

A scalable data collection, characterization, and accounting framework for urban material stocks

Hadi Arbabi  | Maud Lanau  | Xinyi Li | Gregory Meyers  | Menglin Dai | Martin Mayfield | Danielle Densley Tingley

Department of Civil and Structural Engineering, The University of Sheffield, Sheffield, UK

Correspondence

Hadi Arbabi, Department of Civil and Structural Engineering, Sir Frederick Mappin Building, Mappin Street, Sheffield, S1 3JD, UK.
Email: h.arbabi@sheffield.ac.uk

Funding information

The Multispectral Advanced Research Vehicle was supported under EPSRC UK Collaboratorium for Research in Infrastructure & Cities: Urban Observatories (Strand B) [EP/P016782/1]; HA and DDT time was supported under EPSRC Active Building Centre [EP/S016627/1]; XL and GM time was supported under EPSRC UKCRIC-CORONA: City Observatory Research Platform for Innovation and Analytics [EP/R013411/1]; ML and DDT time was supported under EPSRC Multi-Scale, Circular Economic Potential of Non-Residential Building Scale [EP/S029273/1].

Editor Managing Review: Niko Heeren

Abstract

Building stocks represent an extensive reservoir of secondary resources. However, common bottom-up characterization of these, often based on archetypal classification of buildings and their corresponding material intensity, are still not suitable to adequately inform circular economic strategies. Indeed, these approaches typically result in a loss of building-specific details, and a building stock characterization in terms of material mass, for example, glass, rather than component, for example, window. To deliver this higher resolution of details, a scalable approach to urban stock characterization, that enables a bottom-up estimation of building stocks at the building component level, is needed. In this paper, we present a framework to automate the characterization of urban stock. By using and combining a mobile-sensing approach with computer vision, urban stocks can be captured as 3D surface maps allowing the identification and semantic classification of stock objects, components, and materials. We demonstrate the potential of this framework through a case study of a neighborhood in Sheffield, UK, by using a prototype workflow comprising a custom-made mobile-sensing platform and an existing suite of neural network models to calculate an estimate count of buildings external doors and windows. The prototype implementation of the framework achieves comparable total and building-level component counts with those achieved through manual human counts. Such automated estimation of components enables an understanding of opportunities across the circular economic hierarchies and informs stakeholders across the supply chain to better prepare for the implementation of circular strategies including building refurbishments.

KEYWORDS

bottom-up analysis, building material stock, environmental modeling, industrial ecology, machine learning, mobile mapping

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Industrial Ecology* published by Wiley Periodicals LLC on behalf of Yale University

1 | INTRODUCTION

With 189 nations participating, the Paris Agreement encourages a reporting mechanism for emissions in an attempt to avoid global temperature rises of over 2°C (2015). Given that somewhere between 20% and 40% of emissions are attributable to the built environment, global material stock characterization and accounting is essential for decarbonization and avoiding an extreme climate change fallout (World Economic Forum & Boston Consulting Group, 2019). Achieving these climate goals will require radical action across nations, cities, and sectors, but progress has stalled. In the United Kingdom, for example, meeting the country's new net-zero emissions ambition by 2050 will be a considerable challenge, that requires urgent action (Committee on Climate Change, 2019).

Decarbonization of the built environment must tackle two critical areas: (i) embodied impacts and (ii) operational impacts. The first is in part delivered by a shift to the circular economy, as this would reduce long-term material consumption. In addition to the critical roles the built environment plays in our socio-economic metabolism (Haberl et al., 2017), its stocks are also an extensive repository of secondary resources. Anthropogenic stocks have undergone a 23-fold increase and are still on the rise (Krausmann et al., 2017). These anthropogenic resource reservoirs are thus an opportunity for resource recovery through circular economic strategies, especially in urban areas, which are characterized by a dense accumulation of built environment stocks.

Material stock and flow accounting can assist in predicting future material demand based on the current stock age and likely replacement rates (Tazi et al., 2021). This is especially important for those nations with older building stocks, such as the United Kingdom, where 85% of the building stock which will exist in 2050 is predicted to have already been built (Edwards & Townsend, 2011). This also means that operational impacts need to be reduced through energy retrofit of the existing stock. Thus, the more we understand about this stock, the better it can be maintained and *mined* in the future.

1.1 | Building stock accounting

Characterizing built environment stocks can be done through two main approaches, namely top-down and bottom-up. The top-down approach relies on mass balance and lifetime distribution to model the accumulation of material stock within a system over time. This approach has proved useful to achieve an overview of stock dynamics over long periods of time, allowing the identification of patterns and drivers that can be used to benchmark future stock accumulation (Lanau et al., 2019). The bottom-up approach, although time intensive, is preferred when generating detailed information on the physical arrangement of the system under study (Recalde et al., 2008). Bottom-up approaches consist of counting a given stock *piece by piece*, differentiated in terms of materials and stock composition. Spatial differentiation of results can also be achieved through integration with geographical information systems (GIS). Such GIS-based bottom-up approaches were first used by Tanikawa and Hashimoto (2009), where they studied the accumulation of materials in the built environment of two neighborhoods in Manchester, UK and Wakayama, Japan. Since then, a number of studies have been conducted at different spatial scales, on all parts of the built environment (Augiseau, 2017; Lanau & Liu, 2020; Tanikawa et al., 2015) or specific parts, such as roads (Guo et al., 2014), subway (Lederer et al., 2016), pipe networks (Wallsten et al., 2013), and cable networks (Krook et al., 2011). A number of case studies also focus on solely characterizing building stocks of cities, such as Vienna, Austria (Kleemann et al., 2017); Esch sur Alzette, Luxembourg (Mastrucci et al., 2017); Grenada (Symmes et al., 2020); Padua, Italy (Miatto et al., 2017); Melbourne, Australia (Stephan & Athanassiadis, 2017); Chiclayo, Peru (Mesta et al., 2019), as well as nations such as Germany (Ortlepp et al., 2016, 2018).

1.2 | Delivering the circular economy

For the circular economy, the building stock has proved to be a critical part of the built environment as it hosts a wide variety of easy-to-access and easy-to-recover materials above the ground (Lanau & Liu, 2020). The implementation of a circular economy faces a number of challenges industry wide, including a limited awareness across the supply chain (Adams et al., 2017), concerns about the consistency of flows of returned goods, and an unclear market demand for secondary resources (Guldmann & Huulgaard, 2020). Overcoming such barriers requires highly detailed modeling of building stocks so that stakeholders throughout the supply chain can obtain the exact quantity and quality of secondary resources, for example, construction bricks or panes of glass, that would be recoverable from a specific building nearing demolition (Arbabi et al., 2020). To enable this, estimates of material quantity and quality need to be spatially explicit and be measured at a building level. However, because building inventory datasets are heterogeneous in terms of construction type, periods, and use, buildings are often classified into archetypes, according to their characteristics to homogenize the datasets resulting in a loss of detail. These geo-located and building-specific details are, however, crucial to circular economic strategies.

Another shortcoming of common bottom-up modeling of building stocks for circular economy is the quantification of stocks in terms of materials mass rather than building components. In a circular economic paradigm, component reuse is always preferable to material recycling (Hyman et al., 2013). Component-level information is critical when aiming to estimate circular economic potential of a building stock. The number of windows shows the potential for future replacement, remanufacturing, and reuse, whereas mass of glass only shows the scope for recycling (Arora et al., 2019). Achieving a truly high-resolution stock characterization would require inspection of individual stock buildings. Undertaking such inspections across whole cities can be prohibitively time intensive using current approaches.

The more recent uses of remote sensing can and has, to some extent, mitigated some of these associated problems (Mao et al., 2020). These, however, are still lacking in the resolution they offer in stock characterization, especially with respect to components at a building level (Peled & Fishman, 2021). Current archetype-based approaches to stock accounting require increasingly more detailed archetypes to increase precision of the results. This inherently poses a data collection challenge and raises the level of uncertainty at building level (Ortlepp et al., 2018). Barring a few choice countries, including Denmark and Germany, such inventories, or data required to simply assemble them, do not readily exist in many others (Lanau et al., 2019). Given the size of the challenge, that is, building-specific characterization of the entirety of the stock involving buildings numbering in the millions for a given city or nation, there is a need for a scalable approach to urban stock characterization that enables a bottom-up accounting of stock at component level.

2 | BUILDING STOCK CHARACTERIZATION: RELEVANCE, CHALLENGES, AND OPPORTUNITIES FOR A SCALABLE FRAMEWORK

We propose a framework integrating mobile-sensing approaches and workflow automation in urban stock modeling to start addressing the scalability in stocks characterization. First, we define scalability by borrowing a definition common to systems engineering (Bondi, 2000; Jogalekar & Woodside, 2000; Weinstock & Goodenough, 2006). Scalability of an approach is measured as the extent to which it can be repeatedly extended to handle increasing workload with an optimized cost-effectiveness and without additional resource penalties. The methods by which we characterize stocks would need to become both faster at a building level and more efficient at a city/country level. The use of mobile/remote-sensing, computer vision, and deep learning methods are likely candidates to achieve *scalability*.

2.1 | A scalable framework

Computer vision can be defined as the “task of learning the qualitative representation of visual elements in their raw form in order to quantify them” (Ibrahim et al., 2020). In the last decade, boosted by the development of deep learning, computer vision has become an efficient way of modeling different aspects of cities. With regard to the built environment, the applications of computer vision can be categorized into two groups: seeing cities from above and from street level (Ibrahim et al., 2020). Satellite remote-sensing approaches have focused on night-time light for the estimation of in-use stock of metals and other materials in infrastructures and buildings (Rauch, 2009; Takahashi et al., 2009, 2010; Liang et al., 2017; Peled & Fishman, 2021). Seeing closer to the street level, however, has not yet been used to inform building stock research extensively, despite the quickly maturing literature developed as part of the efforts undertaken developing autonomous vehicles (Zhang et al., 2020). So far, its applications include the assessment of land-use (Srivastava et al., 2019), urban vegetation cover (Seiferling et al., 2017), or the detection of potholes on roads (Dhiman & Klette, 2019).

As the overall framework in Figure 1 illustrates, a mobile-sensing approach with computer vision and machine learning to construct 3D urban surface maps would allow us to identify and classify stock objects, components, and materials specific to individual buildings. The framework consists of an initial data collection stage using a confederation of sensors. A suite of machine learning and computer vision approaches are then used to both reconstruct the 3D geometry of the urban scene and detect stock components and materials. Finally, the 3D reconstruction and detected components/materials are fused to generate a semantically labeled urban model that enables quantification of the stock components and material at a building level.

Our intentions in this work are not to develop/advocate a set of specific computer vision models for the sole purpose of stock accounting. The crux of the proposed framework rests in the argument for the integration of the mobile-sensing and computer vision approaches and use of automation to increase the pace and spatial resolution of stock accounting at a building/component level. The data streams suggested in the figure are, thus, informed by the existing literature on what currently can and ultimately could be extracted from each sensory capture: LiDAR enables accurate measurements of size and geometry of buildings and their facade components (Ackere et al., 2019), visual imaging satisfies minimum requirements to detect and recognize the facade components (Dai et al., 2021), and thermal and hyperspectral imaging enable detection of components' material composition and wear condition (Ziolkowski et al., 2018; Yao et al., 2020; Cho et al., 2018).

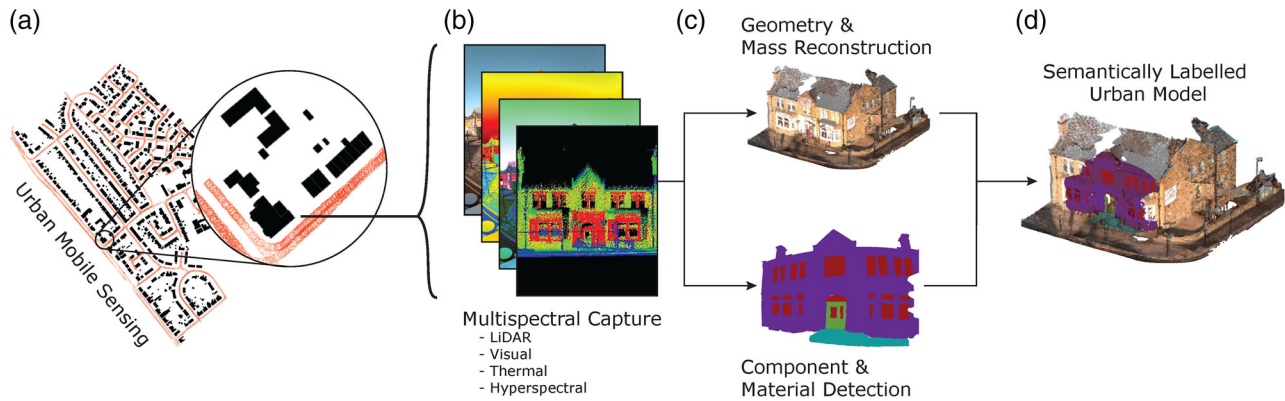


FIGURE 1 Framework schematic for a scalable spatially explicit and high-resolution urban stock characterization using multi-sensor mobile mapping (a, b) to enable reconstruction of building geometry and detection of components and materials (c, d) for a semantically labelled urban surface model (d)

2.2 | Data collection and management

The implementation of such a framework could experience some practical obstacles and/or policy implications. The main and immediate obstacle is ensuring a line-of-sight with buildings when using mobile-sensing approaches for data collection. For a complete implementation, the data captured would need to have full building coverage. Using only drive-by methods, for example, Google Street View or the case study in the next section, will result in quantifying only what the vehicle can observe from the street, leaving out the façade components that would generally be expected to be on the backside of the buildings or those on the front that are obstructed by vegetation or miscellaneous urban furniture. While this does not impact the validity of the information extracted for each building, a mixture of drive-by and fly-by imaging would be required to ensure full building envelope coverage to avoid systemic underestimation of components and stock. Drone-based fly-by building detection approaches are now fairly advanced in reconstructing building geometry from visual images (H. Huang et al., 2020). Remaining challenges in complementing drive-by methods with fly-by imaging are those relating to swarm dynamics. These would be crucial for scalability in the context of an automated deployment of a number of low-cost unmanned aerial vehicles to achieve the same pace and ease of en masse data collection as drive-by methods (Bouffanais, 2016).

A secondary group of implications would involve the collection, management, and governance of such integrated and automated mobile/remotesensing frameworks. In terms of data collection, the proposed framework remains scalable, deployable, and manageable in terms of human effort and time needed to undertake the mobile sensing and maintaining the stock model. This inherently suggests potential for a public body or the academic community collectively shouldering computing costs and the management of such urban resources under an opensource open data framework. We should, however, note the high capital costs of developing and deploying a custom imaging vehicle with all four sensory requirements which is currently required for the implementation of the framework through drive-by imaging might limit development opportunities for commercial ventures. Although we currently do not take a position on a preferred overall collection or governance in reference to the implementation of the framework, we do believe the academic community engaged in developing and undertaking urban building stock accounting should engage in a dialogue on developing harmonized approaches to collection and reporting as already advocated by Heeren and Fishman (2019) in the case of material intensity surveys.

2.3 | Generalizability and scalability

An issue to further expand on is the suitability and limits of computer vision and machine learning for achieving a scalable approach to material stock characterization. The potential of these methods for increasing the speed at which stock accounting can be performed is more easily demonstrable given the progress made so far in image recognition in other engineering and medical applications (Brynjolfsson & Mitchell, 2017). The larger unknown is the generalizability potential of existing methods in the context of material accounting.

Building façade segmentation dates back a few decades. The current cutting edge of both building façade and the broader urban scene segmentation studies relies mostly on deploying and expanding methods using convolutional neural networks of various architectures (Badrinarayanan et al., 2016; Femiani et al., 2018; Fu et al., 2019; Zhao et al., 2017; Zolanvari et al., 2018; Schmitz et al., 2019). The community has also developed a series of public and often-used test datasets including ECP (Teboul et al., 2010) and Graz (Riemenschneider et al., 2012) for façade segmentation and Cityscapes (Cordts et al., 2016), Mapillary Vistas (Neuhof et al., 2017), and ApolloScape (X. Huang et al., 2018) for urban scene segmentation. The existing datasets for building façade and urban scene detection each pose their own particular challenges when adapted for purposes of building

stocks accounting. In the context of scalability, these challenges concern the continued efficiency of models when applied to building components and geographic context outside the scope of the dataset. The building facade datasets often contain pre-edited images which affects the convenience of their use in training models that would be used with other mobile sensed images that will have a variety of viewing angles and conditions. The urban scene datasets, as mainly developed by the autonomous vehicle community, do not have this problem, but currently suffer from a lack of detailed labeling for components below a building level given their primary use case.

The immediate issues that need addressing for practical realization of the framework are those regarding the minimum size of the training sets needed to identify, with a reasonable accuracy, various building components within different national/geographic context and whether this differs between building components and across regions. Zhu et al. (2020) have recently worked on large-scale architectural asset detection in panoramic images across 17 different cities. Their work suggests that there are some essential qualities defining distinct building components. Although geographical proximity and architectural style, for example, windows frame structure, do provide for internal clustering of elements into subgroups with less internal variability. As such, as a starting point, existing façade detection models could be adapted in conjunction with visual imaging available from services similar to Google Street View. (See the online Supporting Information for a small demonstration on generalizability of façade components across geographic regions.) For practical uses, however, these models are likely to require some initial fine-tuning and retraining, not necessarily to compensate for differences in the building stocks used in original training, but for the embedded differences in the images due to variations in equipment, capture angles, etc.

2.4 | Integration with existing approaches

Given the methodological challenges set out in the previous sections, a scalable framework is still some ways off. On the path toward full implementation of such a framework, intermediary steps can be taken to make use of and maintain compatibility with the existing archetype-based methodological approaches. A high level of differentiation can be reached by using the archetypes as a base only and complementing it with as much information as possibly retrievable from computer vision. For example, most bottom-up approaches base their stock estimation entirely on archetypes. In fact, a hypothetical compromise between existing survey methods and the suggested framework could alternatively focus on training similar computer vision models to assign existing archetypes to buildings first. This could simply involve only the visual imaging of the building stock as already available through services such as the Google Street View. However, with the presented framework and utilizing models developed at component level, additional data will become available to refine results at building level, for example, areas of wall, windows, and doors, instead of being intrinsically estimated through an archetype-based approach as per Table 1.

Our framework can hence raise the precision of results while still keeping the data collection requirements feasible and scalable. Indeed, and most importantly, the main constraint in such a framework is the time required for the drive-by and/or fly-over sensing. This means that compared with existing bottom-up methods, the approach is more easily scalable across large urban areas. It is worth reiterating here that in this work, we are not advocating the use of a particular set of computer vision models or overall methodologies.

2.5 | Stock accounting for circular economy and global decarbonization

Arriving back at the broader potential of a scalable framework, the spatially explicit and component level understanding of material stocks is essential to deliver the circular economy. The circular economy is built on the idea of a continuous loop of materials, across multiple lifecycles, to reduce and eventually eliminate new resource extraction. Critical to the idea is that of maintaining materials at their highest value possible. In the case of the building stock, this would be maintaining materials together as a building in the first instance, meaning that building retrofit is a key part of the agenda. If buildings cannot be made fit for purpose, they should be deconstructed, and the components remanufactured or directly reused. Alternatively, they could be striped back into their component parts and the individual materials reused or more likely in this case recycled. The eventual full implementation of the framework enables derivation of façade materials quantity and quality. For example, the number of bricks that could be salvaged for reuse, or the mass of concrete available for downcycling. These building-specific observations can then be improved by linking in building archetypes to provide further information on what cannot be seen from the outside.

Addressing the larger decarbonization challenge involving buildings, a significant proportion of energy use in buildings is to maintain thermal comfort. Heat loss in older buildings is predominantly a function of the thermal performance of the building facade. This is driven by the material characteristics and method of construction. Our suggested framework would also facilitate a scalable identification of the stock most in need of retrofit. This is helpful to both predict future material demand and its carbon impacts, but also enables consideration of whole life carbon payback times for different insulation types (Li & Tingley, 2021; Moncaster et al., 2013) and could facilitate local authorities in bulk procuring retrofit interventions.

TABLE 1 Typical information required to calculate building stocks with a bottom-up approach

	Data source Component	Dimension			Material composition		
		Computer vision	Inventory	Archetype*	Computer vision	Inventory	Archetype
External walls	Wall	Vs		A	Vs	I+	A
	Window	Vs		A	Vs		A
	Door	Vs		A	Vs		A
Internal walls	Wall			A			A
	Door			A			A
Interm. floors	Nb of levels	Vs	I				A
	Intermediate floor	Va*	I*	A			A
	Flooring	Va*	I*	A			A
Basement	Floor	Va*	I*	A			A
	External walls			A			A
	Internal walls			A			A
	Doors			A			A
Foundations	Slab	Va*	I*	A			A
	Strip footing	Va*		A			A
	Reinforcement			A			A
Roof	Ceiling	Va*	I*	A			A
	Roof cover	Va+Vs		A	Vs, Va	I+	A
	Trusses			A			A
Chimney	Dimensions	Vs		A	Vs		A
Stairs	Type			A			A
	Slope			A			A

For each component, information is needed on its dimensions and material composition. For each, several data sources can be used, depending on availability. Abbreviations: I+, information only available if the inventory is detailed enough; Va, computer vision from above; Vs, computer vision from the street.

*The data source can give an indication, but further assumption might be required. This is the case, for example, with the area of a foundation slab: satellite data gives an indication of a building's footprint that may be used as a proxy for the slab area if no other information is available.

3 | PROOF OF CONCEPT: A PRELIMINARY CASE STUDY IN SHEFFIELD

In this section, we demonstrate the potential of the framework combining mobile sensing and computer vision in a small case study of a local neighborhood. This serves as a demonstration of how such a framework could be operationalized based on methodological components already available within the existing literature. The specifics of the neural network model used in the case study is of secondary importance in this paper.

The case study area is of roughly 2500 inhabitants spanning 2.79 km² in the southwest of Sheffield, UK. In prototyping a demonstrator of the framework, we use a bespoke mobile-sensing platform and an existing cluster of neural network models to calculate an estimate count of the buildings' external doors and windows. The mobile-sensing platform used, the multi-spectral advanced research vehicle, is a van mounted with a custom imaging rig, which enables collection of the four data streams required in the production of multi-spectral texturized 3D surface maps of the captured built environment (Meyers et al., 2019). For the case study presented here, we only make use of the visual imaging stream for demonstration purposes. (Interested researchers and groups should contact the corresponding author for data access.)

3.1 | Component and material detection for a semantically labeled urban model

In prototyping the component detection part of the framework, Figure 1c, we make use of an existing ensemble of convolutional neural networks trained to identify building components such as windows and doors at a pixel level (Dai et al., 2019), an example of which is shown in Figure 2. The incorporated ensemble of models is based on a U-Net architecture which segments input images, Figure 2a, for each component class separately, Figure 2b, before assembling the results together. The U-Net architecture was originally developed for processing and segmentation of medical images and uses a combination of high- and low-level image information to determine pixel-wise classification (Ronneberger et al., 2015), see Figure 2d.

We make use of the model developed by Dai et al. (2019) for this case study for convenience as it has originally been trained on output from the same research vehicle. The images used by Dai et al. are also captured in Sheffield, covering mostly terraced, semi-detached, and detached residential buildings, albeit in a different part of the city without any overlap with the stock used in the case study here. Dai et al. use an eight-category schema for labeling their training set comprising: background (non-building pixels), walls, roofs, chimneys, two variations of windows, and two variations of doors, Figure 2b. 240 images are then used for the training, validation, and testing of the model with a [80%, 5%, 15%] split of

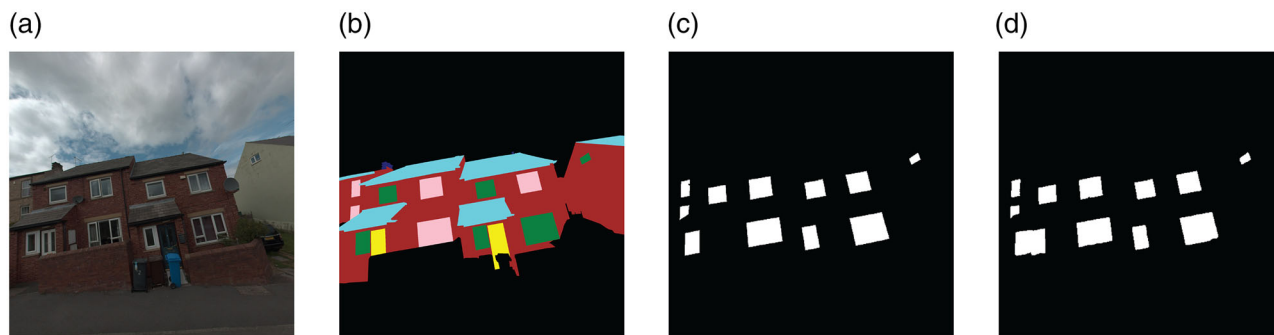


FIGURE 2 Sample of visual input to and output of the component segmentation model. (a) Raw image input, (b) sample annotation of façade component used for training, (c) sample window component truth mask isolated from the full annotation, and (d) sample model predicted window mask

images across the training, validation, and test set. Data augmentation including horizontal flip, vertical and horizontal shifts, and hue adjustment is also applied to the images used by the authors (Krizhevsky et al., 2017).

All models are implemented and trained using existing TensorFlow libraries (Abadi et al., 2016). The results are then pooled together assigning the most confident pixel-level classification from among the categories to each pixel. Here, we should note that we do not retrain the model by Dai et al. specifically based on the data from the neighborhood used in this case study and treat their model as a black-box ready-made package. Further information regarding the specifics of the implementation and training of the ensemble of neural network models used including their accuracy and precision can be found in Dai et al. (2019).

For the purposes of our case study demonstration, we only use the models to detect and present counts for doors and windows since these are more inherently *countable* relative to, say, walls and roofs and as such more instructive in a wider discussion of the framework's practical applications vis-à-vis circular economy and building refurbishments. Regarding broader questions on generalizability of the model used as raised in the previous sections, a toy example of the application of Dai et al. model to Google Street View images from a few different cities outside the United Kingdom is available in the Supporting Information.

3.2 | Other data and methods

The Ordnance Survey maintains and provides building footprint geometry for the United Kingdom (Ordnance Survey, 2020). These are obtained through a combination of aerial LiDAR and photographic imaging. The dataset, as such, includes building height as well as footprint geometry. Given that the focus of this paper is showcasing the potentials of a mobile and remote-sensing framework, we use this dataset here, as Ordnance Survey already provides the product. The same information, however, could have been obtained independently using the LiDAR capture from the mobile-sensing platform used for the visual imaging following existing methodologies (Vayghan et al., 2020; Wang et al., 2019).

In order to obtain estimates of the count of façade windows and doors for each building, we first extract the 100 nearest vehicle positions, to the building centroid, and hence images taken looking in the building's direction. The retrieved image positions and the distribution of average image distance from buildings can be seen in Figure 3a and its inset. Overall, we consider 42,451 unique photos for 1515 structural footprints within the neighborhood. (Raw images used are available from the corresponding author upon reasonable request.) The retrieved photos are then passed to the image segmentation models to extract the number of door and window components for each image. This results in distributions, each effectively based on a hundred sampled values of the count of doors and windows for each building. Since the field-of-view in each image is not limited to a single building, we scale the component count in each image by the number of buildings for which a given image has been retrieved and weigh them inversely with distance from the building such that

$$\bar{X}_n = \left[\frac{1}{100} \sum_i \frac{x_i}{a_i l_{i,n}} \right], \quad (1)$$

where \bar{X}_n is the number of components, doors, or windows, for building n with x_i the number of components estimated for image i . a_i is the number of buildings to which image i is amongst the 100 nearest and $l_{i,n}$ is the distance between image i and the centroid of building n . Note that the ceiling function avoids fractional component counts.

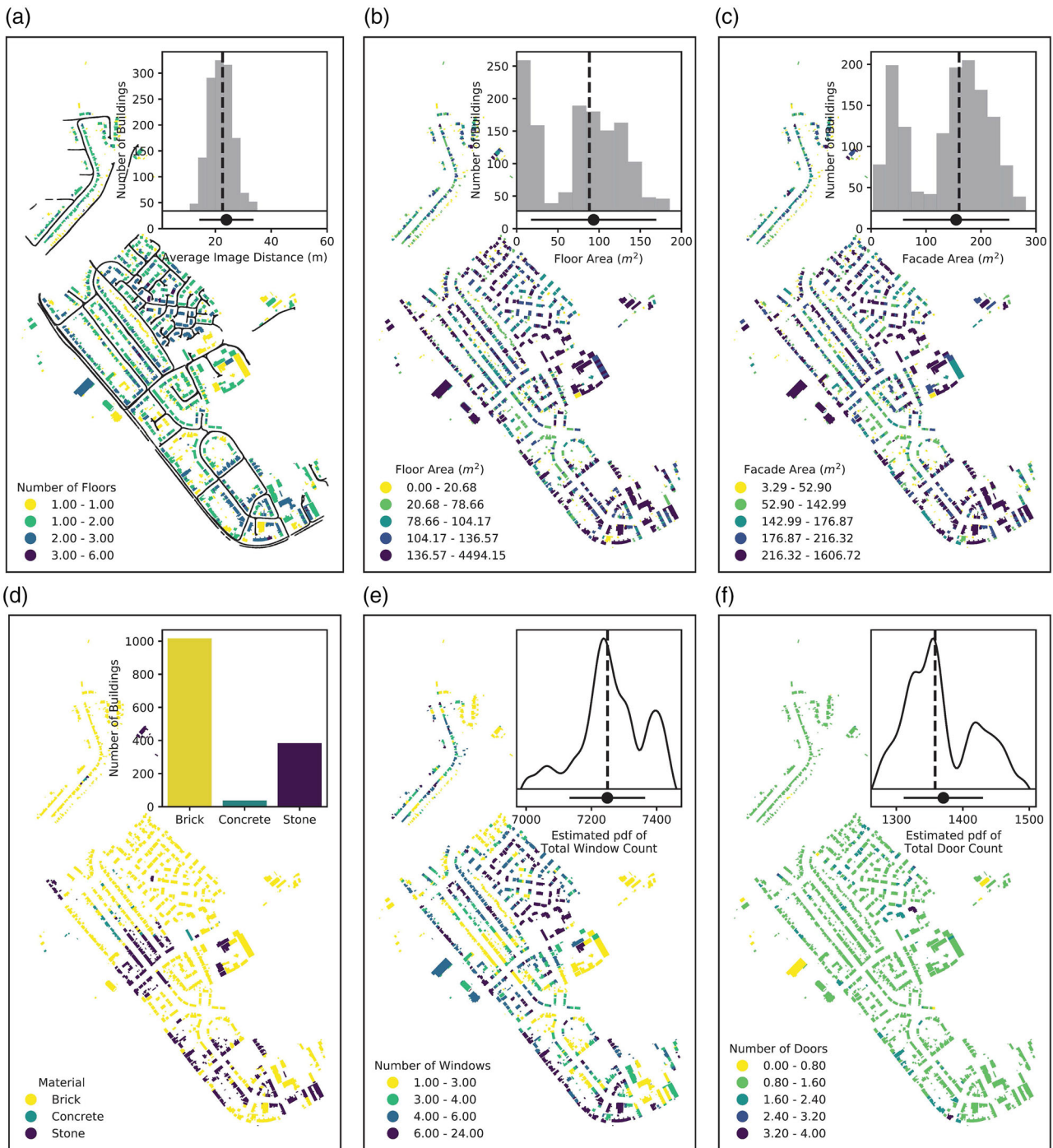


FIGURE 3 Spatial distribution of building characteristics and components. (a) Colormap of floor count along with the positions of the vehicle corresponding to the 42451 images. (b) Total floor area. (c) Façade area as estimated by height and footprint perimeter. (d) Predominant façade material. (e, f) Estimated numbers of windows and doors with insets showing the estimated probability density function of the expected value for the total neighborhood-wide count. For insets, dashed vertical lines denote the median with the point plots indicating mean and standard deviation. Note that colormap breaks are of different scales. Maps contain OS Map data © Crown copyright and database rights 2020. Data from this figure are available in File S2 of the Supporting Information

As the count of components are independent of one another, we estimate the mean and standard deviation of the total number of components across the neighborhood by estimating the sequential convolution of the buildings' distributions of the component count following

$$P(T_n = t) = \int_{-\infty}^{\infty} P(T_{n-1} = x) P(X_n = t - x) dx, \quad (2)$$

where, $P(T_n)$ is the distribution function of the sum of component count, either door or window, of n buildings, $P(T_{n-1})$ is the distribution function of the sum of component count of $n - 1$ buildings, and $P(X_n)$ is that of the component count for the n th building. We use Gaussian kernel estimates of the distributions of the doors and windows for each building in estimating the probability distribution of total number of components in the neighborhood.

3.3 | Building components and materials

Figure 3 shows the spatial distribution of the stock information extracted from the footprint, height, and mobile-sensing platform. The neighborhood is mostly made up of single or double story terraced and semi-detached houses with a handful of medium-rise apartment blocks, Figure 3a. Figures 3b and 3c show the distributions of building floor and façade areas. The bi-modal appearance of these is due to the inclusion of garages and temporary structures which comprise the peaks at the lower values. The rest of the buildings show an average floor and façade area of 100 and 150 m², respectively. Assuming an average wall thickness of 300 mm, the mean neighborhood floor area would be reduced to 90 m², which is in broad agreement with the national average *usable* floor area of 92 m² cited in the English Housing Survey for buildings of a similar age (Ministry of Housing, Communities & Local Government, 2018). Figure 3d provides a breakdown of the prominent façade material, see accuracy and sensitivity section for more details. Finally, Figures 3e and 3f and their insets show spatial distribution of the façade components and their total count across the neighborhood, respectively. We should clarify here that the parameters presented in Figures 3a–c have been extracted from existing building polygons maintained by the Ordnance Survey for convenience. Figures 3d shows material allocation based on the manual examination of the photos performed by two of the authors, see future work section for more information. Figures 3e,f visualizes component count estimates based on the steps described in Section 3.2.

There is, expectedly, a wider variation of component count for windows at a building level than doors. For the majority of buildings in the neighborhood, we detect a single door which is consistent with the number of doors visible on the façade of terraced and semi-detached houses to a street-level observer. (Note that in Figure 3f, the higher number of doors is often observed for corner structures where a better line-of-sight to buildings' back doors is available.) As for windows, the detected number of street-facing windows closely follows the building type, where rows of terraced and semi-detached buildings can be seen with two to five windows, respectively, with the higher window counts observed in clusters belonging to apartments and non-residential buildings.

3.4 | Accuracy and sensitivity

Here, we briefly address some aspects of the prototype's accuracy and sensitivity. We begin by outlining the work undertaken to quantify the accuracy of the component count.

As previously mentioned, the component counts rely on automated processing of 42,451 images. To provide a comparison, we have implemented a manual count of components within a subset of 1366 images. This subset comprises the closest image to each structure's footprint among the 100 originally queried for each polygon. For each image, two of the authors undertook a manual count of the total number of components visible, the individual building component count, and the predominant material of the façade. We should note that manual counters have assessed their images independently and hence there has not been an expectation that their count of components or assessment of the predominant construction material visually apparent should agree. We use the average of these manual counts as an alternative in situ survey assessment against which the prototyped framework can be benchmarked. (Readers are encouraged to refer to the online Supporting Information for the comparison of manual counts by the two authors.)

Figure 4 provides detailed distribution of the modeled and manual component count variations at an image level, with Table 2 showing the summary count and sum of the components for both methods at a building level. Note that neither the figure nor the table provides a quantification of the accuracy or precision of the neural network models used. Rather, they provide a measure of the extent to which the model estimates diverge from manual human counts based on a subset of the images. (Readers should consult Dai et al. (2019) for the accuracy and precision of the neural network models used.) This is important, especially with regard to the manual building-level estimates in Table 2, which are based on single-image information.

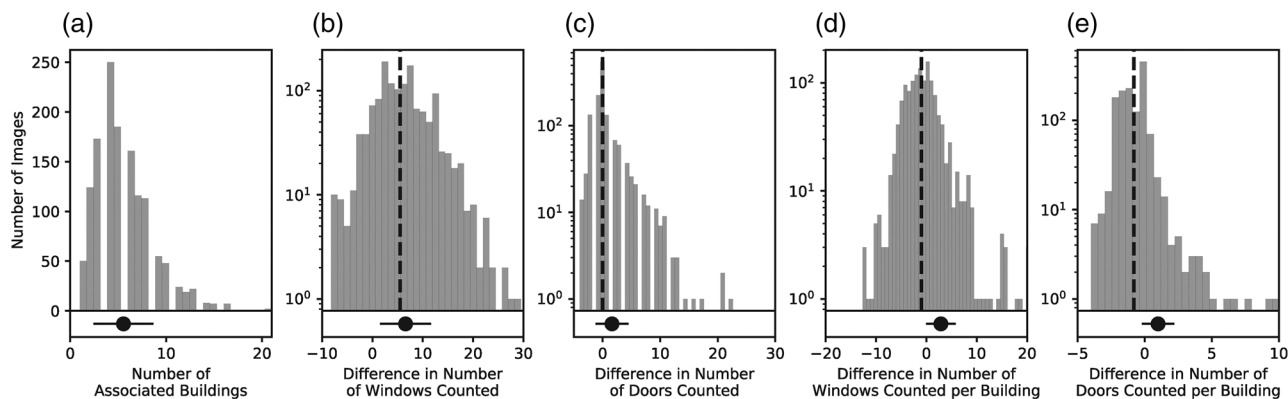


FIGURE 4 Distribution of image-level component counts by the model and its deviation from manual counts. Dashed vertical lines denote the median with the point plots indicating mean and standard deviation. Note that the axes have been truncated excluding single occurrence large values. Data from this figure are available in File S2 of the Supporting Information

TABLE 2 Summary benchmarking of the model estimated counts versus manually assigned counts

	Based on model count	Based on manual count
Number of buildings	1515	
Number of images used	42,451	1366
Mean window count per building (standard deviation)	4.65 (2.63)	4.38 (3.08)
Sum windows (standard deviation)	7248 (112)	6640
Mean door count per building (standard deviation)	1.06 (0.30)	1.09 (0.98)
Sum doors (standard deviation)	1358 (57)	1645

At an image level, the model appears to overestimate the total number of components by up to an average of one and three for per building doors and windows, respectively. This is partly due to the potential multiple counting of door and window components where constituting pixel are non-contiguous. However, the picture is slightly more nuanced, as despite the average values, the majority of the disagreements, for the exception of the total number of windows, involves an underestimation of the number of components, see median and distribution bins in Figure 4c–e. As such, the mean values are skewed by the few instances in which the model has substantially overestimated component counts. (See underlying data in File S2 of the Supporting Information provided for the values not shown in the figure.)

At a building level, the mean difference in the counts for both doors and windows suggest that, despite model simplicity, the model estimated and directly assigned counts broadly agree on an aggregated level, Table 2. The difference between the total door count estimates might appear significantly high at 20%. However, it is worth noting that such high variations, inevitable for relatively low total number of components, still only represent fractional differences in mean component count for an *average* building—compare average 1.06 doors from the model versus 1.09 doors from the manual counts. This is important, since potentially large deviations are currently implicit within a large body of material stocks studies where the use of archetypal values implies large uncertainties in total component count, especially when considering increasingly larger geographic boundaries (Lanau et al., 2019).

3.5 | Future work

In this section, both limitations of this case study and planned future work to further implement the scalable framework are outlined. We have mentioned the major challenges in truly implementing a scalable accounting framework in Section 2. The case study here using street-level imaging showcases a number of these difficulties including underestimation of overall building components due to an incomplete view of buildings.

In practical terms, one of the options for optimal circular economy is building refurbishments. This would reduce the demand for new buildings, reducing embodied impacts. However, there is a risk that if energy efficiency is not a priority, building life extension could maintain operational inefficiencies making carbon emissions from use more challenging to reduce. In building refurbishments terms, the case study presented would only be able to give a crude estimate of likely insulation and does not consider the nuances of different wall types, further work is therefore required to investigate the use of the mobile-sensing platform for the recognition of wall construction from their visual, thermal, and hyperspectral signatures.

This would facilitate estimation of the surface areas that require either external or internal insulation, compared to cavity wall properties which may have already been insulated. Future work also needs to make more targeted use of thermal and hyperspectral imaging to understand heat loss patterns through the building fabric for a quantification of component material type and quality. (See Phan (2012) for an example of a prior attempt at incorporating thermal imaging.) Finally, parallel work is under way at the time of writing this article, focusing on classifying buildings into archetypes for the United Kingdom, something that is currently lacking in a specialized sense for the UK building stock. This will be valuable, and we aim to integrate these archetypes within the framework presented here as set out in Section 2.4.

4 | CONCLUSIONS

In this paper we have demonstrated the application of a framework based on mobile sensing and machine learning to automate the estimation of building stock components and materials. This represents an advance on existing bottom-up accounting of building stocks that rely entirely on archetypes. Indeed, the semantic reconstruction of the built environment in 3D would offer multiple benefits. First, accounting exercises based on such a framework provide building-specific registers of components. At its bare minimum, mobile sensing provides high-fidelity individualized measurement of buildings. Given access to relevant archetypes for an area, the framework can also provide a spatially explicit archetype matching method that provides the same level of information as archetypal average characteristic estimates. Fully functioning, however, the framework enables building-level detection of various facade components and their material, for example, windows and wall, and an understanding of both what is available to *urban mine* in the future based on building-level measurements, as well as indicating likely future material demands for building refurbishments. Additionally, the combined building-specific information on components, their dimensions, and constituting material allow for a much more accurate and spatially resolute quantification of material stock which is currently partially achieved through remote sensed information.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Hadi Arbabi  <https://orcid.org/0000-0001-8518-9022>

Maud Lanau  <https://orcid.org/0000-0001-6315-2608>

Gregory Meyers  <https://orcid.org/0000-0003-4157-3991>

REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ..., & Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning* (pp. 265–283). USENIX Association. <https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi>
- Ackere, S. V., Verbeurgt, J., Sloover, L. D., Wulf, A. D., de Weghe, N. V., & Maeyer, P. D. (2019). Extracting dimensions and locations of doors, windows, and door thresholds out of mobile lidar data using object detection to estimate the impact of floods. *Geoinformation for Disaster Management*, 8, 429–436.
- Adams, K. T., Osmani, M., Thorpe, T., & Thornback, J. (2017). Circular economy in construction: Current awareness, challenges and enablers. *Proceedings of the Institution of Civil Engineers - Waste and Resource Management*, 170(1), 15–24. <https://doi.org/10.1680/jwarm.16.00011>
- Arbabi, H., Punzo, G., Meyers, G., Tan, L. M., Li, Q., Tingley, D. D., & Mayfield, M. (2020). On the use of random graphs in analysing resource utilization in urban systems. *Royal Society Open Science*, 7(4), 200087. <https://doi.org/10.1098/rsos.200087>
- Arora, M., Raspall, F., Cheah, L., & Silva, A. (2019). Residential building material stocks and component-level circularity: The case of Singapore. *Journal of Cleaner Production*, 216, 239–248. <https://doi.org/10.1016/j.jclepro.2019.01.199>
- Augiseau, V. (2017). *La dimension matérielle de l'urbanisation: Flux et stocks de matériaux de construction en Ile-de-France* [These de doctorat, Paris 1]. <http://www.theses.fr/2017PA01H111>
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2016). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv:1511.00561 [Cs]*. <http://arxiv.org/abs/1511.00561>
- Bondi, A. B. (2000). Characteristics of scalability and their impact on performance. In M. Woodside, H. Gomaa, D. Menasce (Eds.), *Proceedings of the 2nd international workshop on software and performance* (pp. 195–203). <https://doi.org/10.1145/350391.350432>
- Bouffanais, R. (2016). Outlook: Can swarms be designed? In R. Bouffanais (Ed.), *Design and control of swarm dynamics* (pp. 105–106). Springer. https://doi.org/10.1007/978-981-287-751-2_7
- Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, 358(6370), 1530–1534. <https://doi.org/10.1126/science.aap8062>
- Cho, Y., Bianchi-Berthouze, N., Marquardt, N., & Julier, S. J. (2018). Deep thermal imaging: Proximate material type recognition in the wild through deep learning of spatial surface temperature patterns. In R. Mandryk, M. Hancock, M. Perry, A. Cox (Eds.), *Proceedings of the 2018 CHI conference on human factors in computing systems* (1–13). <https://doi.org/10.1145/3173574.3173576>
- Committee on Climate Change. (2019). *Net zero: The UK's contribution to stopping global warming*. Committee on Climate Change. <https://www.theccc.org.uk/wp-content/uploads/2019/05/Net-Zero-The-UKs-contribution-to-stopping-global-warming.pdf>
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016). *The cityscapes dataset for semantic urban scene understanding*. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Cordts_The_Cityscapes_Dataset_CVPR_2016_paper.html

- Dai, M., Meyers, G., Tingley, D. D., & Mayfield, M. (2019). Initial investigations into using an ensemble of deep neural networks for building façade image semantic segmentation. *Remote Sensing Technologies and Applications in Urban Environments IV*, 11157, 1115708. <https://doi.org/10.1117/12.2532828>
- Dai, M., Ward, W. O. C., Meyers, G., Tingley, D. D., & Mayfield, M. (2021). Residential building facade segmentation in the urban environment. *Building and Environment*, 199, 107921. <https://doi.org/10.1016/j.buildenv.2021.107921>
- Dhiman, A., & Klette, R. (2019). Pothole detection using computer vision and learning. *IEEE Transactions on Intelligent Transportation Systems*, 21, 3536–3550. <https://doi.org/10.1109/TITS.2019.2931297>
- Edwards, J., & Townsend, A. (2011). *Buildings under refurbishment and retrofit* (Carbon Action 2050 White Papers). Chartered Institute of Building. <https://www.ciob.org/sites/default/files/Buildings%20under%20Refurb%20and%20Retrofit.pdf>
- Femiani, J., Para, W. R., Mitra, N., & Wonka, P. (2018). Facade segmentation in the wild. *arXiv:1805.08634 [Cs]*. <http://arxiv.org/abs/1805.08634>
- Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., & Lu, H. (2019). Dual attention network for scene segmentation. *arXiv:1809.02983 [Cs]*. <http://arxiv.org/abs/1809.02983>
- Guldmann, E., & Huulgaard, R. D. (2020). Barriers to circular business model innovation: A multiple-case study. *Journal of Cleaner Production*, 243, 118160. <https://doi.org/10.1016/j.jclepro.2019.118160>
- Guo, Z., Hu, D., Zhang, F., Huang, G., & Xiao, Q. (2014). An integrated material metabolism model for stocks of urban road system in Beijing, China. *Science of the Total Environment*, 470–471, 883–894. <https://doi.org/10.1016/j.scitotenv.2013.10.041>
- Haberl, H., Wiedenhofer, D., Erb, K.-H., Görg, C., & Krausmann, F. (2017). The material stock–flow–service nexus: A new approach for tackling the decoupling conundrum. *Sustainability*, 9(7), 1049. <https://doi.org/10.3390/su9071049>
- Heeren, N., & Fishman, T. (2019). A database seed for a community-driven material intensity research platform. *Scientific Data*, 6(1), 23. <https://doi.org/10.1038/s41597-019-0021-x>
- Huang, H., Micheline, M., Schmitz, M., Roth, L., & Mayer, H. (2020). LOD3 building reconstruction from multi-source images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2020*, 427–434. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-427-2020>
- Huang, X., Cheng, X., Geng, Q., Cao, B., Zhou, D., Wang, P., Lin, Y., & Yang, R. (2018). *The ApolloScape Dataset for autonomous driving*. https://openaccess.thecvf.com/content_cvpr_2018_workshops/w14/html/Huang_The_ApolloScape_Dataset_CVPR_2018_paper.html
- Hyman, M., Turner, B., & Carpintero, A. (2013). *Guidelines for national waste management strategies: Moving from challenges to opportunities*. United Nations Environment Programme. http://cwm.unitar.org/national-profiles/publications/cw/wm/UNEP_UNITAR_NWMS_English.pdf
- Ibrahim, M. R., Haworth, J., & Cheng, T. (2020). Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities*, 96, 102481. <https://doi.org/10.1016/j.cities.2019.102481>
- Joglekar, P., & Woodside, M. (2000). Evaluating the scalability of distributed systems. *IEEE Transactions on Parallel and Distributed Systems*, 11(6), 589–603. <https://doi.org/10.1109/71.862209>
- Kleemann, F., Lederer, J., Rechberger, H., & Fellner, J. (2017). GIS-based analysis of Vienna's material stock in buildings. *Journal of Industrial Ecology*, 21(2), 368–380. <https://doi.org/10.1111/jiec.12446>
- Krausmann, F., Wiedenhofer, D., Lauk, C., Haas, W., Tanikawa, H., Fishman, T., Miatto, A., Schandl, H., & Haberl, H. (2017). Global socioeconomic material stocks rise 23-fold over the 20th century and require half of annual resource use. *Proceedings of the National Academy of Sciences*, 114(8), 1880–1885. <https://doi.org/10.1073/pnas.1613773114>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Krook, J., Carlsson, A., Eklund, M., Frändegård, P., & Svensson, N. (2011). Urban mining: Hibernating copper stocks in local power grids. *Journal of Cleaner Production*, 19(9–10), 1052–1056. <https://doi.org/10.1016/j.jclepro.2011.01.015>
- Lanau, M., & Liu, G. (2020). Developing an urban resource cadaster for circular economy: A case of Odense, Denmark. *Environmental Science & Technology*, 54(7), 4675–4685. <https://doi.org/10.1021/acs.est.9b07749>
- Lanau, M., Liu, G., Kral, U., Wiedenhofer, D., Keijzer, E., Yu, C., & Ehlert, C. (2019). Taking stock of built environment stock studies: Progress and prospects. *Environmental Science & Technology*, 53(15), 8499–8515. <https://doi.org/10.1021/acs.est.8b06652>
- Lederer, J., Kleemann, F., Ossberger, M., Rechberger, H., & Fellner, J. (2016). Prospecting and exploring anthropogenic resource deposits: The case study of Vienna's subway network. *Journal of Industrial Ecology*, 20(6), 1320–1333. f9pxhk
- Li, X., & Tingley, D. D. (2021). Solid wall insulation of the Victorian house stock in England: A whole life carbon perspective. *Building and Environment*, 191, 107595. <https://doi.org/10.1016/j.buildenv.2021.107595>
- Liang, H., Dong, L., Tanikawa, H., Zhang, N., Gao, Z., & Luo, X. (2017). Feasibility of a new-generation nighttime light data for estimating in-use steel stock of buildings and civil engineering infrastructures. *Resources, Conservation and Recycling*, 123, 11–23. /gbkq8v
- Mao, R., Bao, Y., Huang, Z., Liu, Q., & Liu, G. (2020). High-resolution mapping of the urban built environment stocks in Beijing. *Environmental Science & Technology*, 54(9), 5345–5355. <https://doi.org/10.1021/acs.est.9b07229>
- Mastrucci, A., Marvuglia, A., Popovici, E., Leopold, U., & Benetto, E. (2017). Geospatial characterization of building material stocks for the life cycle assessment of end-of-life scenarios at the urban scale. *Resources, Conservation and Recycling*, 123, 54–66. gbmbnf
- Mesta, C., Kahhat, R., & Santa-Cruz, S. (2019). Geospatial characterization of material stock in the residential sector of a Latin-American City. *Journal of Industrial Ecology*, 23(1), 280–291. <https://doi.org/10.1111/jiec.12723>
- Meyers, G., Zhu, C., Mayfield, M., Tingley, D. D., Willmott, J., & Coca, D. (2019). Designing a vehicle mounted high resolution multi-spectral 3D scanner: Concept design. In P. Pannuto, S. Pan, F. Salim, M. Kjærgaard (Eds.), *Proceedings of the 2nd workshop on data acquisition to analysis*, 16–21. <https://doi.org/10.1145/3359427.3361921>
- Miatto, A., Schandl, H., & Tanikawa, H. (2017). How important are realistic building lifespan assumptions for material stock and demolition waste accounts? *Resources, Conservation and Recycling*, 122, 143–154. 10/gbg3fk
- Ministry of Housing, Communities & Local Government. (2018). *Floor space in English homes* (English Housing Survey, p. 43). HM Government. <https://www.gov.uk/government/publications/floor-space-in-english-homes>
- Moncaster, A., Soulti, E., Mubarak, G., & Symons, K. (2013). Retrofitting solid wall buildings: Energy and carbon costs and savings. In K. Höfler, P. Maydl, & A. Passer (Eds.), *Proceedings of SB13 graz: Construction products and technologies* (pp. 95–101). Verlag der Technischen Universität Graz. <http://oro.open.ac.uk/52843/>

- Neuhold, G., Ollmann, T., Rota Bulo, S., & Kontschieder, P. (2017). *The Mapillary Vistas Dataset for semantic understanding of street scenes*. https://openaccess.thecvf.com/content_iccv_2017/html/Neuhold_The_Mapillary_Vistas_ICCV_2017_paper.html
- Ordnance Survey. (2020). *OS MasterMap topography layer*. Ordnance Survey. <https://www.ordnancesurvey.co.uk/business-government/products/mastermap-topography>
- Ortlepp, R., Gruhler, K., & Schiller, G. (2016). Material stocks in Germany's non-domestic buildings: A new quantification method. *Building Research & Information*, 44(8), 840–862. doi.org/10.1080/09590283.2016.1191444
- Ortlepp, R., Gruhler, K., & Schiller, G. (2018). Materials in Germany's domestic building stock: Calculation model and uncertainties. *Building Research & Information*, 46(2), 164–178. doi.org/10.1080/09590283.2017.1381444
- Paris Agreement. (2015). *United Nations treaty collection*. https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=en
- Peled, Y., & Fishman, T. (2021). Estimation and mapping of the material stocks of buildings of Europe: A novel nighttime lights-based approach. *Resources, Conservation and Recycling*, 169, 105509. <https://doi.org/10.1016/j.resconrec.2021.105509>
- Phan, L. N. (2012). *Automated rapid thermal imaging systems technology* (Thesis). Massachusetts Institute of Technology. <http://dspace.mit.edu/handle/1721.1/75664>
- Rauch, J. N. (2009). Global mapping of Al, Cu, Fe, and Zn in-use stocks and in-ground resources. *Proceedings of the National Academy of Sciences*, 106(45), 18920–18925. <https://doi.org/10.1073/pnas.0900658106>
- Recalde, K., Wang, J., & Graedel, T. E. (2008). Aluminium in-use stocks in the state of Connecticut. *Resources, Conservation and Recycling*, 52(11), 1271–1282. <https://doi.org/10.1016/j.resconrec.2008.07.006>
- Riemenschneider, H., Krispel, U., Thaller, W., Donoser, M., Havemann, S., Fellner, D., & Bischof, H. (2012). Irregular lattices for complex shape grammar facade parsing. *2012 IEEE conference on computer vision and pattern recognition* (pp. 1640–1647). <https://doi.org/10.1109/CVPR.2012.6247857>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28
- Schmitz, M., Huang, H., & Mayer, H. (2019). Comparison of training strategies for convnets on multiple similar datasets for facade segmentation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2-W13, 111–117. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-111-2019>
- Seiferling, I., Naik, N., Ratti, C., & Proulx, R. (2017). Green streets – Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, 165, 93–101. <https://doi.org/10.1016/j.landurbplan.2017.05.010>
- Srivastava, S., Vargas-Muñoz, J. E., & Tuia, D. (2019). Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution. *Remote Sensing of Environment*, 228, 129–143. <https://doi.org/10.1016/j.rse.2019.04.014>
- Stephan, A., & Athanassiadis, A. (2017). Quantifying and mapping embodied environmental requirements of urban building stocks. *Building and Environment*, 114, 187–202. doi.org/10.1016/j.buildenv.2017.09.055
- Symmes, R., Fishman, T., Telesford, J. N., Singh, S. J., Tan, S., & De Kroon, K. (2020). The weight of islands: Leveraging Grenada's material stocks to adapt to climate change. *Journal of Industrial Ecology*, 24(2), 369–382. <https://doi.org/10.1111/jiec.12853>
- Takahashi, K. I., Terakado, R., Nakamura, J., Adachi, Y., Elvidge, C. D., & Matsuno, Y. (2010). In-use stock analysis using satellite nighttime light observation data. *Resources, Conservation and Recycling*, 55(2), 196–200. <https://doi.org/10.1016/j.resconrec.2010.09.008>
- Takahashi, K. I., Terakado, R., Nakamura, J., Daigo, I., Matsuno, Y., & Adachi, Y. (2009). In-use stock of copper analysis using satellite nighttime light observation data. *Materials Transactions*, 50(7), 1871–1874. <https://doi.org/10.2320/matertrans.M2009030>
- Tanikawa, H., Fishman, T., Okuoka, K., & Sugimoto, K. (2015). The weight of society over time and space: A comprehensive account of the construction material stock of Japan, 1945–2010. *Journal of Industrial Ecology*, 19(5), 778–791. doi.org/10.1111/jiec.12262
- Tanikawa, H., & Hashimoto, S. (2009). Urban stock over time: Spatial material stock analysis using 4D-GIS. *Building Research & Information*, 37(5–6), 483–502. <https://doi.org/10.1080/09613210903169394>
- Tazi, N., Idir, R., & Fraj, A. B. (2021). Towards achieving circularity in residential building materials: Potential stock, locks and opportunities. *Journal of Cleaner Production*, 281, 124489. <https://doi.org/10.1016/j.jclepro.2020.124489>
- Teboul, O., Simon, L., Koutsourakis, P., & Paragios, N. (2010). Segmentation of building facades using procedural shape priors. *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 3105–3112). <https://doi.org/10.1109/CVPR.2010.5540068>
- Vayghan, S. S., Salmani, M., Ghasemkhani, N., Pradhan, B., & Alamri, A. (2020). Artificial intelligence techniques in extracting building and tree footprints using aerial imagery and LiDAR data. *Geocarto International*. <https://doi.org/10.1080/10106049.2020.1844311>
- Wallsten, B., Carlsson, A., Frändegård, P., Krook, J., & Svanström, S. (2013). To prospect an urban mine – Assessing the metal recovery potential of infrastructure “cold spots” in Norrköping, Sweden. *Journal of Cleaner Production*, 55, 103–111. <https://doi.org/10.1016/j.jclepro.2012.05.041>
- Wang, C., Ji, M., Wang, J., Wen, W., Li, T., & Sun, Y. (2019). An improved DBSCAN Method for LiDAR data segmentation with automatic Eps estimation. *Sensors*, 19(1), 172. <https://doi.org/10.3390/s19010172>
- Weinstock, C. B., & Goodenough, J. B. (2006). *On system scalability*. Carnegie-Mellon Univ Pittsburgh PA Software Engineering Inst. <https://apps.dtic.mil/sti/citations/ADA457003>
- World Economic Forum & Boston Consulting Group. (2019). *The Net-Zero challenge: Global climate action at a crossroads* [Briefing Paper]. World Economic Forum. http://www3.weforum.org/docs/WEF_The_Net_Zero_Challenge_Part1.pdf
- Yao, W., Lian, C., & Bruzzone, L. (2020). ClusterCNN: Clustering-based feature learning for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*. <https://doi.org/10.1109/LGRS.2020.3010837>
- Zhang, M., Fu, R., Guo, Y., & Wang, L. (2020). Moving object classification using 3D point cloud in urban traffic environment. *Journal of Advanced Transportation*, 2020, 1583129. <https://doi.org/10.1155/2020/1583129>
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. *arXiv:1612.01105 [Cs]*. <http://arxiv.org/abs/1612.01105>
- Zhu, P., Para, W. R., Fruehstueck, A., Femiani, J., & Wonka, P. (2020). Large scale architectural asset extraction from panoramic imagery. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/TVCG.2020.3010694>
- Ziolkowski, P., Niedostatkiewicz, M., & Szulwic, J. (2018). Deterioration of masonry structure investigated by a thermal camera with a qualitative assessment of thermal image processing. *International Multidisciplinary Scientific GeoConference*, 18(2,3), 97–104. <https://doi.org/10.5593/sgem2018/2.3/S10.013>

Zolanvari, S. M. I., Laefer, D. F., & Natanzi, A. S. (2018). Three-dimensional building façade segmentation and opening area detection from point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 143, 134–149. <https://doi.org/10.1016/j.isprsjprs.2018.04.004>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Arbabi H, Lanau M, Li X, Meyers G, Dai M, Mayfield M, Tingley DD. A scalable data collection, characterization, and accounting framework for urban material stocks. *J Ind Ecol*. 2022;26:58–71. <https://doi.org/10.1111/jiec.13198>