Global maps of 3D built-up patterns for urban morphological analysis

Mengmeng Li\textsuperscript{a,b}, Yuan Wang\textsuperscript{a,c}, Job F. Rosier\textsuperscript{a}, Peter H. Verburg\textsuperscript{a,d}, Jasper van Vliet\textsuperscript{a,*}

\textsuperscript{a} Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1111, 1081HV Amsterdam, the Netherlands
\textsuperscript{b} Department of Land Management, Zhejiang University, Hangzhou 310058, China
\textsuperscript{c} School of Geographic Sciences, Hebei Normal University, Shijiazhuang 050024, China
\textsuperscript{d} Swiss Federal Research Institute WSL, Zürcherstrasse 111, Birmensdorf CH-8903, Switzerland

A R T I C L E   I N F O

Keywords:
- Building height
- Urban form
- SAR
- Urban heterogeneity
- Land use intensity

A B S T R A C T

Horizontal and vertical patterns of built-up land are essential to analyse a range of environmental change impacts, such as exposure to natural hazards, urban heat islands, and trapping air pollution, as well as for decision making in this context. However, while data on horizontal patterns are abundant, they are relatively rare for vertical patterns. Here, we present global maps of 3D built-up patterns at a 1-km\textsuperscript{2} resolution for the nominal year 2015. These data are estimated using random forest models, fed with a wide range of spatial data and trained on reference data from all continents except Antarctica. Independent testing indicates that R\textsuperscript{2} values of the global models for built-up footprint, height, and volume equal 0.89, 0.73, and 0.84, respectively. Our results show that buildings worldwide are 6.16-m high on average, and total building volume is 1645 km\textsuperscript{3}, which is the equivalent of a solid cube of 12 km on each side. Yet, we find large variations in 3D built-up patterns, both within and across world regions. In particular, floor space per person exceeds 200 m\textsuperscript{2} in both Oceania and North America, while it is only 29 m\textsuperscript{2} in South Asia and 38 m\textsuperscript{2} in Sub-Saharan Africa. Our results provide novel insights into the global distribution of 3D built-up patterns and offer new opportunities for the assessments of urban environmental impacts. The global data for building footprint, height and volume can be downloaded from https://doi.org/10.34894/4QAGYL.

1. Introduction

Solutions to sustainability challenges such as climate change and biodiversity loss critically depend on the development of human settlements (McDonald et al., 2019; Seto et al., 2012). Human settlements provide shelters to the vast majority of the global population and are the location of economic development. Yet, they are also an important source of greenhouse gas emissions (Ramaswami et al., 2016), and contribute to the loss of biodiversity habitat (He et al., 2014; Ren et al., 2022). At the same time, human settlements are also increasingly affected by environmental change, such as river flooding (Winsemius et al., 2015) and urban heat island effects (Chapman et al., 2017; Guo et al., 2022). The impacts of human settlements on environmental change as well as the impacts of environmental change on human settlements depend on their extent and location, but also on their 3D patterns (Seto and Pandey, 2019). For example, compact tall buildings may deteriorate the urban thermal environment (Manoli et al., 2020) and enforce the concentration of air pollutants (Llaguno-Munitxa and Bou-Zeid, 2020). In contrast, low-density urban development generally increases travel kilometres (Ewing and Cervero, 2010) and may also affect food production and biodiversity conservation through additional land take (van Vliet, 2019). Ongoing uncertainty on environmental impacts of built-up area and its change could partly be attributed to a lack of representation of urban vertical pattern (Middel et al., 2014).

To represent the spatial heterogeneity of 3D built-up patterns, urban climatologists, for example, often characterize urban morphology using a few discrete landscape classes, mostly known as Local Climate Zones (Demuzere et al., 2022; Stewart and Oke, 2012). However, it is increasingly acknowledged that continuous characterization of 3D built-up patterns is essential to gauge the subtle variations in urban morphology (Lipson et al., 2022). Compared to 3D building information for individual buildings which are crucial for urban analytics (Biljecki and Chow, 2022; Labetski et al., 2022), gridded 3D built-up datasets at a coarser resolution are directly supportive for representing urban land-use conditions. In particular, 3D built-up patterns are linked with urban density or land-use intensity (Angel et al., 2021a), which has been studied extensively in the fields of urban planning and land use science (Angel et al., 2021b; Dovey and Pafta, 2013; Li et al., 2022).

\* Corresponding author.
E-mail address: jasper.van.vliet@vu.nl (J. van Vliet).

https://doi.org/10.1016/j.jag.2022.103048
Received 11 August 2022; Received in revised form 27 September 2022; Accepted 4 October 2022
Available online 7 October 2022
1569-8432/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Large-scale gridded datasets of 3D built-up patterns are derived using both indirect measurements and estimation approaches (Esch et al., 2022; Franzit et al., 2021; Huang et al., 2022; Lao et al., 2021; Liias and Stavrou, 2016; Yang and Zhao, 2022). Indirect measurements relate satellite imagery, specifically SAR data, directly to the height of buildings. Recently, the first global map of 3D building height was produced using this approach, indicating building height for the year ≈ 2013 (Esch et al., 2022). The advantage of this approach is that it does not require any reference data for the generation of results, although reference data remains essential for accuracy assessment. However, the disadvantage is the large computational requirements, making it impossible to reproduce or repeat. Estimation approaches relate to supervised classification approaches that relate explanatory variables to observed horizontal and vertical spatial patterns (Cao and Huang, 2021; Franzit et al., 2021; Li et al., 2020). The advantage of this approach is the relatively small computational requirements. Yet, such estimations require reference data for training validating and testing a model. While such data are available for a few regions, it is relatively scarce or absent elsewhere, hampering global mapping thus far (Li et al., 2020).

In this study, we present global maps of 3D built-up patterns at a 1-km² resolution for the nominal year 2015. These data are estimated using a random forest model that is trained with a unique set of reference data across the world, and including both urban and rural areas. These reference data combine readily available data on these properties, mainly from North America, Europe and China, with data derived from 3D city models developed for urban planning. To increase the representativeness, we complemented our reference data with manually classified ground truth for more than 10,000 tiles of 1-km² Google satellite imagery where street-view images are also available, covering smaller towns and villages. Together, our sample contained 79,186, 71,163, and 71,079 locations for building footprint, height, and volume, respectively. These reference samples are distributed across all continents except Antarctica. We estimated the three properties of 3D built-up patterns using three separate ensembles of random forest models. In order to account for large differences between world regions, we included socioeconomic indicators, such as the Gini index and GDP at a country level, in addition to these data at a pixel level. We first split our reference data to independently train, validate, and test our models. After the assessment of model accuracy, we trained our models with all available reference data to predict building footprint, height, and volume globally. These processes are elaborated in the following Materials and methods section.

2. Materials and methods

We estimated the footprint, average height, and volume for buildings, respectively, at a 1-km² resolution across the globe. For each property, we developed a separate ensemble of random forest models, trained on a set of between 71,079 and 79,186 points of reference data, depending on the respective property. We fed the models with a wide range of geospatial data, and satellite imagery. As illustrated in supplementary Fig. S1, our method is an improved version of the approach presented in Li et al. (2020), including optimized model structure, additional explanatory variables, and updated input data.

To ensure computational efficiency and to reduce noise, we masked all input data to cover only areas that have at least some impervious surface. For this mask, we used the World Settlement Footprint (WSF-2015, see Marconcini et al. (2020)), which outlines 10.41 million 1-km² grids that include at least some built-up land in the year 2015. Our results were estimated in a Mollweide equal area projection, and all input data were first re-projected into that coordinate reference system before any further analysis. We used the WSF-2015 because it has a high accuracy and robustness that outperforms other comparable datasets (Marconcini et al., 2020). We also applied this mask for collecting ground truth data and in the processing of explanatory variables.

2.1. Explanatory variables

To select explanatory variables, we used three criteria: First, we include only variables that are expected to provide information on 3D built-up patterns. Second, we selected dataset that was available for the year 2015, or at least close to that year, to ensure temporal consistency. Third, we only wanted to include data that was directly observed rather than being downsampled, to ensure independency of these variables. These three criteria yielded 35 datasets (see Table S1), which can be further grouped into four categories according to their imaging modes or sources: Optical remote sensing data, radar data (SAR), remote sensing-derived indices (RS-derived), and other data.

Optical Remote Sensing data includes cloud-free Landsat-8 imagery for all available spectral bands for the year 2015. We included optical data because reflectance values of various bandwidths are associated with the extent and intra-environment of impervious areas (Yuan and Bauer, 2007). We did not use the WSF-2015 as a mask for optical remote sensing data, since Landsat-8 imagery is not only responsive to buildings, but also to vegetation and water, which might include valuable information for this analysis. Thus, we expect that reflectance values at a larger scale provide valuable information on building structure at a smaller scale. For each Landsat band, the median value for each pixel is computed in cloud-free and shadow-free conditions, to produce the representative values for the timespan of a year while excluding outliers, followed by the aggregation of these median values into a 1-km² resolution using a mean function (see Fig. S2).

Radar (SAR) data was used from Sentinel-1 imagery. We expect that radar data are relevant for the estimations of building height building volume, because radar backscatter signal has been found responsive to surface roughness (Zhu and Bamler, 2010). We used imagery from the winter seasons (December-February for the Northern hemisphere, and June-August for the Southern hemisphere) around the years 2015 and 2016, in order to limit the influence of vegetation on SAR backscatter (Chen et al., 2020). Apart from buildings, other objects could also yield backscatter, such as topographic relief and vegetation outside the built environment. Therefore, we only accounted for backscatter coefficients of within the WSF-2015 derived mask. Exploratory data analysis showed that higher buildings are often replaced because of the side-looking SAR configuration. Therefore, we applied a 20-m buffer to the WSF-2015 mask (i.e., 2 pixels in Sentinel-1 images), see Fig. S2 for more detail. Polarization modes VV and VH were calculated for all pixels. When SAR data was missing in a location for the target date, we used data from one month earlier or later. For each 10-m pixel i in an individual tile, we firstly averaged time-series backscatter coefficients available in the study period for SAR-VV(i), SAR-VH(i), and the maximum of the two, i.e., SAR-MAX(i). Then, we aggregated the mean backscatter coefficients for 10-m SAR-VV(i), SAR-VH(i), and SAR-MAX(i) into the 1-km² grids using the mean of values within the built and buffered areas, yielding three explanatory variable layers, i.e., SAR-VV, SAR-VH, and SAR-MAX.

RS-derived variables include the vegetation index (EVI), land surface temperature (LST), night-time light (NTL), and several Landsat-derived indices including Normalized Difference Built-up Index (NDBI), Normalized Difference Vegetation Index (NDVI), Normalized difference Bare Land Index (NBLI), and Urban Index (UI). Here, the MODIS EVI product, which is available for every 16 days, is used to represent vegetation conditions. All EVI layers available for the year 2015 are then aggregated into three variables using maximum, mean, and minimum functions, separately. The NDBI derived from the MODI1A2 product, in which daytime and night-time are independently stored. We averaged all LST layers for daytime and night-time in 2015, respectively. For the Landsat-derived indices, we used NDVI, NBLI, NDVI, and UI as explanatory variables. The variable NTL was derived from VIBS night-time light, which was available as monthly layers in the form of 500 m × 500 m grids. We synthesized all these monthly layers in 2015 into largest night-time light intensity for each grid, and then aggregated them into 1-km² grids using a mean function (see Fig. S2 for details).
In addition to satellite imagery and derived products, we also used other data as explanatory variables, which include impervious area, accessibility, road networks, topography, GDP per capita, and Gini coefficient. We expect that impervious surface density, road density, and accessibility are correlated with 3D built-up patterns thus providing information indirectly, while we expect that topography could be used to correct the signal from SAR backscatter (van der Wal et al., 2005). We expect that the WSF-2015 (impervious surface) correlates strongly with building footprint, and thus provide valuable information for building volume. Moreover, we used accessibility-to-cities to represent travel time to the nearest populated settlements in 2015 (Weiss et al., 2018). Vector road data were used to produce five road-density maps that ranged from highways to local roads (Meijer et al., 2018). A density map for all roads combined was also included in building the model. We used the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) as the source for DEM, which was further used to derive slope and aspect.

We used data on Gross Domestic Product (GDP) per capita and Gini index values for (sub)national administrative units (World Bank, Gen-naioi et al. (2013), and (Solt, 2020)). These variables were not included in Li et al (2020) on which this study builds. Yet, we expect that GDP per capita is able to capture the heterogeneity of buildings across countries from an economic perspective, and this heterogeneity is much larger on a global scale than for the countries included in Li et al (2020). We did not use gridded GDP data (Kummu et al., 2018), as their downscaling is based on other gridded datasets similar to those used in this study, which would create redundancy and possibly circularity. Gini coefficients reflect economic inequality within a country, and we expect this to provide information about the probability of findings in specific types of urban structure.

2.2. Reference data

To train our models, we collected reference data from multiple sources worldwide. These reference data include publicly available data, commercially available data, and a large number of data points that were generated manually for this study. We only included reference data that represents building properties for a range of three years from the nominal year of 2015, i.e., between 2012 and 2018.

Building heights for Europe for the year 2012 (BuildingHeight-2012) were taken from Copernicus Global Land Service (https://land.copernicus.eu), and were further integrated with building footprints collected from OpenStreetMap (OSM). We only consider 1-km² grids where the proportion of OSM building footprints with valid height values reported in BuildingHeight-2012 exceeds 80% of all the building footprint area. We expected that this threshold could exclude grids where a large proportion of new buildings were constructed after the production of the building height dataset, given that BuildingHeight-2012 was produced often earlier than the updated OSM buildings. For the U.S., we collected publicly available datasets from local governmental websites for the year 2015 (https://hub.arcgis.com). These datasets include vector data of building footprints with vertical properties for 27 urban areas that ranged from megacities like New York and Los Angeles to counties that only include small villages in remote areas. Building height data for China was available for large cities for the year 2015, expressed as the number of floors (https://www.amap.com). Here, we assume that each floor is 3 m high (Leichtle et al., 2019). Our preliminary evaluation suggested a relatively low model performance for China. We manually overlaid these data with the building footprints derived from Google Maps’ VHR satellite imagery and found that the low model performance was most likely caused by the underrepresentation of buildings on the peripheries of cities. Therefore, we manually removed the 1-km² grids that contained such large omissions. For other World regions, little or no reference data was publicly available. Therefore, we acquired 3D building data from Visicom (https://visicomdata.com/), for selected areas across the world. These building data exist in vector format with Level of Detail-2 quality (LoD-2, i.e., multiple heights per building, see Biljecki et al. (2016) for more detailed descriptions), and were originally developed to assist urban planning in respective regions. From this source we acquired data of 43 cities distributed over all regions outside Europe, the United States, and China, and predominantly in less-developed regions such as Latin America, Southeast Asia, and Africa.

Publicly accessible reference data for 3D building structure are predominantly available for larger cities. Therefore, we complemented these reference grids with empirical 1-km² grids that represent smaller settlements. For this, we used Google API to randomly download over 10,000 Very High Resolution (VHR) satellite images far from large cities (travel time to populated settlements > 10 min, and impervious surface density > 0). For the visual interpretation process, we used a 50 × 50 fishnet to manually specify building footprint, as well as Google Street View to estimate building height. Together, we collected a large number of training sites ranging from 71,079 to 79,186 for building footprint, building height, and building volume, distributed over all continents except Antarctica (Fig. 53, and Table 52).

2.3. Experimental set-up and validation

We built three Random Forest (RF) model ensembles to predict building footprint, height, and volume, respectively. Each model ensemble consists of 100 independent RF models. These models together yielded 100 predictions for each of the 10.41 million grids in unknown areas. We use the mean values as their final predictions, and quantify our model uncertainty as the Coefficient of Variation (CV) of the 100 predictions.

In general, tree number in a RF model is positively associated with model performance. Yet, the improvement becomes less prominent as the tree number increases. Therefore, after initial experiments, we set the tree number as 150, and set the minimum sample number at each leaf node as 5.

For each model ensemble, we first set apart 20% of our reference data for independent testing, and used the remaining 80% for training and validation. For each RF model, the 80% training and validation sample was randomly divided into two subsets of 70% and 10% for training and validation, respectively. Mean values of the 100 predictions were assigned to the final predictions of 20% test samples, of which observed values were compared with their predictions to independently assess model robustness.

We calculated the importance of input variables for each of the 100 models in the three model ensembles. The mean of all 100 importance values of a specific variable was assigned to its final importance value. Variable importance ranges from 0 to 1, and the sum of importance values for all variables equals 1. To reduce overfitting, we iteratively removed the least important variable until all variable importance values were ≥ 0.5%.

Subsequently, we compared the mean value of the 100 predictions with observed values from the 20% test collection to assess the overall accuracy of our model ensembles. We assessed model performance at a global scale, as well as for ten World regions separately.

The world regions (Canada and United States, China, Europe, South Asia, Latin America, Middle-East and North America, Oceania, Russia and Central Asia, Southeast Asia, and Sub-Saharan Africa) were taken from the World Bank, with further subdivision of (i) East Asia and Pacific and (ii) Europe and Central Asia to reflect the variations in both urban development and socioeconomic dynamics (see Fig. S4). Specifically, the East Asia and Pacific region was further subdivided into China, Southeast Asia, and Oceania, which follows the observation of rapid development of urban areas in China, accounting for 23% of global built-up area expansion between 1992 and 2015 (van Vliet, 2019). At the same time, the differences in wealth and lifestyle between Oceania and Southeast Asia merit a further subdivision of these areas in our analyses. Moreover, Europe was separated from Russia and Central Asia given that...
Europe has experienced rapid urban growth in recent years, but this development was relatively marginal in Russia and Central Asia. Therefore, our division represents coherent groups of countries from a socioeconomic point of view. It should be noted that the world regions were only used for presenting and discussing the results in an aggregate way, while we used one global model for our pixel-level estimates.

2.4. Analysing the distribution of building footprint, height, and volume

Following the Eqs. (1)–(3), we calculate the global sum of building footprint ($F_{\text{sum}}$), average of building height ($H_{\text{ave}}$), and global sum of building volume ($V_{\text{sum}}$), respectively.

$$F_{\text{sum}} = \sum_{i=1}^{N} f_i \quad (1)$$

$$H_{\text{ave}} = \frac{\sum_{i=1}^{N} (f_i \times h_i)}{\sum_{i=1}^{N} f_i} \quad (2)$$

$$V_{\text{sum}} = \sum_{i=1}^{N} v_i \quad (3)$$

where $f_i$, $h_i$, and $v_i$ are building footprint, height, and volume of the pixel $i$, respectively. $N$ is the total number of pixels with impervious surface presents according to the WSF-2015.

We analysed 3D built-up patterns along the urban–rural gradient in selected cities and in more rural areas across the ten world regions, to illustrate the variation in urban morphology. Specifically, average building footprint per pixel, average building height and average building volume per pixel are used to delineate changes in 3D urban structure along urban–rural gradient, in which water surface was excluded based on the permanent water layer from Pekel et al. (2016) to facilitate comparison between inland cities and coastal cities. We analyse the three properties per 1-km buffer ring with a total distance of 50 km, where urban centres are represented by the centroid points derived from the polygon geometries in GHS Urban Centre Database 2015 (Fiorczyk et al., 2019). Moreover, to delineate 3D built-up patterns in more rural areas, we analyse frequency of the 1-km pixels with a relative small footprint and low buildings for ten world regions.

Finally, we assessed the occupation buildings per person for the ten world regions (Fig. S4). For this, we summarize total population, total building footprint, average building height, and total building volume for the ten world regions, followed by building footprint per person, building volume per person, and consequently floor area per capita (assuming that each floor is 3-meter-high on average). Population in the year 2015 for each country was obtained from the United Nations Population Division (https://population.un.org/wpp/), and was then aggregated into the ten world regions.

3. Results

3.1. Global distribution of 3D built-up patterns

Our results show a total building footprint area of 264 thousand km$^2$,
a total building volume of 1645 km$^3$, and an average building height of 6.16 m, globally. This total building footprint is about the size of New Zealand, while the total building volume is equivalent to a solid cube of almost 12 km on each side, which is enough to fill Lake Ontario. Globally, the correlation between building height and building footprint at a pixel level is 0.57, indicating the relevance of mapping these properties separately.

The distribution of building footprint and height further suggests that sparse and low-rise buildings dominate the globe, and it varies substantially across space, both between world regions, and along the rural–urban gradient. The variation in building height and footprint is illustrated in Fig. 1, showing areas in New York and Paris with a dense footprint and high-rise buildings, mixed patterns in Beijing, and relatively dense but low urban in Lagos and Java. Generally, in low and medium income regions such as South Asia, Latin America, Southeast Asia, and Sub-Saharan Africa, there are proportionally more pixels characterized by dense building footprint but low height, including for example informal settlement development therein. Whereas in well-developed regions, particularly West Europe and East China, pixels with dense building footprint tend to have higher buildings. In Canada and USA and Oceania, however, urban vertical growth is less prominent, which could be attributed partly to the availability of land for development.

Along the urban–rural gradient in larger cities, urban centres generally have denser and higher buildings, and both decrease with increasing distance from the centre (Fig. 2). Yet, exceptions exist, such as Kano City in Nigeria, where buildings in the centre are low compared to other cities and this height barely decreases with the distance from the centre (Fig. 2b). At the other extreme, in the centre of New York, buildings are very high, and their height drops dramatically between 5 and 8 km from the centre. Distance decay patterns are also visible in the distribution of building footprints. This pattern is clearest in the centre of Jakarta, where buildings are moderately high but their footprint is up to half of the land surface. Mumbai, on the other hand, shows the lowest peak in building footprint, and this peak is not around its centroid, but in a ring between 3 and 12 km from the urban centroid, largely owning to the mismatch between urban centroid and downtown area as a result of the disorderly urban expansion and land fragmentation around coastal area.

Of the ten world regions, building footprint per person is the highest in Oceania, followed by Canada and USA (Table 1). In contrast, the building footprint per person in Sub-Saharan Africa is only 23 m$^2$, and in South Asia it is only 18 m$^2$. Buildings in Sub-Saharan Africa are the lowest on average (4.77 m), whereas buildings in Middle-East and Northern Africa are the highest (7.36 m). Building volume per person ranges from 86 m$^3$ in South Asia to 682 m$^3$ in Oceania, which is the equivalent to a floor area of 29 m$^2$ and 227 m$^2$, respectively, when assuming an average floor height of 3 m.

Although sparse and low-rise buildings dominate rural areas globally, we see different patterns across the ten world regions (Fig. 3). In Sub-Saharan Africa, for example, there are many locations with dense low-rise buildings, represented by a high density of pixels with low building height. Conversely, in Europe there are more areas with sparser but higher buildings, typically with 2 floors. We also see mixed patterns in South Asia and Southeast Asia.

### 3.2. Model performance and uncertainty

The exclusion of variables variable importance values were $\geq 0.5\%$ process led to 22, 31, and 25 variables for estimating building footprint,
height, and volume, respectively (Table S3). Results of the variable importance analysis mostly confirm our prior hypotheses of the relevance of the selected input data, as 31 out of 35 variables have a variable importance > 0.5% for at least one of the three models. The exclusion of less important variables has minimal effect on model performance (Fig. S5).

Independent testing indicates that $R^2$ values of the global models for building footprint, height, and volume equal 0.89, 0.73, and 0.84, respectively, but with variations across world regions (Figs. 4, S6, and S7). Overall, model uncertainty reflected by CV values for footprint shows an inverse U-shape, whereas model uncertainties for height and volume constantly increase (Fig. S8). The variations of model uncertainty between different world regions are not identical, yet uncertainties are roughly 50% lower than a previous study of building height, footprint and volume (Li et al., 2020), indicating the robustness of our models, which is likely a consequence of the large and more heterogeneous sample of reference data.

4. Discussion

The global maps of 3D built-up patterns provide an image of the
We find that building footprint per person and building volume per person in low and medium income countries (LMIC) is generally much lower than in the U.S., Europe, and Oceania, which likely reflects the difference in wealth between these regions. Consequently, these differences also show the different per person contribution to the global competition for land (van Vliet, 2019). In response, high-rise buildings and compact urban development have been proposed as pathways for more sustainable urban development (Cortinovis et al., 2019). Our results reveal hotspots of dense and high buildings in urban centres, predominantly in Europe, the U.S., China, and the Middle-East. However, our results also reveal much larger areas that are sparsely covered with buildings and that are on average only one or two stories high. In the U.S., for example, these results suggest urban sprawl, while in LMIC these often reveal patterns of mixed urban and agricultural use (Agergaard et al., 2019).

The global maps presented in this paper have a higher accuracy, show a smaller systematic error, and are more accurate (Table S4) than the recent WSF-3D dataset (Esch et al., 2022), currently the only available global data of building height, footprint and volume. The lower Root Mean Square Error (RMSE) and lower Mean Absolute Error (MAE) of the data presented here can likely be explained by the coarser resolution of our maps, as errors within a single pixel cancel out upon aggregation. However, this is not the case for systematic errors (SE). The SE of \(-0.00\) km\(^2\)/km\(^2\), \(-0.05\) m, and \(-0.03 \times 10^3\) m\(^3\)/km\(^2\) for building footprint, height, and volume, respectively, indicate very little bias in our estimations, which is in contrast with the WSF-3D. In addition to the WSF-3D data, two global datasets of Local Climate Zones have been presented recently (Demuzere et al., 2022; Zhu et al., 2022). Local Climate Zones are defined as discrete classes rather than continuous values, which hampers a direct comparison between accuracy metrics.

Both night-time light and Landsat band 5 have a larger variable importance than the different SAR bands included, despite SAR being widely recognized as responding to surface roughness and thus building height (Frolking et al., 2013). Also, interestingly, the Gini index and the GDP have a large explanatory power at the pixel level, despite these values being provided at the national or subnational level only. Both observations show the relevance of supervised classification algorithms over direct measurements, as they allow to incorporate a wider range of input data that potentially explain 3D built-up patterns. At the same time, Random Forests are relatively simple machine learning algorithms, and our approach only used average values of the various input data at a 1-km\(^2\) resolution, ignoring the variation and patterns that might exist in the input data within this 1-km\(^2\) pixel. Consistently, we expect that Deep Learning approaches, such as Convolutional Neural Networks, might provide more accurate results, due to their capacity to detect spatial patterns in satellite imagery. Yet, such approaches are computationally demanding thus reducing the simplicity of our approach.

To train and validate our models, we manually collected a large number of reference grids in rural regions, which are highly labour-intensive. In recent years, crowdsourcing platforms such as Geo-Wiki...
have been available for collecting reference data from the visual interpretation of satellite and aerial imagery (See et al., 2022a). Moreover, the distribution of 3D built-up patterns is highly dependent on the WSF-2015 dataset. While WSF-2015 is one of the most accurate global built-up layers (Marconcini et al., 2020), it can still be improved through extensive engagement of more volunteers and stakeholders using the crowdsourcing platforms as a tool (See et al., 2022b).

The global maps of building height, footprint and volume are publicly available and can serve as an input for future studies on urban sustainability, including analyses of urban form (Taubenböck et al., 2020), exposure to natural hazards (Paputryno et al., 2020), urban climate impacts (Cao et al., 2022; Gago et al., 2013), and energy consumption within the built-environment (Creutzig et al., 2015). Moreover, information on 3D built-up patterns will also benefit population density mapping (Leyk et al., 2019) and identification of local climate zones (Demuzere et al., 2019). Finally, these data could provide valuable input for more informed policies and assessments at regional to global scales, thus avoiding the misinterpretation that all built-up land is similar.

Author contributions

ML and JvV developed the model and conceived the study. ML, YW and JFR prepared and processed the data. ML visualized the outputs. JvV supervised this study. All authors contributed to the preparation of this manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The generated maps, as well as the input data and algorithms are publicly available at https://doi.org/10.34894/4QAGYL. Original data can be accessed using links provided in this article.

Acknowledgement

We would like to acknowledge Elio Koks and Hannes Taubenböck for the discussion on methodological improvement in earlier stages of the project. ML was supported by China Scholarship Council (201706510011) and the Major Projects of Social Science Foundation of China (212DA064). JvV was supported by Netherlands Organization for the development on methodological improvement in earlier stages of the project. ML was supported by China Scholarship Council (21ZDA064). JvV was supported by Netherlands Organization for the development on methodological improvement in earlier stages of the project. 

Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2022.103048.

References

