per hectare, enabling consistent spatial comparison across the study area (Freire *et al.*, 2020). The gridded surface highlights intra-urban population distribution, revealing distinct patterns such as dense urban cores, mixed-use areas, and emerging peri-urban developments. This level of detail is crucial for urban planning and disaster risk management (Lloyd, Sorichetta and Tatem, 2017).

The aggregation is implemented based on the principle that each building's estimated population was allocated to the grid cell(s) it intersects. For buildings spanning multiple cells, the population was proportionally distributed based on the footprint area overlapping each cell. The total population in each grid cell was then computed by summing all contributing building estimates. The building-level population estimates were then aggregated into a 100 m × 100 m grid to produce a continuous population density surface, highlighting fine-scale settlement patterns across Hue City.

## 3. Results and discussions

### 3.1. Building-based population redistribution

The building-based population redistribution was implemented in QGIS using a series of spatial analysis tools, as described below:

1. Residential Building Selection

Overlay the building footprints with the residential land-use layer using the “Intersection” tool to generate a new *Residential\_House* layer containing only residential buildings.

2. Assigning Administrative Units to Buildings

Use “Join Attributes by Nearest” to link each residential building to its corresponding ward polygon from the administrative boundary layer, producing a *Residential\_House\_with\_WardID* layer with a unique ward ID attribute for each building.

3. Aggregating Total Residential Building Area per Ward

Apply the “Aggregate” tool to summarize the total residential building area by ward ID (SUM of footprint areas), creating an *Aggregated\_House\_in\_Ward* layer.

4. Updating Ward-Level Population Layer

5. Join the aggregated building area data to the ward administrative layer using “Join Attributes by Field Value” based on the ward ID, adding a new attribute *Total\_House\_Area* to each ward polygon.

6. Calculating Average Residential Area per Person

With the “Field Calculator” compute a new attribute *Area\_per\_Person* for each ward by dividing *Total\_House\_Area* by the known ward population.

7. Assigning Area-per-Person Values to Buildings

Join the *Area\_per\_Person* attribute from the *Aggregated\_House\_in\_Ward* layer back to the *Residential\_House\_with\_WardID* layer using “Join Attributes by Field Value”

## 8. Estimating Population per Building

Finally, estimate the population for each building using the “*Field Calculator*” by dividing the building footprint area (*Building Area*) by the corresponding *Area per Person* value.

The results of the building-based population redistribution for Hue City are illustrated in **Error! Reference source not found.**, which presents a high-resolution spatial distribution of estimated population density across the city. The analysis highlights clear patterns of urban concentration and peripheral sparsity. In the central part of the city, particularly around the Imperial Citadel, the wards of Phú Xuân and Thuận Hòa show exceptionally high densities of residential structures. This area, as the historical and administrative heart of Hue, contains a dense network of small to medium-sized housing units closely packed along narrow streets. Consequently, the estimated population in these central wards is significantly higher than in other parts of the city. This outcome not only reflects the actual concentration of inhabitants but also confirms the suitability of using building footprint area as a proxy for population distribution at fine spatial scales..

In contrast, the population density markedly decreases in the peripheral and mountainous regions, such as the former A Luoi and Nam Dong districts, located in the western and southern parts of Hue City. These areas are characterized by sparse settlements, rugged topography, and limited infrastructure, leading to lower concentrations of residential buildings and hence, lower estimated populations. This sharp gradient between the densely populated core and the sparsely inhabited outskirts highlights the effectiveness of the dasymetric mapping technique in capturing intra-urban demographic variation. One advantage of the dasymetric redistribution method used in this study is to allocate population based on the spatial extent of residential buildings while excluding non-residential zones, therefore enhances the reliability of the population estimates compared to traditional census maps. The results not only validate the methodology but also emphasize the importance of spatially detailed data in urban planning and resource allocation, especially in rapidly urbanizing regions like Hue. The fine-scale population surface generated through this approach provides critical insights for urban planners, emergency response teams, and infrastructure developers.

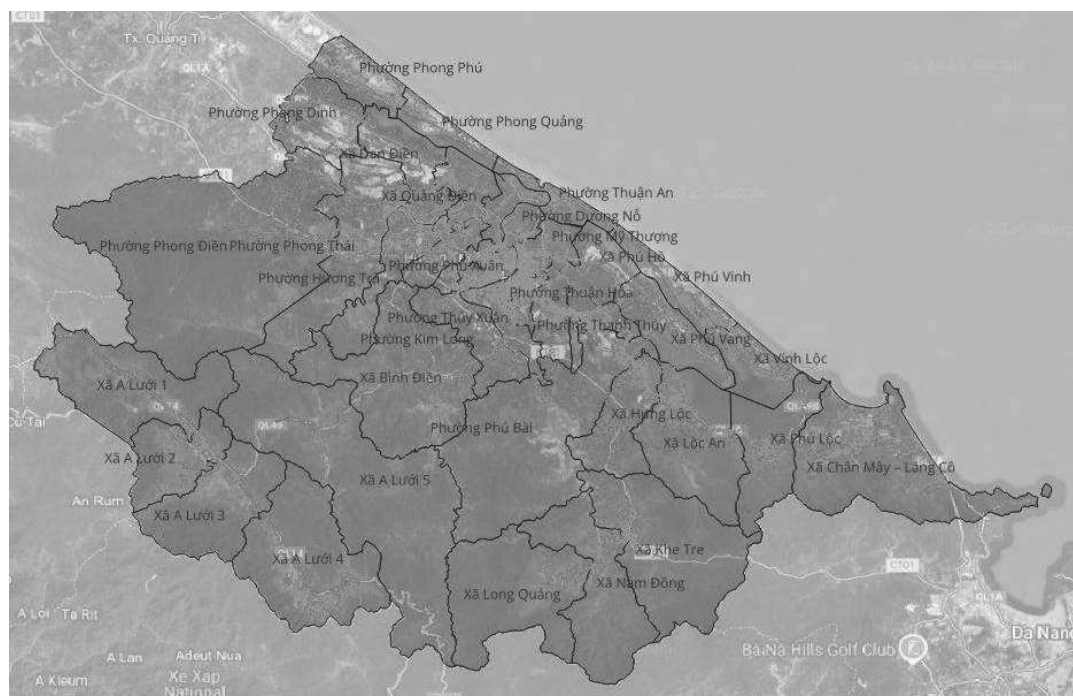


**Fig. 4.** Population redistribution based on the building areas

### 3.2. Population redistribution in raster format

Another widely adopted method for representing population distribution is the raster-based approach, in which the geographic space is divided into a regular grid of equally sized cells, and each cell is assigned a population value. This method is especially common because it provides a continuous spatial representation of human presence, facilitating integration with other raster datasets and enabling advanced spatial analyses across thematic layers such as risk zones, infrastructure, or environmental features.

Building upon the results of the building footprint-based population redistribution, a 100-meter resolution population raster was generated for Hue City (Fig. 5). The choice of 100-meter resolution for the raster-based population distribution map strikes a crucial balance between spatial detail and computational efficiency. This resolution is widely adopted in global and regional population mapping efforts (e.g., WorldPop, Global Human Settlement Layer - GHSL, High Resolution Settlement Layer - HRSL) and is particularly suitable for urban planning, disaster management, and environmental assessments, where understanding human presence at a neighborhood or block level is essential (Pesaresi *et al.*, 2016; Ehrlich *et al.*, 2019). This raster surface provides a high-resolution view of the population distribution, preserving fine-scale spatial variation while offering compatibility with a wide range of analytical tools and modeling frameworks. The pixel values represent the estimated number of people residing within each  $100 \times 100$  m cell, derived by aggregating the estimated population values from individual residential buildings.

To further enhance interpretability and support practical applications, the raster population density values were classified into three density categories: Low-density areas, defined as cells with fewer than 50 persons per hectare, Medium-density areas, ranging from 50 to 200 persons per hectare and High-density areas, with over 200 persons per hectare. These classifications are visualized in Fig. 6, revealing significant spatial contrasts across the city. High-density zones are concentrated in the urban core—especially around the historic Citadel and wards such as Phú Xuân and Thuận Hóa—where residential building concentration is highest. Another notable concentration of population appears in the coastal areas of Hue City, especially in ward of Thuận An and commune of Phu Vinh. These are traditional fishing communities where many fishermen and their families reside in compact, densely built-up settlements along the shoreline. Despite the limited land area, these zones exhibit high population densities due to the clustering of residential buildings, driven by historical land-use patterns and occupational needs. The small parcel sizes and intensive land utilization result in dense habitation patterns comparable to inner urban areas. This highlights the importance of considering socioeconomic and cultural factors, such as livelihood-based clustering, when modeling population distribution, as such communities may otherwise be underestimated in conventional mapping approaches. In contrast, low-density areas are prevalent in peripheral and mountainous regions, such as the former districts of A Lưới and Nam Đông, reflecting their sparse settlement patterns.

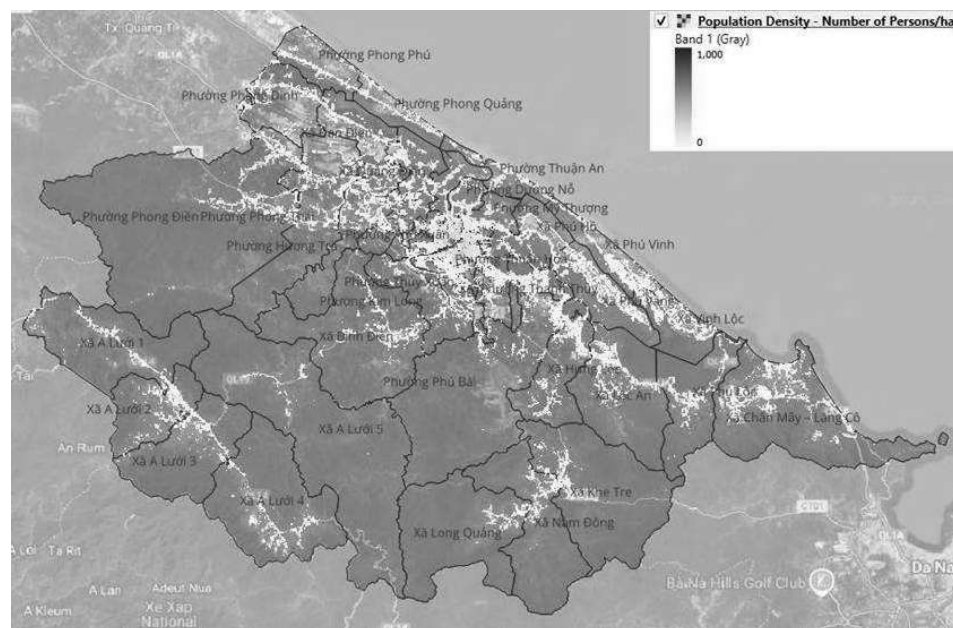


**Fig. 5.** Population density calculated by number of persons live in one hectar

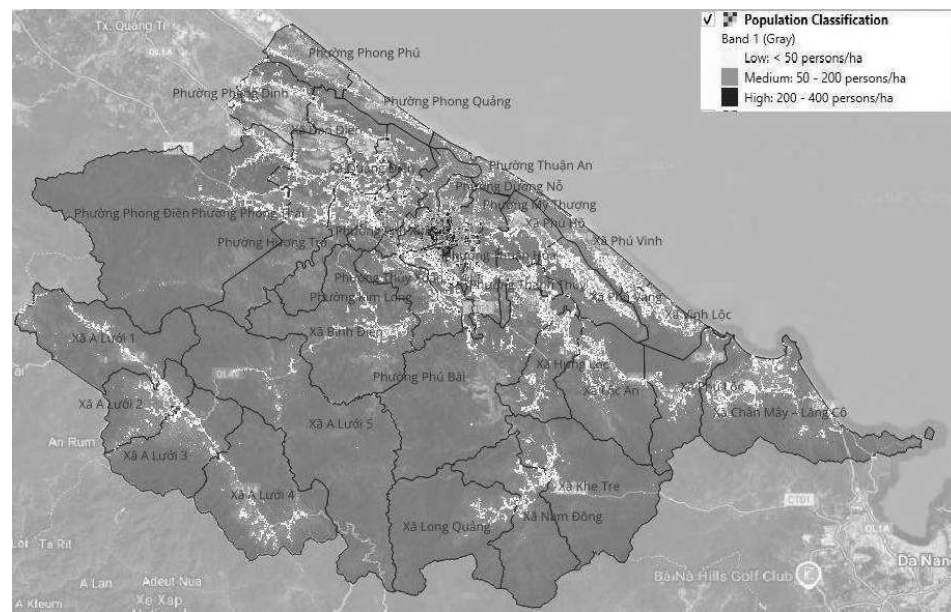


**Fig. 6** Density classification for 100x100m cells of Hue City

This raster-based population model not only improves visual and analytical accessibility but also supports more granular planning and risk assessment, particularly for emergency preparedness, urban service optimization, and infrastructure development.

### 3.3. Comparison of estimated population with WorldPop data

To assess the reliability of population estimates derived from Google Building Footprint data, a visual comparison was conducted using WorldPop data as a benchmark. WorldPop is a widely recognized global population dataset that provides gridded population estimates at relatively high resolution, and is often used for demographic analyses, planning, and policy support. In this study, the comparison focuses on Hue City, with Fig. 7(a) showing the population distribution based on Google Building Footprints, and Fig. 7(b) presenting the corresponding distribution from the WorldPop dataset.

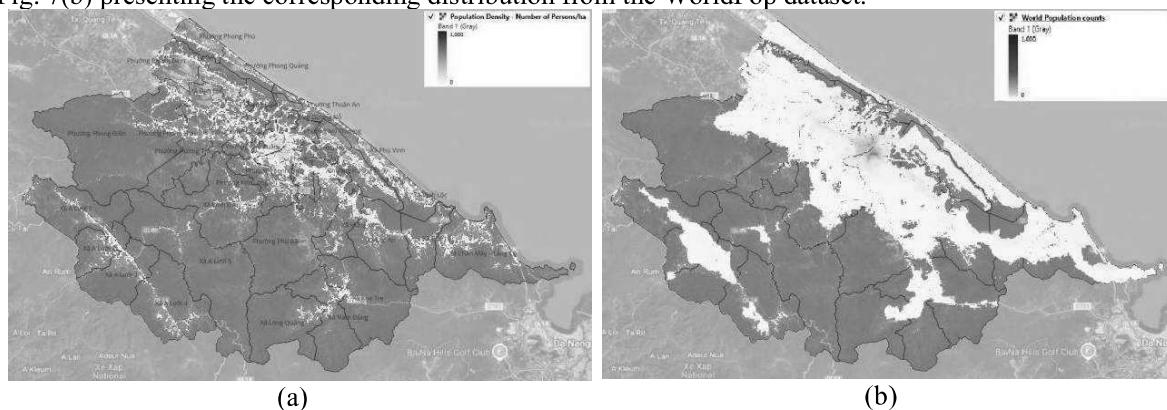


**Fig. 7.** Comparison of estimated population using google footprint and world pop data. (a) Population distribution by Google Building Footprints with ancillary data, (b) WorldPop data of Hue City.

The results reveal notable differences between the two population surfaces. The Google-based population map exhibits a finer granularity and appears to more precisely delineate actual residential areas. This precision is due to the nature of the input data: the Google Building Footprint dataset captures the physical presence of buildings at high resolution, allowing for a more accurate identification of inhabited structures. As a result, population is redistributed only to areas with known residential buildings, avoiding overestimation in non-residential or uninhabited regions.

In contrast, the WorldPop data tends to generalize population distribution over broader land areas. In many cases, it assigns population values to zones that include agricultural fields, forested land, or other undeveloped areas. Visual inspection against satellite imagery confirms that some of these zones are unlikely to be inhabited. This overestimation may stem from WorldPop's methodology, which relies

heavily on land cover classification and ancillary spatial covariates that can misclassify or overgeneralize land use, particularly in peri-urban and rural environments where settlements are more dispersed.

### **3.4. Impact of Google Building Footprints' accuracy to the estimated population distribution area**

Although Google Building Footprint data offers a valuable and high-resolution representation of built-up areas, it does not always accurately distinguish between residential and non-residential structures. This limitation can introduce notable errors in population estimation, particularly in contexts that demand high accuracy, such as infrastructure development, urban planning, and disaster risk management. In such cases, relying solely on footprint data may result in either overestimation or underestimation of population figures. To enhance reliability, ancillary data—such as land use zoning maps, cadastral records, or ground truth observations—are essential for filtering and validating residential buildings.

A significant challenge arises from the misclassification of certain non-residential features as buildings. For instance, in many regions of Hue City and similar urban-rural transitional zones, rectangular-shaped features such as family tombs, agricultural plots with fences, or even industrial yard markers may be mistakenly interpreted by automated systems as building footprints. These inaccuracies can distort the spatial distribution of estimated populations, especially in rural or peri-urban areas where non-residential structures are common but mimic the geometric characteristics of dwellings.

To further enhance the accuracy of population estimation using building footprint data, improvements can be made by integrating additional data sources—particularly high-resolution remote sensing imagery. Optical satellite images, aerial photographs, or LiDAR data can offer more contextual information to help differentiate between residential structures and non-residential features, such as tombs, greenhouses, or fenced agricultural plots. These supplementary layers provide crucial spectral, structural, and textural cues that are not captured in vector-based building footprint datasets alone.

In addition, the application of artificial intelligence (AI) models—particularly deep learning techniques—holds significant promise in addressing the misclassification problem. By training convolutional neural networks (CNNs) or object detection models on annotated remote sensing imagery, it is possible to automate the classification of buildings with higher precision. These models can learn to identify and exclude non-residential or artificial structures that resemble buildings in shape but differ in usage. When combined with spatial filters and official land use data, AI-enhanced analysis not only reduces false positives but also contributes to more reliable population redistribution.

Together, these improvements pave the way for a more robust framework for intra-urban population estimation. By leveraging both data fusion and intelligent classification, planners and researchers can better support sustainable urban development, resource allocation, and emergency preparedness in dynamic urban environments like Hue City.

## **4. Conclusions**

This study demonstrates a robust approach for estimating intra-urban population distribution by integrating Google Building Footprints with official ward-level census data and ancillary spatial layers such as land use data. By applying dasymetric mapping techniques and aggregating results into a 100 m × 100 m raster grid, the study presents a high-resolution population density surface that reveals settlement intensity patterns across the city, from densely populated central wards to sparsely inhabited mountainous regions.

A comparative analysis with WorldPop data further validates the strengths of our approach. While WorldPop provides consistent national-scale estimates, its tendency to allocate population to uninhabited areas such as forests, farmlands, or ceremonial grounds. The inclusion of these objects into the data for estimation of population can result in spatial inaccuracies at the urban scale.

While building footprint data offers significant advantages in delineating inhabited areas, it is not without limitations. Misclassifications, such as tombs or agricultural structures being identified as buildings can affect accuracy. To address this, the study emphasizes the importance of incorporating auxiliary datasets and the suggestion of applying AI models for distinguishing true residential buildings from similar-looking objects.

## **Acknowledgment**

This research was supported by the project “Giải pháp trích xuất dữ liệu nhà tự động từ ảnh chụp bằng máy bay không người lái (UAV) phục vụ công tác thành lập dữ liệu nền thông tin địa lý” under the funding of the Ministry of Natural Resources and Environment of Vietnam (Project Code: TNMT.2024.04.07). The authors gratefully acknowledge the project's financial and technical support, which made this study possible.

## References

- Boo, G. *et al.* (2022) 'High-resolution population estimation using household survey data and building footprints', *Nature communications*. Nature Publishing Group UK London, 13(1), p. 1330.
- Corbane, C. *et al.* (2019) 'Automated global delineation of human settlements from 40 years of Landsat satellite data archives', *Big Earth Data*. Taylor & Francis, 3(2), pp. 140–169.
- Ehrlich, D. *et al.* (2018) 'Remote sensing derived built-up area and population density to quantify global exposure to five natural hazards over time', *Remote Sensing*, 10(9). doi: 10.3390/rs10091378.
- Ehrlich, D. *et al.* (2019) *GHSL data package 2019 – Public release GHS P2019*. Publications Office. doi: doi/10.2760/062975.
- Eicher, C. L. and Brewer, C. A. (2001) 'Dasymetric mapping and areal interpolation: Implementation and evaluation', *Cartography and Geographic Information Science*. Taylor & Francis, 28(2), pp. 125–138.
- Fetai, B. *et al.* (2022) 'Inconsistencies in cadastral boundary data—digitisation and maintenance', *Land*. MDPI, 11(12), p. 2318.
- Freire, S. *et al.* (2020) 'Enhanced data and methods for improving open and free global population grids: putting 'leaving no one behind' into practice', *International Journal of Digital Earth*. Taylor & Francis, 13(1), pp. 61–77.
- Huang, X. *et al.* (2021) 'A 100 m population grid in the CONUS by disaggregating census data with open-source Microsoft building footprints', *Big Earth Data*. Taylor & Francis, 5(1), pp. 112–133.
- Leyk, S. *et al.* (2019) 'The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use', *Earth System Science Data*. Copernicus Publications Göttingen, Germany, 11(3), pp. 1385–1409.
- Lloyd, C. T., Sorichetta, A. and Tatem, A. J. (2017) 'High resolution global gridded data for use in population studies', *Scientific data*. Nature Publishing Group, 4(1), pp. 1–17.
- Mennis, J. (2009) 'Dasymetric mapping for estimating population in small areas', *Geography Compass*. Wiley Online Library, 3(2), pp. 727–745.
- Patel, N. N. *et al.* (2017) 'Improving Large Area Population Mapping Using Geotweet Densities', *Transactions in GIS*, 21(2). doi: 10.1111/tgis.12214.
- Pesaresi, M. *et al.* (2016) *Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014*. Publications Office of the European Union Luxembourg.
- Stevens, F. R. *et al.* (2015) 'Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data', *PLOS ONE*. Public Library of Science, 10(2), p. e0107042. Available at: <https://doi.org/10.1371/journal.pone.0107042>.
- Tiecke, T. G. *et al.* (2017) 'Mapping the world population one building at a time', *arXiv preprint arXiv:1712.05839*.
- Ural, S., Hussain, E. and Shan, J. (2011) 'Building population mapping with aerial imagery and GIS data', *International Journal of Applied Earth Observation and Geoinformation*. Elsevier, 13(6), pp. 841–852. doi: 10.1016/J.JAG.2011.06.004.