VU VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

Advancing large-scale analysis of human settlements and their dynamics

Li, Mengmeng

2023

DOI (link to publisher)
10.5463/thesis.9

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Li, M. (2023). *Advancing large-scale analysis of human settlements and their dynamics*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam]. Ipskamp Printing. https://doi.org/10.5463/thesis.9

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

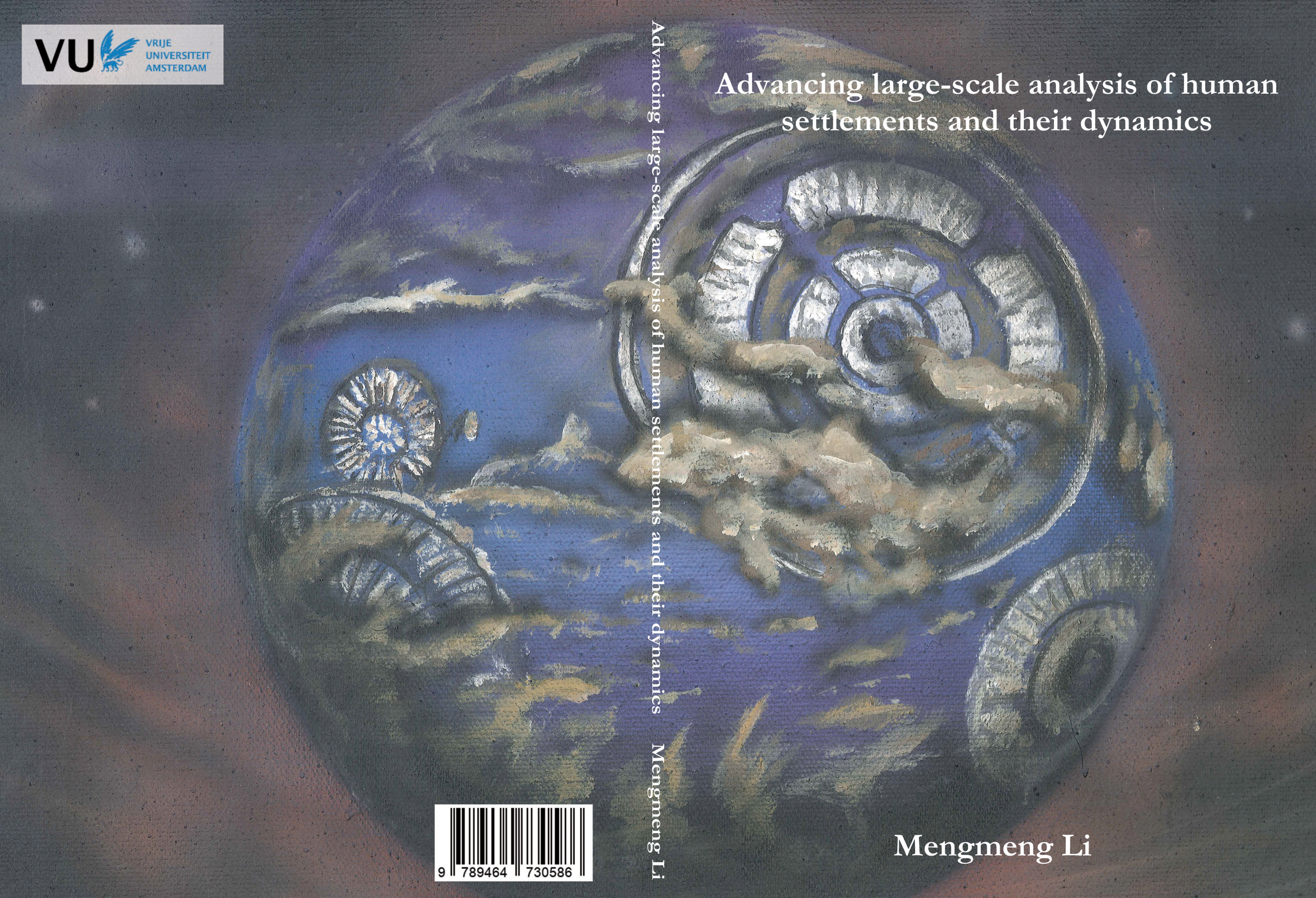
E-mail address:
vuresearchportal.ub@vu.nl

Download date: 03. Mar. 2023

VU VRIJE UNIVERSITEIT AMSTERDAM

Advancing large-scale analysis of human settlements and their dynamics

Advancing large-scale analysis of human settlements and their dynamics Mengmeng Li



Mengmeng Li

9 789464 730586

ADVANCING LARGE-SCALE ANALYSIS OF HUMAN SETTLEMENTS AND THEIR DYNAMICS

Mengmeng Li

Advancing large-scale analysis of human settlements and their dynamics

PhD thesis, Vrije Universiteit Amsterdam, The Netherlands
ISBN: 978-94-6473-058-6
DOI: 10.5463/thesis.9

© Mengmeng Li, Amsterdam, March 2023

Printed by:
Ipskamp Printing

Cover story:

This cover image depicts the prototype of Ecumenopolis, a term coined in 1967 by Constantinos Doxiadis, the father of *ekistics*, to represent his vision that global human settlements would expand and merge in the future, and there would eventually be one single giant city as a progression from urbanization and population growth, among others. (Painted by Mingkui Hu with inspiration from Stellaris)

The candidate was supported by the China Scholarship Council (Grant No. 201706510011). The research for this thesis was carried out primarily at the Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam.

VRIJE UNIVERSITEIT

ADVANCING LARGE-SCALE ANALYSIS OF HUMAN SETTLEMENTS AND THEIR DYNAMICS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philiosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op woensdag 15 maart 2023 om 9.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door
Mengmeng Li
geboren te Henan, China

promotoren:	prof.dr.ir. P.H. Verburg dr.ir. J. van Vliet
promotiecommissie:	prof.dr. P.J. Ward dr. E. Koomen prof.dr. Y. Ban prof.dr. K. Pfeffer prof.dr. C. He

To my family

Contents

Summary

Literally, human settlements are *the place* where *people* live. Human settlements are often divided into *urban settlements* and *rural settlements*, although there is no consensus on how to exactly differentiate the two. The share of the world population living in cities has increased rapidly in recent years, surpassing the 50% threshold only recently. Cities are expected to accommodate 2.5 billion additional people by 2050, leading to an urbanization rate reaching 68%. Meanwhile, micro-urbanization and land abandonment are also happening worldwide in small settlements. Cities play a vital role in global economic development and they are hubs for innovation. Consequently, cities are also the place where human resources and capital concentrate, further amplifying their importance. Yet, as a consequence of the concentration of people and their activities, human settlements also take up an increasing amount of land. As such, human settlements play an increasingly important role in the global competition for land.

Due to the importance for a range of sustainability challenges, it is important to understand the spatial dynamics of human settlements. The rapid expansion of built-up land is among the most extensive global land changes, even though built-up land occupies only a small fraction of the terrestrial biosphere. Moreover, the different ways in which human settlements are manifested are crucially important for their environmental and socioeconomic impacts. For example, the enormous concentration of population in urban slums can be interpreted as highly intensive land use, thus minimizing the amount of land used for residential activities. Yet, it also represents low-quality living conditions, which suggest unsustainable socioeconomic outcomes. Urban sprawl, in contrast, requires a relatively large amount of land per person, thus threatening food production and biodiversity habitat, while the quality of life is generally much higher therein. Yet, current analysis of human settlements heavily relies on land cover datasets, which typically have only one single class to represent human settlements, i.e., built-up land, impervious surface, or urban land. Consequently, the analysis of human settlements does often not account for the heterogeneity within urban environment or their subtle changes over time. This simplistic representation severely limits our understanding of change processes in human settlements, as well as our capacity to assess socioeconomic and environmental impacts.

Land systems, which represent the outcome of human interactions with natural environment, are widely acknowledged as the key to overcoming the simplicity of

land cover-based analysis. Therefore, mapping land systems could underpin the representation of nature-environment interactions in the integrated social-ecological systems. However, existing classification schemes of land systems are predominantly applied to characterize natural and/or agricultural landscapes. Although a few studies have incorporated multiple classes representing human settlements, these did not include urban land systems. Yet, this approach offers ample opportunities to explore the heterogeneity within human settlements.

This thesis aims to advance large-scale analysis of human settlements and their dynamics through the lens of land systems, with a specific focus on the role of land-use intensity. To achieve this objective, this dissertation addresses the following three research questions (RQs):

RQ1: How can we represent the diversity of settlement systems?

RQ2: What can we learn from Earth Observation to better characterize urban land-use intensity?

RQ3: How do urban land-use intensity and settlement systems change over time?

Chapter 2 explores the use of human settlement systems as an approach to understanding their variation in space and changes over time. To do so, three properties derived from built-up land (the density of built-up land, the density of built-up clusters, and the size of the largest cluster) are used to characterize settlement systems in China. Results show that settlement systems exist along a density gradient, and their change trajectories are typically gradual and incremental. Moreover, this chapter shows that the vast majority of all built-up land is included in a number of village systems, despite a focus of many studies on large urban areas. In addition, results indicate that the total increase in built-up land in village landscapes outweighs that of dense urban regions. This chapter suggests that we should characterize human settlements more comprehensively to advance the analysis of human settlements, going beyond the emergence of new built-up land in a few mega-cities only.

In **Chapter 3**, urban land-use intensity is operationalized by the horizontal and vertical spatial patterns of buildings. Particularly, I trained three random forest models to estimate building footprint, height, and volume, respectively, at a 1-km resolution for Europe, the US, and China. Random forest models were trained based on reference data from a few available 3D city models. Compared to 3D building information for individual buildings, which is crucial for urban analytics, the gridded 3D building datasets exist at a coarser resolution which is directly supportive of

representing urban land-use conditions at larger scales. The random forest models are fed with optical remote sensing imagery, SAR imagery, remote sensing derived indices, and other spatial data. Subsequently, these trained models are used to estimate the same variables for all the other locations in Europe, China, and the US. The models yield R^2 values of 0.90, 0.81, and 0.88 for building footprint, height, and volume, respectively, for all three continents combined. The correlation between building footprint and building height at a pixel level was 0.66, illustrating the relevance of mapping these properties independently.

Chapter 4 builds on the methodological approach presented in chapter 3. Specifically, it presents an improved approach to mapping 3D built-up patterns (i.e., 3D building structure), and applies this to map building footprint, height, and volume at a global scale. The methodological improvement includes an optimized model structure, additional explanatory variables, and updated input data. The latter includes reference data from all world continents. Subsequently, I analysed 3D built-up patterns along the urban–rural gradient in selected cities, as well as in more rural areas across ten world regions, to illustrate the variation in urban morphology. These results show a large variation in the built-up volume per person across major world regions, with the volume in South Asia almost an order of magnitude smaller than the volume in Oceania. Moreover, I find distance decay functions from the centre of the city to its outskirts for all three properties for major cities in all continents. Yet, again, the height, footprint (density), and volume differ drastically across these cities. The generated data offers new opportunities for the assessments of urban environmental impacts at a global scale.

Chapter 5 uses built-up land per person as an operationalization for urban land-use intensity, in order to investigate its temporal dynamics at a global scale. Results suggest that the decrease of urban land-use intensity relates to 38.3%, 49.6%, and 37.5% of the built-up land expansion in the three periods during 1975-2015 (i.e., 1975-1990, 1990-2000, and 2000-2015), but with large local variations. Chapter 5 also finds that centers of large cities intensify in all three periods, while their rural counterparts show an opposite direction. In the Global South, densification often happens in regions where human settlements are already used intensively, suggesting potential trade-offs with other living standards.

These chapters represent the recent advancements in large-scale analysis of human settlements by revealing a large variation in urban fabric. Urban densification is widely acknowledged as one of the tangible solutions to satisfy the increased land demand for human settlement while conserving other land, suggesting the relevance of these findings to inform sustainable development. Nevertheless, local settlement

trajectories towards intensive forms should also be guided in a large-scale context with broad considerations, including the quality of life for inhabitants, because these trade-offs and synergies remain largely unexplored in this analysis. I argue that the representation of urban land-use change should incorporate such variation in integrated assessment models to inform decision-making.

Nederlandstalige samenvatting

Letterlijk zijn menselijke nederzettingen de plaats waar mensen wonen. Menselijke nederzettingen worden vaak verdeeld in stedelijke nederzettingen en landelijke nederzettingen, hoewel er geen consensus bestaat over hoe de twee precies kunnen worden onderscheiden. Het aandeel van de wereldbevolking dat in steden woont, is de afgelopen jaren snel toegenomen en heeft pas onlangs de drempel van 50% overschreden. Steden zullen naar verwachting tegen 2050 2,5 miljard extra mensen huisvesten, wat leidt tot een verstedelijkingspercentage van 68%. Ondertussen vinden ook wereldwijd micro-urbanisatie en landverlating plaats in kleine nederzettingen. Steden spelen een vitale rol in de wereldwijde economische ontwikkeling en zijn hubs voor innovatie. Bijgevolg zijn steden ook de plaats waar menselijke hulpbronnen en kapitaal zich concentreren, waardoor hun belang nog groter wordt. Maar als gevolg van de concentratie van mensen en hun activiteiten nemen ook menselijke nederzettingen steeds meer land in beslag. Als zodanig spelen menselijke nederzettingen een steeds belangrijkere rol in de wereldwijde concurrentie om land.

Vanwege het belang voor een reeks duurzaamheidsuitdagingen, is het belangrijk om de ruimtelijke dynamiek van menselijke nederzettingen te begrijpen. De snelle uitbreiding van bebouwd land is een van de meest uitgebreide wereldwijde landveranderingen, ook al beslaat bebouwd land slechts een klein deel van de terrestrische biosfeer. Bovendien zijn de verschillende manieren waarop menselijke nederzettingen zich manifesteren van cruciaal belang voor hun milieu- en sociaaleconomische effecten. Zo kan de enorme concentratie van de bevolking in stedelijke sloppenwijken worden geïnterpreteerd als zeer intensief landgebruik, waardoor de hoeveelheid land die wordt gebruikt voor woonactiviteiten wordt geminimaliseerd. Toch vertegenwoordigt het ook levensomstandigheden van lage kwaliteit, die wijzen op onhoudbare sociaaleconomische resultaten. Stedelijke wildgroei vereist daarentegen een relatief grote hoeveelheid land per persoon, waardoor de voedselproductie en de biodiversiteit in de leefomgeving worden bedreigd, terwijl de kwaliteit van leven daar over het algemeen veel hoger is. Toch leunt de huidige analyse van menselijke nederzettingen sterk op datasets over landbedekking, die doorgaans slechts één enkele klasse hebben om menselijke nederzettingen weer te geven, d.w.z. bebouwd land, ondoordringbaar oppervlak of stedelijk land. Bijgevolg houdt de analyse van menselijke nederzettingen vaak geen rekening met de heterogeniteit binnen de stedelijke omgeving of hun subtiele veranderingen in de loop van de tijd. Deze simplistische weergave beperkt ons begrip

van veranderingsprocessen in menselijke nederzettingen ernstig, evenals ons vermogen om sociaaleconomische en milieueffecten te beoordelen.

Landsystemen, die het resultaat zijn van menselijke interacties met de natuurlijke omgeving, worden algemeen erkend als de sleutel tot het overwinnen van de eenvoud van op landbedekking gebaseerde analyses. Daarom zou het in kaart brengen van landsystemen de representatie van natuur-milieu-interacties in de geïntegreerde sociaal-ecologische systemen kunnen ondersteunen. Bestaande classificatieschema's van landsystemen worden echter voornamelijk toegepast om natuurlijke en/of agrarische landschappen te karakteriseren. Hoewel in enkele studies meerdere klassen zijn opgenomen die menselijke nederzettingen vertegenwoordigen, omvatten deze geen stedelijke landsystemen. Toch biedt deze benadering volop mogelijkheden om de heterogeniteit binnen menselijke nederzettingen te onderzoeken.

Dit proefschrift beoogt grootschalige analyse van menselijke nederzettingen en hun dynamiek te bevorderen door de lens van landsystemen, met een specifieke focus op de rol van landgebruiksintensiteit. Om dit doel te bereiken, behandelt dit proefschrift de volgende drie onderzoeksvragen:

OV1: Hoe kunnen we de diversiteit van afwikkelingssystemen weergeven?

OV2: Wat kunnen we leren van aardobservatie om de intensiteit van stedelijk landgebruik beter te karakteriseren?

OV3: Hoe veranderen de intensiteit van stedelijk landgebruik en vestigingssystemen in de loop van de tijd?

Hoofdstuk 2 onderzoekt het gebruik van menselijke vestigingssystemen als een benadering om hun variatie in de ruimte en veranderingen in de tijd te begrijpen. Om dit te doen, worden drie eigenschappen die zijn afgeleid van bebouwd land (de dichtheid van bebouwd land, de dichtheid van bebouwde clusters en de grootte van het grootste cluster) gebruikt om vestigingssystemen in China te karakteriseren. De resultaten tonen aan dat vestigingssystemen langs een dichtheidsgradiënt bestaan en dat hun veranderingstrajecten typisch geleidelijk en incrementeel zijn. Bovendien laat dit hoofdstuk zien dat het overgrote deel van alle bebouwde grond deel uitmaakt van een aantal dorpssystemen, ondanks de focus van veel studies op grote stedelijke gebieden. Bovendien geven de resultaten aan dat de totale toename van bebouwd land in dorpslandschappen groter is dan die van dichtbevolkte stedelijke gebieden. Dit hoofdstuk suggereert dat we menselijke nederzettingen uitgebreider moeten karakteriseren om de analyse van menselijke nederzettingen vooruit te helpen, waarbij

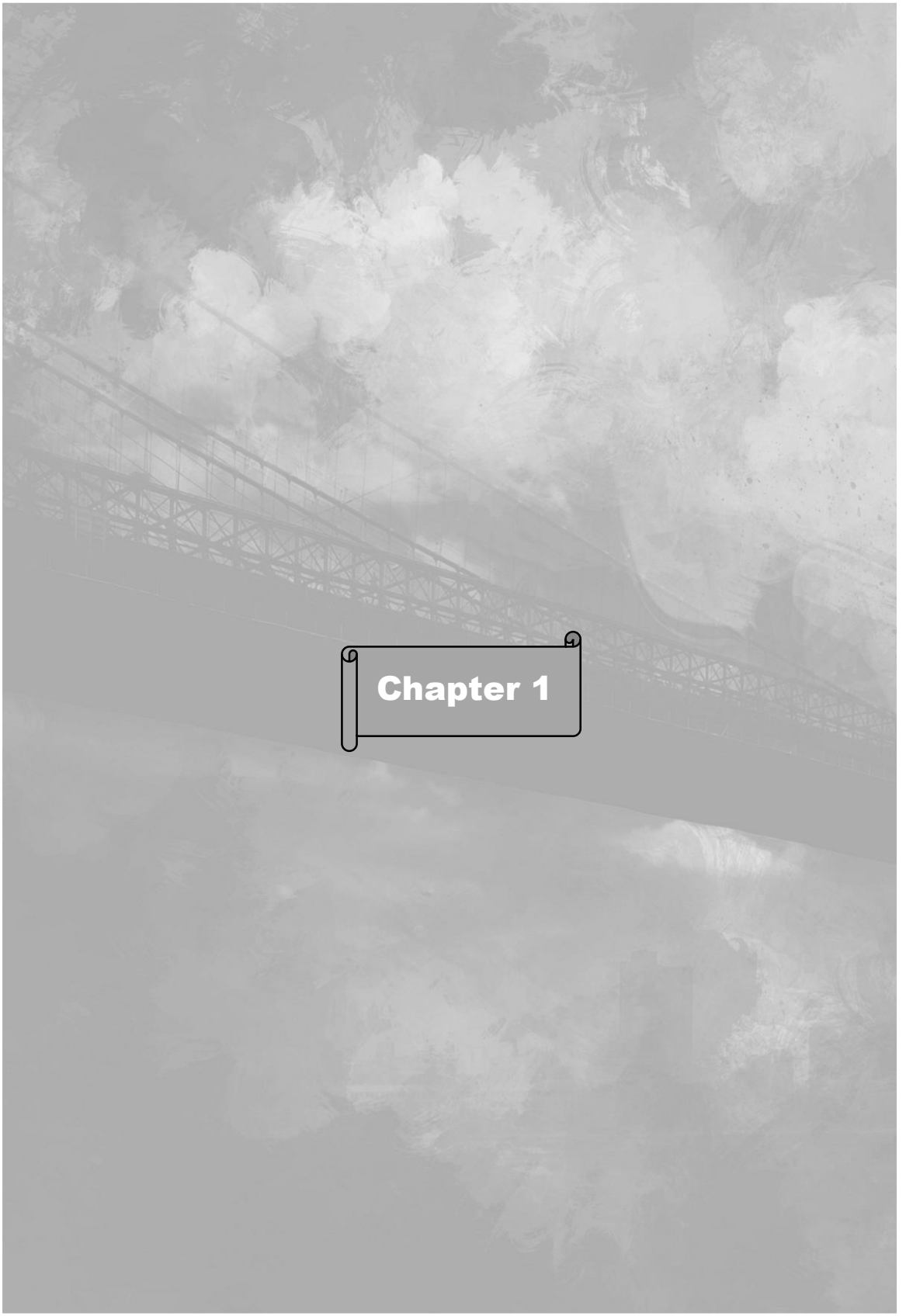
we verder gaan dan de opkomst van nieuw bebouwd land in slechts enkele megasteden.

In **hoofdstuk 3** wordt de stedelijke landgebruiksintensiteit geoperationaliseerd door de horizontale en verticale ruimtelijke patronen van gebouwen. Ik heb met name drie willekeurige bosmodellen getraind om respectievelijk de voetafdruk, hoogte en het volume van gebouwen te schatten met een resolutie van 1 km voor Europa, de VS en China. Willekeurige bosmodellen werden getraind op basis van referentiegegevens van enkele beschikbare 3D-stadsmodellen. Vergeleken met 3D-bouwinformatie voor individuele gebouwen, wat cruciaal is voor stedelijke analyses, bestaan de gerasterde 3D-bouwdatasets met een grovere resolutie die direct ondersteuning biedt voor het weergeven van stedelijke landgebruiksomstandigheden op grotere schaal. De willekeurige bosmodellen worden gevoed met optische teledetectiebeelden, SAR-beelden, van teledetectie afgeleide indices en andere ruimtelijke gegevens. Vervolgens worden deze getrainde modellen gebruikt om dezelfde variabelen te schatten voor alle andere locaties in Europa, China en de VS. De modellen leveren R2-waarden op van 0,90, 0,81 en 0,88 voor respectievelijk de voetafdruk, de hoogte en het volume van het gebouw voor alle drie de continenten samen. De correlatie tussen de voetafdruk van het gebouw en de hoogte van het gebouw op pixelniveau was 0,66, wat de relevantie illustreert van het onafhankelijk in kaart brengen van deze eigenschappen.

Hoofdstuk 4 bouwt voort op de methodologische benadering die in hoofdstuk 3 wordt gepresenteerd. Het presenteert met name een verbeterde benadering voor het in kaart brengen van 3D-bebouwde patronen (d.w.z. 3D-bouwstructuur), en past deze toe om de voetafdruk, hoogte en het volume van gebouwen op wereldschaal in kaart te brengen. De methodologische verbetering omvat een geoptimaliseerde modelstructuur, aanvullende verklarende variabelen en bijgewerkte invoergegevens. Deze laatste omvat referentiegegevens van alle wereldcontinenten. Vervolgens analyseerde ik 3D-bebouwde patronen langs de stad-landelijke gradiënt in geselecteerde steden, evenals in meer landelijke gebieden in tien wereldregio's, om de variatie in stedelijke morfologie te illustreren. Deze resultaten laten een grote variatie zien in het opgebouwde volume per persoon in de belangrijkste wereldregio's, waarbij het volume in Zuid-Azië bijna een orde van grootte kleiner is dan het volume in Oceanië. Bovendien vind ik afstandsvervalfuncties van het centrum van de stad naar de buitenwijken voor alle drie de eigendommen voor grote steden in alle continenten. Maar nogmaals, de hoogte, voetafdruk (dichtheid) en volume verschillen drastisch tussen deze steden. De gegenereerde gegevens bieden nieuwe mogelijkheden voor de beoordeling van stedelijke milieueffecten op wereldschaal.

Hoofdstuk 5 gebruikt bebouwd land per persoon als operationalisering voor de intensiteit van stedelijk landgebruik, om de temporele dynamiek ervan op wereldschaal te onderzoeken. De resultaten suggereren dat de afname van de intensiteit van stedelijk landgebruik betrekking heeft op 38,3%, 49,6% en 37,5% van de bebouwde landuitbreiding in de drie perioden van 1975-2015 (d.w.z. 1975-1990, 1990-2000 en 2000-2015), maar met grote lokale variaties. Hoofdstuk 5 constateert ook dat centra van grote steden in alle drie de perioden intensiveren, terwijl hun tegenhangers op het platteland een tegengestelde richting laten zien. In het Globale Zuiden vindt verdichting vaak plaats in regio's waar menselijke nederzettingen al intensief worden gebruikt, wat wijst op mogelijke afwegingen met andere levensstandaarden.

Deze hoofdstukken vertegenwoordigen de recente vorderingen in grootschalige analyse van menselijke nederzettingen door een grote variatie in stedelijk weefsel aan het licht te brengen. Stedelijke verdichting wordt algemeen erkend als een van de tastbare oplossingen om te voldoen aan de toegenomen vraag naar land voor menselijke vestiging en tegelijkertijd ander land te behouden, wat de relevantie van deze bevindingen suggereert voor duurzame ontwikkeling. Desalniettemin moeten lokale vestigingstrajecten naar intensieve vormen ook worden begeleid in een grootschalige context met brede overwegingen, inclusief de levenskwaliteit van de bewoners, omdat deze afwegingen en synergieën grotendeels onontgonnen blijven in deze analyse. Ik betoog dat de representatie van verandering in stedelijk landgebruik dergelijke variatie in geïntegreerde beoordelingsmodellen moet opnemen om de besluitvorming te informeren.


Chapter 1

1 | Introduction

1.1 A brief description of human settlements

Human settlements are literally *the place* where *people* live. To this end, people may construct shelter and associated infrastructure, and most importantly, acquire extra territory for resource exploitation to sustain their livelihoods (Ellis et al., 2013), which altogether prompt the archaeological debate on the causality of human settlements that emerged to facilitate anthropogenic activities such as hunter-gathering and farming (Cumming et al., 2014; Weissbrod et al., 2017). In other words, human settlements are the entity that integrates and assimilates human beings living there, physical elements such as shelter and infrastructure, as well as functions and services to which their physical components could provide support. Accordingly, human geographers acknowledge that human settlements can be further distinguished by their population size, spatial pattern, dominant production activity, regional centrality and functionality, among others.

Since the Industrial Revolution, the development of human settlements, particularly the urbanization process featured by rural-urban migration, has profoundly reshaped the Earth's environment as well as our socioeconomic system (Eitelberg et al., 2016; Seto and Ramankutty, 2016). Global assessments suggest that more than 75% of Earth's terrestrial surface has been modified by human activities (Ellis and Ramankutty, 2008), despite of a relatively small fraction of territory directly used for human residency (Liu et al., 2014c). In comparison, agricultural production has long been acknowledged as one of the most prominent human alterations of our planet (Foley et al., 2011; Lambin and Meyfroidt, 2011). In recent decades, however, the expansion speed of human settlements is approaching the magnitude of agricultural expansion (van Vliet, 2019), which underpins a growing demand for in-depth analysis of human settlement trajectories, in order to better consolidate existing human settlements and develop new land elsewhere for human settlements.

Given that human settlements are often surrounded by valuable land such as fertile cropland and natural habitat, the projected expansion of human settlements in the near future due to population increase and rural-urban migration challenges our pathways toward global sustainability (Forman and Wu, 2016; Liu and Li, 2017). For instance, van Vliet et al. (2017) projected that global built-up land will double between 2000 and 2040, leading to a spatial displacement of up to 35 Mha cropland to fulfill the demand for food. With the incorporation of different trajectories of human settlements, however, Wang et al. (2021a) find that rural-urban migration in China

could release a large amount of cropland from rural built-up land between 2015 and 2050. Therefore, crop production in 2050 would be higher than that in 2015 even considering the relatively low productivity in reclaimed cropland elsewhere. These inconsistent or even contradictory findings imply that steering human settlements toward an appropriate trajectory could be the leverage point for addressing food insecurity.

Often, human settlements are dichotomously divided into *urban settlements* and *rural settlements*, there is no consensus yet on how to exactly differentiate the two. In reality, human settlements exist along a continuum of land uses that ranges from deep rural to dense urban (Seto et al., 2012b; van Vliet et al., 2020), and awareness of this gradient is essential in local practices for global and regional sustainability. So far, cities play a vital role in global development of human settlements as they serve as the engine for economic growth and are the hubs for innovation. Cities are expected to house 2.5 billion additional people by 2050, with an urbanization rate reaching 68% (United Nations, 2018). On the one hand, cities are the place where human resources and capital concentrate and are therefore viewed as the horsepower for development (Liu and Li, 2017). On the other hand, countryside communities in many parts of the world are suffering from land abandonment and/or land use dis-intensification (van der Zanden et al., 2016).

In the context of globalization, especially the flow of population and commodity-driven agricultural expansion in a tele-coupled world, human settlements play an increasingly central role in planetary sustainability. Sustainable Development Goal 11 (SDG11), which explicitly targets a sustainable future of human settlements, requires systematic solutions to challenges emerging with the development of human settlements. In particular, we need to consider how the use of our territory can be arranged to support sustainable development.

1.2 Representation of human settlements in land science

Early documentations on human settlement studies, as many other geographic studies in the same period, are rather theoretical and descriptive. The classic theories on human settlements developed in the past centuries, which serve as backbones for human settlement studies nowadays, are dominated by theoretical descriptions of their spatial pattern and/or interaction with human activities. For example, von Thünen (1826) laid the foundation of land use theory by introducing a model to explain the spatial interaction of human activities that determines the location for

facilities, in a sense that public goods and services are optimally accessed to the population. Over one hundred years later, Christaller (1933) built a mathematical geometry model to explain the distribution of market services in a commercial system, or human settlements in a residential system. Later on, Fujita and Krugman (1995) developed a monopolistic competition model and examined conditions under which all manufactured goods will be produced in a single town. The distance-decay function along urban-rural gradients or between paired cities for one or more spatial interactions is integral to these models (Van der Horst, 2006).

Since the wide application of remote sensing techniques after 1980s, land cover has prevailed in large-scale analyses of human settlements (Shaw et al., 2020). Land cover, representing biophysical attributes of the Earth's surface (Turner et al., 1994), often illustrates how a specific region is spatially covered by physical land types such as forests and wetlands, among others. Land cover can be derived from satellite and/or aerial imagery, while land use that reflects human purpose applied to land cover cannot be directly interpreted (Lambin et al., 2001). Land cover in the context of human settlements mostly refers to built-up land, impervious surface, or artificial land, assuming it reflects and/or relates to the place where most human residential activities take place. Binary land cover products that differentiate human settlements from the others stem from the classification of remote sensing images, of which pixels are typically determined by their dominant land cover (Bontemps et al., 2013; Schneider et al., 2009; van Vliet et al., 2019). For the past decades, researchers have attempted to quantify how human settlements has developed over time (Weng, 2012). Yet, their approaches are mostly confined to the analysis of locations where new human settlements have taken place, and/or where they are likely to emerge in the future.

Land use is associated with the economic and cultural activities through which people interact with land and terrestrial ecosystems (Meyfroidt et al., 2018), and it is the accumulation of spatial arrangements, activities, and inputs that people undertake in a certain land cover type (Watson et al., 2000). Land use is an advanced representation of land cover by incorporating subtle discrepancies in human activities. For example, land used for aesthetics in urban forestry parks but is less accessible to people and primary forest designated as protected areas are both forest-covered in appearance, but they have different functions to people and to nature. Land use in the context of human settlements relates to the biophysical land cover (typically, built-up land or impervious surface) used for human activities such as residency, business, and recreation, which can be further elaborated, for example by its spatial pattern, density, and mix.

More recently, the focus of land use studies shifted from traditional land use/land cover (LULC) to land system science. According to Verburg et al. (2013), land systems "*represent the terrestrial component of the Earth system and encompass all processes and activities related to the human use of land, including socioeconomic, technological and organizational investments and arrangements, as well as the benefits gained from land and the unintended social and ecological outcomes of societal activities*". Beyond the concept of mosaic landscapes in land classification (Messerli et al., 2009), land system science also addresses the nuanced differences with which a specific land type is used. For instance, with the recognition of human alternation on our planet, Ellis and Ramankutty (2008) identified eighteen *anthropogenic biomes* (also known as "*anthromes*") through a synthetic analysis of population, land use, and land cover. Moreover, van Asselen and Verburg (2012) identified thirty land systems globally by synthesizing land cover, livestock, and agricultural intensity data with a hierarchical classification method. Similarly, Václavík et al. (2013) identified twelve land system archetypes based on a wide range of environmental and socioeconomic indicators. At a regional scale, Malek and Verburg (2017) mapped land systems in the Mediterranean region by integrating data on land cover and socioeconomics, which eventually outlines multifunctional agro-silvo-pastoral mosaic landscapes that host various human alterations that are normally not well represented by land cover products. Similarly, Ornetsmüller et al. (2018) mapped land systems for Laos by synthesizing multisource datasets, which enables the delineation of land uses such as shifting cultivation that is otherwise underrepresented in land cover products. These land system representations underpin the nature-environment interactions in the integrated social-ecological systems (McGinnis and Ostrom, 2014). However, these classification schemes of land systems are predominantly designated to characterize natural and/or agricultural landscapes, despite of many classes representing human settlements, in particular, the *urban* and *dense settlements* as well as a number of *village* classes in Ellis and Ramankutty (2008), which are used to analyse changes in natural and agricultural land, but they do not serve well to investigate the development of human settlements themselves.

A nuanced representation of human settlements in land system science could help address global sustainability challenges. On the one hand, a considerable share of the global sustainability crisis can be attributed to inappropriate uses of land, among which human settlements are one of the most intensive forms of land used by human (Foley et al., 2005). On the other hand, the manifestation of human settlements that across individual sites, neighbourhoods, blocks, urban areas, and regions is the prerequisite for developing practical solutions for sustainability crisis (Seto and Pandey, 2019), as there is an increasing consensus that there are no one-size-fits-all solutions for the challenges such as greenhouse gas emission and food insecurity, but

instead to apply appropriate and specific mitigating measures for different types of social groups that characterize the distinction of human settlements (Baiocchi et al., 2015; Hemerijckx et al., 2022).

Given the solution-oriented nature of land system science (Verburg et al., 2013), there is a growing demand for a better representation of human settlements in order to formulate and implement practical solutions for global sustainability crisis. For example, Tierolf et al. (2021) recently developed a land system classification for human settlements in the Southeast Asia to explore future preferable development pathways (i.e., expansion or intensification) in the context of exposure to river flood risk, and further analyses suggest that preferable options are context dependent. Specifically, for some countries, flood risk was largest in the expansion scenario, while for other countries in the Southeast Asia, flood risk was largest in the intensification scenario (Tierolf et al., 2021). These different outcomes of human settlement trajectories cannot be differentiated without the nuanced land system classification for human settlements.

1.3 Characterizing human settlements with Earth Observation data

1.3.1 Global studies

The development of global human settlement products from Earth Observation and ancillary data is more or less temporally synchronized with the progress in other thematic products such as cropland (Ramankutty et al., 2008), forest (Hansen et al., 2013), and water bodies (Pekel et al., 2016), which all adheres to the development of remote sensing platforms. Table. 1.1 gives an overview of global studies in mapping human settlement, where MODIS, Landsat, and DMSP/OLS are mostly utilized sources of satellite imagery. In general, human settlement mapping studies are developing towards higher frequency, higher resolution, and higher accuracy, particularly, particularly after the release of public computational platforms such as Google Earth Engine (Gorelick et al., 2017).

Human settlements in these remote sensing products are primarily represented by their land cover, e.g., built-up land, or impervious surfaces. The characterization of land cover in these products can be dichotomously divided into human settlements and non-human settlements. As a consequence, many change analyses are limited to the mere conversion of non-human settlements to human settlements.

Table 1.1. Representative studies illustrating the development of mapping human settlements at a global scale based on Earth Observation techniques. An extended summary of studies can be found in, for example, Potere et al. (2009), Schneider et al. (2010), and Sabo et al. (2018). Note that the selected studies below are conducted exclusively for mapping human settlements with two exceptions: Bartholomé and Belward (2007) and Gong et al. (2013), in which human settlements are depicted as *cities* and *impervious surfaces,* respectively, among many other LULC classes.

Publication	Platform/data source	Resolution	Period
Elvidge et al. (2007)	DMSP/OLS NTL	1000 m	2000
Bartholomé and Belward (2007)	SPOT	1000 m	2000
Schneider et al. (2009)	MODIS	500 m	~2002
Schneider et al. (2010)	MODIS	500 m	~2002
Gong et al. (2013)	Landsat	30 m	~2005
Zhou et al. (2015)	DMSP/OLS NTL	1000 m	2000
Esch et al. (2017)	TanDEM-X and TerraSAR-X	~12 m	~2012
Liu et al. (2018)	Landsat	30 m	1990-2010
Florczyk et al. (2019a)	Landsat	30 m	1975-2015
Marconcini et al. (2020)	Landsat, and Sentinel-1 SAR	10 m	2015

1.3.2 Recent advances in regional and global studies

Only recently, a small number of studies have characterized human settlements beyond the presence and/or absence of built-up areas, including both numerical and categorical representations. Here, the numerical aspect primarily refers to the quantitative methods developed to characterize the 3D urban morphology. While the categorical aspect is associated with the description of urban environment by categorical classification (e.g., climate-relevant classes, and urban land-use types).

Whilst monitoring urban extent from local to global scales has been prominent in Earth Observation studies, there is an increasing need to characterize urban development beyond two-dimensional spatial patterns to comprehensively assess urban sustainability. Earth Observation researchers recently characterized urban environment by the vertical dimension of buildings. For example, Frantz et al. (2021) mapped building heights across Germany by combining dual-polarized Sentinel-1A/B and multispectral Sentinel-2A/B. Cao and Huang (2021) estimated building height based on multi-view imagery in 42 Chinese cities. More recently, Esch et al. (2022) present the World Settlement Footprint 3D product that characterizes urban 3D morphology within all human settlements.

Large-scale mapping of Local Climate Zones (LCZs) and urban land-use types became possible only a few years ago. LCZ is a classification scheme that accounts for vegetation and the vertical configuration of buildings has been widely adopted by urban climate scientists to investigate the association between urban morphology and thermal environment (Stewart and Oke, 2012). In recent years, the LCZ typology is also used to explicitly illustrate how human settlements grow vertically. For example, Chen et al. (2020a) mapped the LCZs from 1985 to 2018 in Denmark, and results highlight the divergent trajectories of urban development between Copenhagen and Aarhus of Denmark. As of the year 2022, there are two globally available datasets on LCZs, produced by Zhu et al. (2022) and Demuzere et al. (2022), respectively. The former dataset covers all cities across the globe with a population greater than 300 thousand (Zhu et al., 2022), while the latter dataset illustrates not only human settlements, but also natural and agricultural landscapes that are far away from human settlements (Demuzere et al., 2022). Both datasets facilitate urban morphology analysis at a global scale. In addition, mapping land-use types is recently implemented at large scales. For example, Gong et al. (2020a) produced the first map of land-use types in major Chinese cities, which provides opportunities to analyze the spatial pattern and configuration of different land-use types. By fusing Earth Observation data and socioeconomic data, Rosier et al. (2022) mapped urban land use in the Netherlands with high overall accuracy.

In some cases, the heterogeneity of human settlements is represented in a broader context. For example, Stokes and Seto (2019) characterized the divergence of urbanization process by comparing urban infrastructural transitions in India and the US, concluding that the mere analysis of horizontal built-up land expansion could under-represent important nuances such as rural electrification and informal settlement growth. By employing spectral-temporal metrics and regression-based un-mixing, Schug et al. (2020) explored how to map the fractions of built-up land and vegetation at nation-wide scales, and their new product in which human settlements are represented by continuous built-up land fractions allows for better characterization of human settlements and identification of subtle change trajectories (e.g., densification). The concept of mapping land cover fractions coincides with earlier fragmented studies applied at various scales (Dai et al., 2010; Huang et al., 2016; Michishita et al., 2012; Wetherley et al., 2017).

The heterogeneity of human settlements also involves many other aspects, for example, land size and land tenure (Cao et al., 2020), which may not be accessible merely from remote sensing imaginaries. Emerging information acquisition techniques such as social sensing that based on individual-centered data may provide unprecedented tools for addressing these issues (Liu et al., 2015). However, as human

settlements act as the linkage of physical, ecological, and socioeconomic systems across spatial scales, Zhu et al. (2019) point out that efforts to separate them into various components or categories will always be limited, and absolute delineations of purely social or biophysical are not possible. The other key but earlier review on urban remote sensing by Weng (2012) highlights the application of hyperspectral imaging in mapping building material in urban environment.

1.4 Knowledge gap

Whilst the dynamics of human settlements have been investigated from local to global scales, two key issues are particularly under-presented, which may affect the conclusions drawn from these studies. Specifically, the analysis of human settlements heavily relies on land cover datasets, which typically have only one single class representing human settlements, and the analysis of human settlements, therefore, does often not account for heterogeneity within the urban environment and its subtle changes over time.

These land-cover-based characterizations of human settlements are insufficient to portray the conditions under which human settlements could be steered to more sustainable trajectories. Despite that LCZ scheme provides a potential solution for addressing such issue, LCZ classes are more climate-driven and are less capable to represent land-use pattern in a broader context. In land system science, land systems that account for more than land cover have been classified at various scales (Dou et al., 2021; Levers et al., 2015; van Asselen and Verburg, 2012; van der Zanden et al., 2016). These land-system products are generated primarily for agricultural or natural land, and human settlements are not well presented. Therefore, change trajectories of land systems are not well applied in the context of human settlements.

The other omission in most work on human settlements is the absence of attention for land use intensity, despite that management intensity has been widely investigated in land system science, which is, however, addressed primarily in agricultural systems (Erb et al., 2013; van Asselen and Verburg, 2012). In light of this, human settlements can also be characterized by their land-use intensity. Yet, we do not have much data and knowledge on the characterization and analysis of urban land-use intensity.

1.5 Research objective and outline of this thesis

In light of the aforementioned knowledge gap, the objective of this thesis is:

To advance large-scale analysis of human settlements and their dynamics through the lens of land system with a specific focus on the role of land-use intensity.

To achieve this objective, this dissertation will address the following three research questions (RQs):

RQ1: How can we represent the diversity of settlement systems?

RQ2: What can we learn from Earth Observation to better characterize urban land-use intensity?

RQ3: How do urban land-use intensity and settlement systems change over time?

These three RQs are addressed in the consecutive four chapters (see Figure 1.1). In Chapter 2, human settlements are mainly characterized by land cover component, and analyze their change trajectories. In Chapters 3-5, human settlements are represented with the incorporation of urban land-use intensity, which allows for a more nuanced analysis of human settlements and their dynamics.

Figure 1.1: Schematic display of the four research chapters in this thesis.

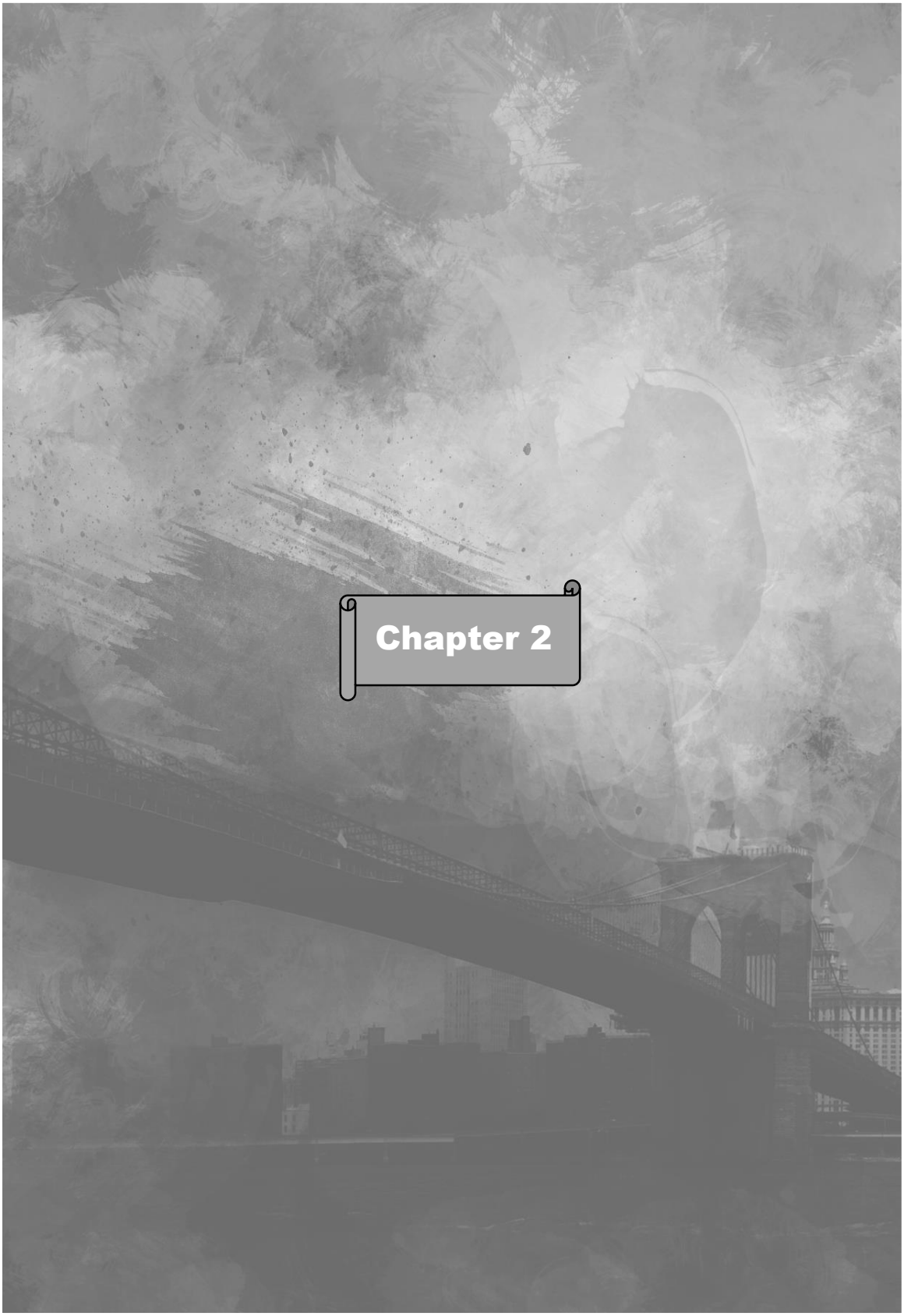
Chapter 2 is associated with **RQ1** and **RQ3**. As pointed out in Section 1.4, spatial analysis of human settlements is often limited to the mere presence and/or absence of built-up land and the changes therein, while subtler differences between various settlement systems are ignored. In Chapter 2, an experimental test is conducted to classify settlement systems in China. Specifically, built-up density, cluster density and cluster size are used to delineate settlement system, and analyze the observed change trajectories. On the one hand, the classification of settlement system could help us better understand the status and evolution of human settlement by accounting for land cover and land use pattern, rather than merely illustrating the locations of new

built-up land that emerged. On the other hand, the wide range of settlement system allows for further quantification of built-up land and its trajectories along urban-rural gradient, which could test the incompleteness of studies focusing on large cities only.

RQ2 is addressed in Chapters 3, 4 & 5 by accounting for 3D building structure, and built-up land per capita as the operationalization of urban land-use intensity. In Chapter 3, a new approach is developed to mapping 3D building structure, i.e., building footprint, height, and volume at a continental scale. Chapter 4 basically further improves the new approach proposed developed in Chapter 3 and applies it at a global scale. Different from the former chapter, this chapter emphasizes the analysis of heterogeneity in 3D building structure across space. In Chapter 5, built-up land per person is used to measure urban land-use intensity at multiple points of time.

In addition to RQ2, Chapter 5 is also associated with **RQ3**. Given the limitation in mapping 3D building structure at two points of time, built-up land per person is used alternatively as the operationalization for measuring urban land-use intensity in Chapter 5. Specifically, the long-term discussion on decreasing land-use intensity in cities worldwide is revisited. Apart from the changes in urban land-use intensity, built-up land change is decomposed into growth related to population dynamics and growth related to changes in land take per person, for more than 75000 administrative regions, typically representing municipalities or counties. By incorporating human settlements of different size and their peripheries, this chapter builds on the knowledge in scientific community that urban land-use intensity has been decreasing for centuries, to illustrate how existing concentrated studies in large cities only may hamper knowledge generalization.

The final Chapter 6 synthesizes the most important results and concludes the main findings of previous chapters as well as knowledge gaps this thesis has filled. Furthermore, perspectives for future study and policy implications are discussed.


Chapter 2

2 | Mapping settlement systems in China and their change trajectories between 1990 and 2010

A wide variety of settlement systems exist, ranging from small villages to large metropolises. However, spatial analyses are typically confined to the mere presence or absence of built-up land, and the changes therein, while subtler differences between various settlement systems are ignored. In this paper we study the spatial distribution of Chinese settlements in terms of their built-up land, cluster density and cluster size, as well as the changes therein between 1990 and 2010. Subsequently, we use these three properties to delineate settlement systems, and analyze the observed change trajectories between 1990 and 2010. We find that roughly 70% of all built-up land and more than 50% of all new built-up land added between 1990 and 2010 is included in village landscapes, which challenges the current focus on studying mega-cities. We also find that settlement changes mostly follow small and incremental steps towards denser urban systems, following multiple different development trajectories. Specifically, rural villages seldom convert into urban systems directly, but typically increase gradually to towns and sub-urban landscapes. The conception of settlement systems provides a first step towards a more comprehensive understanding of human settlements and their trajectories, which can inform targeted land use planning and the development of policies that more explicitly accounts for diversity in settlement types.

This chapter is published as:

Li, M., van Vliet, J., Ke, X., & Verburg, P. H. (2019). Mapping settlement systems in China and their change trajectories between 1990 and 2010. *Habitat International*, 94, 102069. DOI: 10.1016/j.habitatint.2019.102069

2.1 Introduction

The vast majority of the global alterations in land cover relate to the conversion of natural ecosystems into agricultural land. Human settlements initially emerged as an integral part of these agricultural systems. However, as settlements developed, they became increasingly disconnected from their resource base (Cumming et al., 2014). Due to population growth, migration, and economic development, settlements have been growing across the globe, especially in recent decades (Angel et al., 2011), to the point that built-up land is now competing for land with food production and other land demands (Bren d'Amour et al., 2017; Lambin and Meyfroidt, 2011; van Vliet, 2019). As the world population is expected to continue growing for at least the next few decades, there is a need for a nuanced understanding of the different forms of human settlements and their change processes over time, in order to inform land use policies and allow landscape planning to fully account for the diversity in forms of human settlement.

Urban growth has been studied extensively in recent years, mostly in terms of the expansion of built-up land and its spatial patterns. In these studies, the characterization of land is typically divided into urban and non-urban land, and as a result many studies are based on the conversion of non-built-up land to built-up land (Long et al., 2014; Mozumder et al., 2016; Yue et al., 2013), and the environmental impacts of these conversions (Angel et al., 2011; Poelmans et al., 2011; Seto et al., 2012a). In particular, the spectacular growth of large metropolises has received a considerable amount of attention (Ianoş and Jones, 2019; Schneider and Mertes, 2014; Xu et al., 2018b). Furthermore, several studies also use derivatives of built-up land, especially landscape metrics, to further characterize the environmental impacts of human settlements (Alberti et al., 2007; Ramachandra et al., 2015). Yet, the analysis of changes between built-up and non-built-up land only ignores the spatial pattern of built-up land and as a result, we do not know how changes in built-up land affect these patterns. Moreover, the focus on large urban areas obscures our view on other changes in built-up land affecting smaller settlements.

Increasingly land use typologies are used to characterize typical combinations or land uses or land covers and to study the changes therein, aiming to provide actionable solutions to sustainability challenges (Baiocchi et al., 2015; Loures and Vaz, 2018; Rounsevell et al., 2012). Notably, Václavík et al. (2013) identified 12 land system archetypes globally, based on a variety of variables, pointing out that land systems resemble across the globe but spatial pattern differs at subnational scales. At a national scale Ornetsmüller et al. (2018) identified land systems in Lao PDR based amongst

others on the typical combination of agricultural land and forested land, in order to analyse gradual shifts from swidden-forest mosaics to permanently managed agricultural land. However, settlement systems are not normally represented in these studies, as they are mostly focused on agricultural landscapes. Although different urban types have been identified (Dadashpoor and Ahani, 2019; Schwarz, 2010; Wu et al., 2018), there are only a few investigations on spatial patterns in urban land. For example, focusing on vacant land within urban landscape, Kim et al. (2018) developed a vacant land typology for Roanoke in Virginia, which can be potentially used for optimizing vacant parcel configurations at city scale. Krehl and Siedentop (2018) theoretically developed a classification scheme for urban land based on land use, built environment and others, which was subsequently applied in four German urban regions. However, characterizing settlement typologies at landscape and national scales is also crucial, as these scales are predominantly targeted by land use policies (Cullum et al., 2016; Levers et al., 2015).

The objective of this study is to identify settlement systems at a national scale, specifically for China, and analyse their recent change trajectories. To that effect we first analyse the distribution of built-up land, but also the size and density of clusters of built-up land, in order to provide a comprehensive picture of settlement changes. Subsequently, we use these characteristics to develop a characterization of settlement systems, and analyse the different change trajectories between 1990 and 2010. China is currently the most populated country in the world and it has developed quickly from a relatively rural to a relatively urban society in recent decades. As a result of this societal transition, it has experienced unprecedented urban growth, fuelled by a large-scale migration from rural to urban areas (Kourtit et al., 2014; Shi et al., 2017). This urbanization has led to large-scale expansion of built-up land, which has been well-studied for some of the larger metropolitan areas, including Beijing-Tianjin-Hebei region, Yangtze River Delta, Pearl River Delta, and some other areas (Li et al., 2018b; Schneider et al., 2015; Tian et al., 2011; Wu et al., 2015). Yet, there is much less understanding of the dynamics of smaller towns and villages in China. With the characterization of settlement systems, this study analyses multiple different change trajectories, and thus provides a more comprehensive view of settlement changes in China by also explicitly addressing the dynamics outside the main urban areas.

2.2 Materials and methods

2.2.1 Analysing settlement patterns and their changes over time

We use three properties to characterize settlement systems: the density of built-up land, the density of built-up clusters, and the size of the largest cluster (Figure 2.1a). These three properties are selected because they represent different characteristics of settlement systems. In particular, the density of built-up land resembles the traditional focus on the distribution of built-up and non-built-up land, the density of built-up clusters characterizes the extent to which built-up land is dispersed or concentrated within a landscape, and the size of the largest cluster is important to identify the hierarchy in human settlements as determined by their size and their influence on neighbouring settlements. All three properties are calculated at ~ 2-km resolution based on China's Land-Use/Cover Datasets (CLUDs) for 1990, 2000 and 2010 (see Table 2.1) in an Albers equal area projection, with standard parallels of 25°N and 47°N. The exact spatial resolution of output data is 2010 meters, as this is an exact multiple of the 30-meter input data, to which we refer as 2-km cells hereafter. This resolution is chosen because it reflects the landscape level at which settlement systems exist, rather than the local land cover at a location. While CLUD is also available for 2015, this data is not used in this study, since exploratory data analysis reveals that it is not consistent with data from other years. The CLUDs include six land use classes with 25 subclasses. We characterize and analyse settlement systems primarily based on urban land (subclass-51) and rural residential land (subclass-52) only. These two subclasses mainly differ in administrative properties, but since we are interested in land cover rather than administrative classifications, they are combined in this study and henceforth referred to as *built-up area*. Both classes represent built-up area following the classification of Liu et al. (2014c), as they predominantly consist of impervious surface, complemented with vegetated area, barren land, and water. These classes exclude lands used for industrial purposes such as quarries, factories, mining, as well as transportation outside cities (Kuang et al., 2016), which might otherwise obfuscate the delineation of human settlements. All analyses are conducted for contiguous China, i.e., excluding Taiwan, the islands in South China Sea and some other small islands around the continent, but including Hong Kong and Macau.

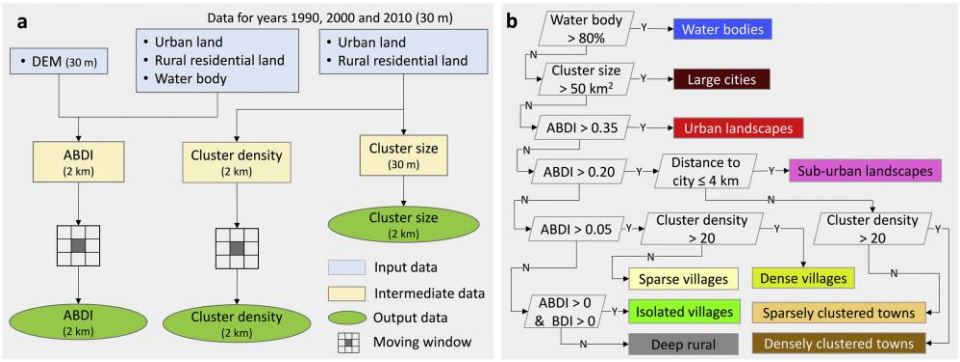


Figure 2.1: (a) Flowchart of generating the three properties to characterize settlement systems; (b) Expert-based classification scheme used to categorize settlement systems in China. "City" refers to the combination of large cities and urban landscapes. BDI is the ratio of built-up area to the total area of a 2×2 km cell, as such, BDI > 0 excludes cells that no built-up land exists within. To enhance visual interpretation of settlement systems distribution in the resulted maps, deep rural and water bodies are also presented as two separated categories.

To calculate built-up density, we define the Adjusted Built-up Density Index (ABDI) as the built-up area within a 3×3 moving window around a 2-km cell, divided by the total area that is suitable for built-up land therein. Suitable land here refers to all areas that are not water or permanent snow, and that have a slope that does not exceed 25%. The 25% threshold is based on the *Code for Vertical Planning on Urban and Rural Development Land of China* (CJJ 83-2016), which indicates that areas with a slope greater than 25% are not suitable for construction. A moving window approach is applied because settlement systems relate to the scale at which the system or process exists or responds (McGarigal and Cushman, 2005), which we interpret to be larger than the 2-km cells on which we define settlement systems.

Cluster density indicates the number of clusters of built-up area within a 3×3 moving window around each 2-km cell. To compute cluster density, we firstly convert original 30-meter CLUDs data into simplified polygons, effectively generating patches of built-up area, and subsequently generate feature points within these polygons that represent these patches. Patches are defined by direct as well as diagonal neighbours in this procedure. Then, we calculate the number of points within each 2-km cell and sum these results using a moving window method. The representation of clusters as points avoids double-counting patches that are spread over multiple 2-km cells.

Table 2.1: Datasets used in this study.

Dataset	Selection/process	Resolution	Period	Source	Reference
Built-up area (from China's land-use/Cover datasets (CLUDs))	Classes 51 (urban land) and 52 (rural residential land)	30 meters	1990, 2000, 2010	Institute of Geographical Sciences and Natural Resources Research	Liu et al. (2003)
Water/permanent snow cover (from China's land-use/Cover datasets (CLUDs))	Classes 41-46	30 meters	1990, 2000, 2010	Institute of Geographical Sciences and Natural Resources Research	Liu et al. (2003)
DEM	Derive slope	30 meters	/	Aster GDEM v2, USGS Earth Explorer	Tachikawa et al. (2011)
Administrative boundaries	/	Vector	/	Resource and Environment Data Cloud Platform	http://www.resdc.cn/

Cluster size is calculated as the size of the largest patch of built-up area that has at least some area in a 2-km cell. Similar to the analysis of cluster density, patches are identified based on direct as well as diagonal neighbours, using the algorithm settings of Map Comparison Kit (Visser and de Nijs, 2006).

We calculate ABDI, cluster density, and cluster size for the years 1990, 2000, and 2010, and analyse changes in each of the three properties, as well as the relation between different types of changes. First, we analyse how changes in ABDI, cluster density, and cluster size relate to the starting situation, e.g. in order to analyse whether new built-up area appears in already built-up areas, or instead in sparser areas. Then, we compare results for different years to explore changes in these properties over time. To explore the relation between changes in these properties, we categorize changes in each of the properties on a 2-km cell level as decrease, stable or increase. Subsequently, for each pair of properties, we analyse how changes in one property relate to changes in another property, in order to characterize spatial patterns of change. The thresholds for ABDI, cluster density, and cluster size are selected using a histogram-based method (see for example Yu et al. (2018)). Specifically, we create histograms showing the distribution of changes, and identify the boundary between change and no change based on the absolute amount of change, so that at least 2% of all cells were identified as changes. This value of 2% is chosen for all classes to make their changes comparable, while at the same time accounting for the relatively static nature of built-up area. Associated threshold values for each change category are presented in Table 2.2.

Table 2.2: Thresholds for classifying changes in ABDI, cluster density and cluster size.

Property	Threshold	Proportion of the study area that is stable	
		1990-2000	2000-2010
ABDI	± 0.02	97.59%	97.91%
Cluster density	± 1	97.78%	97.72%
Cluster size	± 0.09 km²	97.53%	97.78%

The accuracy of spatial analyses is highly dependent on the accuracy of the input data. The CLUDs maps have been validated, yielding an overall accuracy of more than 90% (Liu et al., 2014a). Yet, due to our aggregation into tiles of 2-km, we expect that the accuracy at this resolution is effectively higher, and hence sufficient for the purpose of this study.

2.2.2 Mapping settlement systems and their changes over time

We use a decision tree (Figure 2.1b) to map settlement systems based on the three characteristics that are also used for the analysis of settlement patterns: ABDI, cluster density and cluster size. In addition, we use the relation between clusters in the neighbourhood (i.e., proximity). Decision trees are expert-systems, which are preferred over statistical clustering methods as the latter are found sensitive to the selected distance metric and criteria for determining the order of clustering (van der Zanden et al., 2016). We identify archetypical settlement systems in cells with a 2-km resolution, similar to the analysis of settlement patterns described above. For all cells that are not water, we first identify those that contain *large cities*, i.e. clusters of built-up area larger than 50 km². Subsequently, *urban landscapes* are defined as cells with an ABDI above 0.35, indicating landscapes with a predominantly urban character. *Large cities* and urban landscapes together, are used to generate a proximity layer for detecting *sub-urban landscapes*. The remaining areas where categorized as clustered towns and clustered villages, based on their ABDI and cluster density. *Isolated villages* are characterized by their small fraction (< 0.05) of built-up area as indicated by lower threshold of ABDI, while cells with no built-up area at the 2-km resolution are represented as *deep rural*. We tested a range of threshold values for identifying settlement systems and found that minor changes in threshold values did not yield considerably different patterns, but only marginal displacement of the class boundaries, suggesting a robust classification method.

To analyse changes between settlement systems, we calculated two change matrices, for the periods 1990-2000 and 2000-2010, respectively, and identified the major conversions relative to each settlement system. These change matrices are subsequently used to identify settlement change trajectories, based on the relative occurrence of changes from one settlement system to another.

2.3 Results

2.3.1 Spatial pattern of built-up area and its changes in China

ABDI, cluster density and cluster size for year 2010 are presented in Figure 2.2 (see supplementary Figure A1 for years 1990 and 2000). High values for ABDI are, not unexpectedly, concentrated around the larger urban areas, such as Beijing, Shanghai, Guangzhou, and Hong Kong. Yet, Figure 2.2a also shows that there is a very large part of China that has at least some built-up area outside these large urban areas, albeit mostly with an ABDI lower than 0.20. Cluster density, on the other hand, shows the

highest values not in the urban areas, but in the intensive agricultural areas, such as the North China Plain, Northeast China Plain, and the Chengdu Plain (Figure 2.2b). The distribution of cluster sizes reflects the location of large urban areas, as these are reflected in the largest cluster sizes, while the cluster size quickly reduces farther away from these large urban areas (Figure 2.2c).

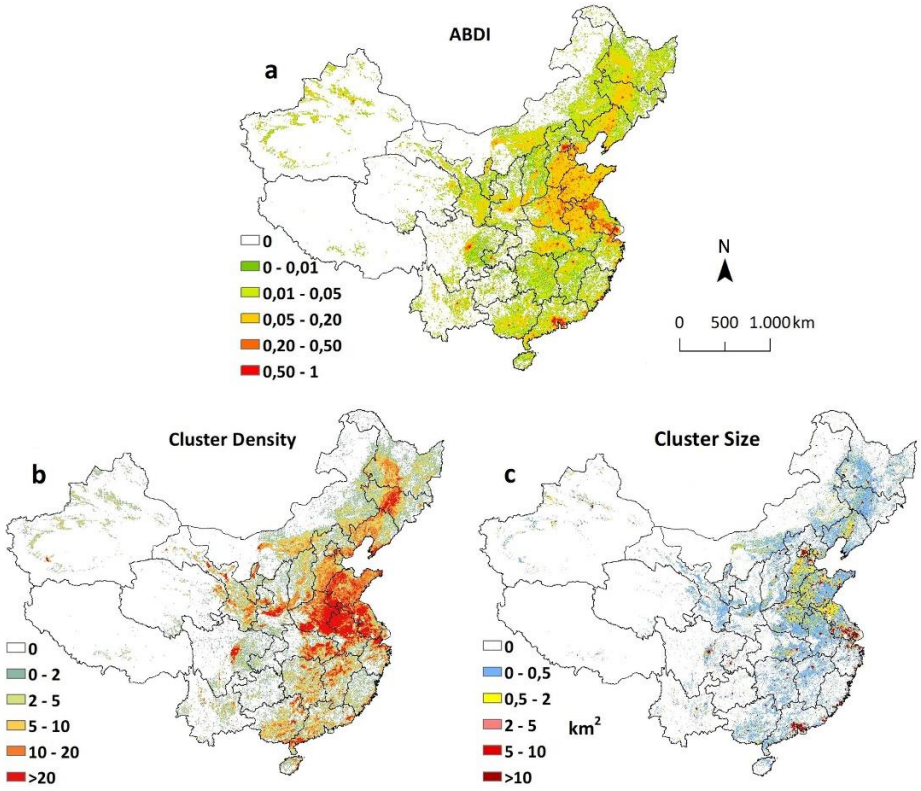


Figure 2.2: ABDI, cluster density and cluster size for year 2010 (a-c). Comparable figures for 1990 and 2000 are provided in the supplementary material. Only the study area of contiguous China is shown.

The combined changes in ABDI, cluster density and cluster size reflect a spatial restructuring process of settlement systems. As is illustrated in Figure 2.3, most non-stable areas show an increase in both ABDI and cluster size, while patterns in terms of cluster density are spatially heterogeneous. Two typical combinations of ABDI and cluster density changes were found (Figure 2.3a). The combination of increased ABDI and decreased cluster density reflects urban areas growing connected, while a higher cluster density in combination with a higher ABDI indicates the appearance of new urban clusters (with or without edge-growth). Figure 2.3b indicates that an

increase in ABDI mostly coincides with an increase in cluster size, which is typically the result of edge growth of existing urban areas, which affects both in the same direction. The relation between cluster density and cluster size varies mostly between two typical combinations. The combination of lower cluster density and larger cluster size typically reflects consolidation, i.e. formerly isolated clusters growing together. The opposite, a higher cluster density in combination with a larger cluster size, indicates a combination of different processes, such as edge growth and the appearance of new scattered clusters. There is little difference between the changes in the both periods.

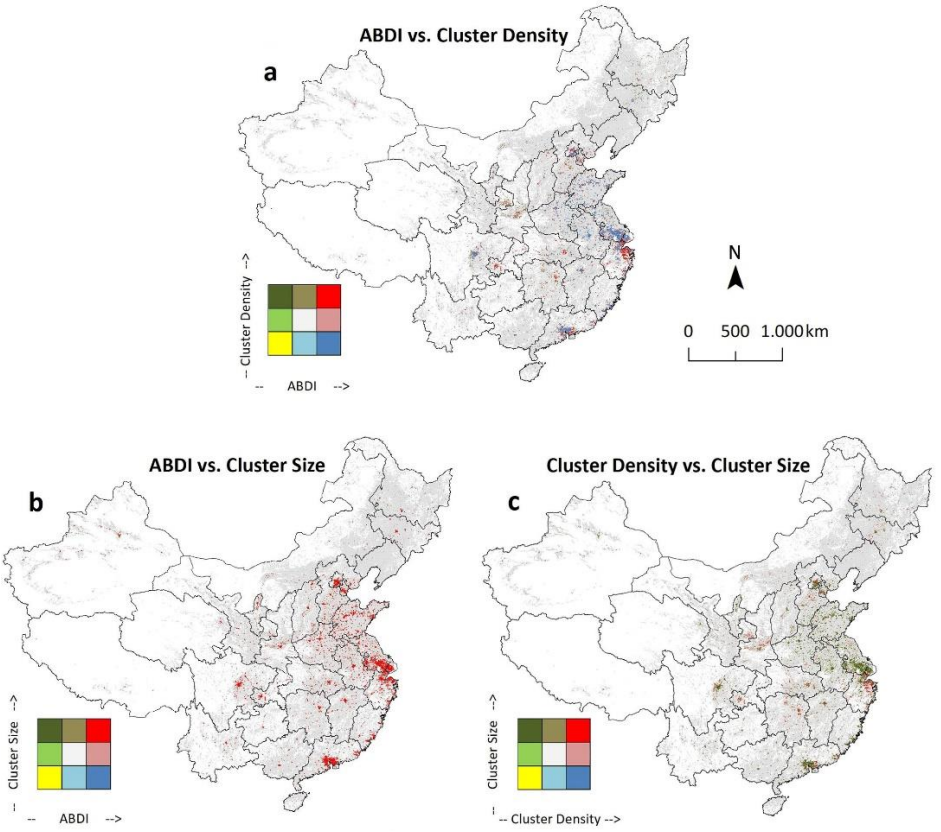


Figure 2.3: Relations between changes in ABDI, cluster density, and cluster size during 2000-2010. Comparable figures for changes between 1990 and 200 are provided in Figure A2. Only the study area of contiguous China is shown.

Analyses of changes in ABDI, cluster density, and cluster size as a function of their original values for these indicators reveal some typical patterns. The average increase in ABDI is largest for cells that are already have an ABDI between 0.5 and 0.8, while

the average increases in cells with a lower ABDI is much smaller (Figure 2.4a). It should be noted however that the number of cells with a low ABDI at the start of the period is much larger, and that therefore these cells still accommodate the majority of new built-up area (see supplementary Figure A3). Cluster density can increase over time due to the appearance of new clusters but also decrease as a result of clusters growing together. The net effect is that for most cluster densities, the average change is near zero, except for cells that start without any built-up area, as the cluster density can only increase here (Figure 2.4b). Changes in cluster size show that large clusters often also experience large changes in size, while small clusters typically experience only smaller changes (Figure 2.4c, also see supplementary Figure A4). Additionally, some other small clusters witness a considerable increase in size, especially between 2000 and 2010, which indicates the consolidation process as mentioned that small clusters in close proximity to large ones get connected to each other, resulting in a large increase in the size of the largest clusters as well as a decrease in cluster density.

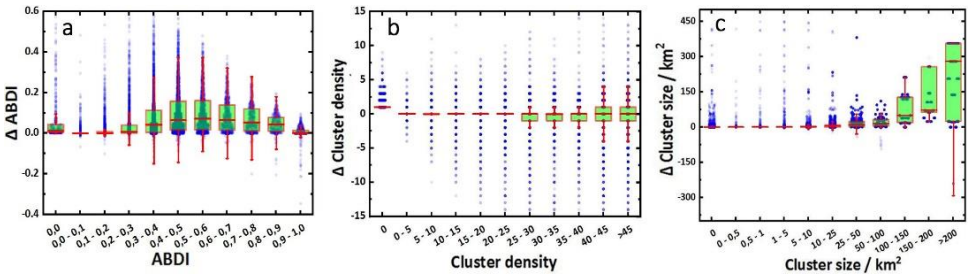


Figure 2.4: Changes in ABDI, cluster density and cluster size as a function of their initial value during 2000-2010. The horizontal bars in each boxplot correspond to the 25th, 50th, and 75th percentiles. The whiskers extend to 1.5 times the interquartile ranges. Supplementary Figure A4 provides results for the period 1990-2000, and a more detailed view on changes in cluster size is shown in Supplementary Figure A5.

2.3.2 Settlement systems and their changes over time

Combining the characteristics of built-up area allows the categorization of different settlement systems in China. About 20% of all land in China can be characterized as a settlement system, mostly concentrated in the eastern part of the country (Figure 2.5). Conversely, *deep rural*, including both wild land and agricultural areas without any built-up land, accounts for almost 80% of the total area, which is mostly located in western part of the country. The distribution of settlement systems as well as built-up area within each system is presented in Figure 2.6. In 1990, 8.50% of all land in China is characterized as *isolated villages*, comprising 17.58% of all built-up surfaces. *Sparse villages* take up only 6.89% of all land, yet incorporate 35.38% of built-up area,

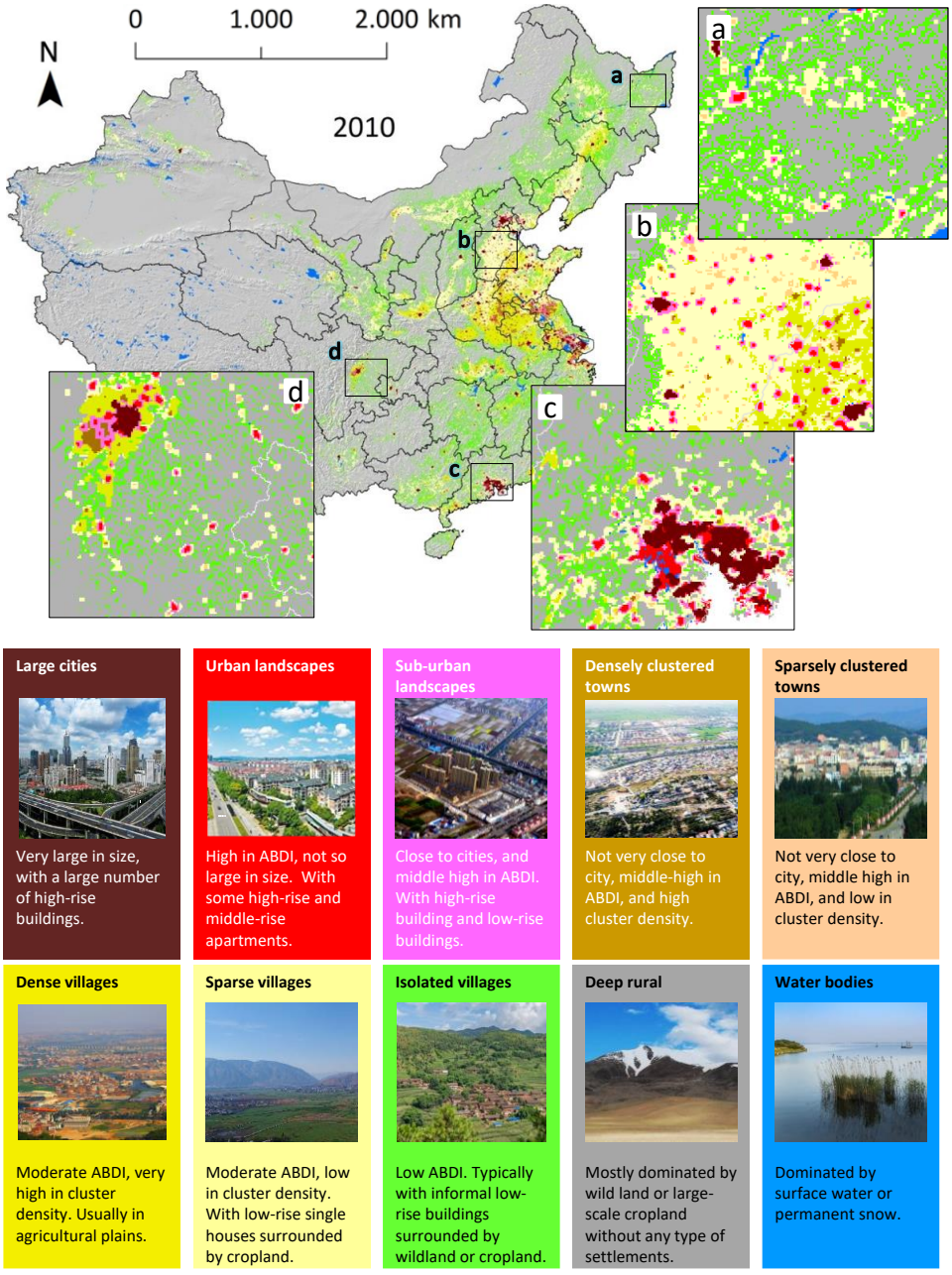


Figure 2.5: Settlement systems in 2010. Comparable settlement system maps for 1990 and 2000 are provided in supplementary Figure A6. As a holistic map, deep rural and water bodies are also presented to enhance visual interpretation of settlement systems distribution. Only the study area of contiguous China is shown.

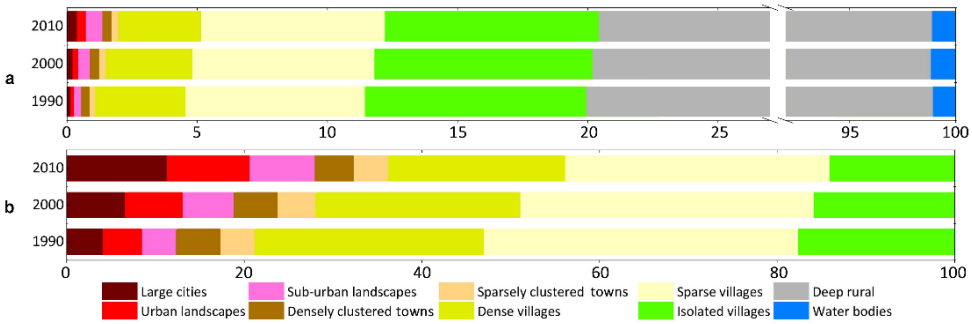


Figure 2.6: a) Percentage of all land occupied by each settlement system. b) Percentage of built-up area included in each settlement system. No built-up area is found in deep rural area which is consistent with its definition.

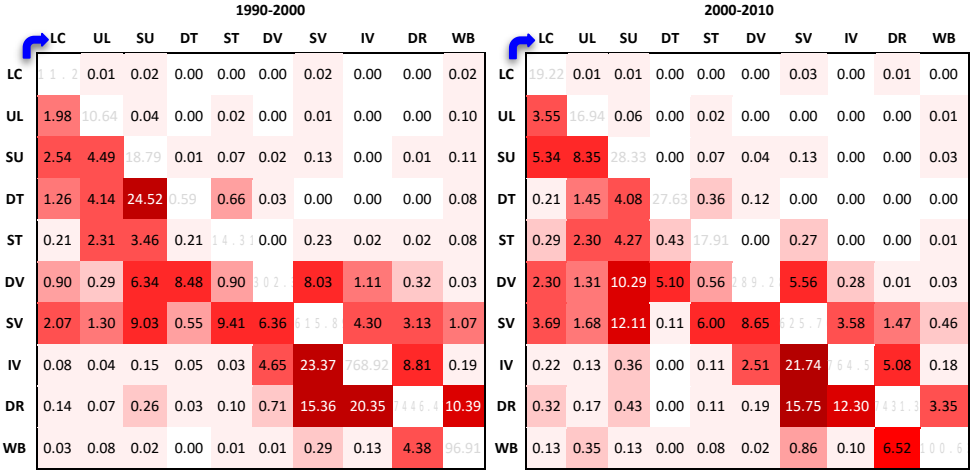


Figure 2.7: Settlement trajectory matrices for 1990-2000 and 2000-2010, with values in 1000 km² of land area. The shades of red correspond with the size of the observed changes. A darker pane highlights a larger area that is transformed. Legend: LC = large cities, UL = urban landscapes, SU = sub-urban landscapes, DT = densely clustered towns, ST = sparsely clustered towns, DV = dense villages, SV = sparse villages, IV = isolated villages, DR = deep rural, WB = water bodies.

the largest share among all land systems. *Dense villages* are responsible for 3.46% of all land, while its proportion of built-up area reaches to 25.83%. Other settlement systems including *sparsely clustered towns*, *densely clustered towns, sub-urban landscapes*, *urban landscapes*, and *large cities* together take up 1.08% of the land area, but account for 21.23% of all built-up area. In 2010, village landscapes (*isolated villages*, *sparse villages*, and *dense villages*) accounted for 18.46% of the land, which is only slightly less than in 1990. The

proportion of built-up area included in these three settlement systems has decreased from 78.80% in 1990 to 63.75% in 2010. *Sparsely clustered towns* and *densely clustered towns* combined account for 0.62% of the land in 2010, which is slightly more than that proportion of 0.55% in 1990. Conversely, proportion of built-up area in these two types of towns show a slight drop from 8.83% to 8.27%. The proportion of China's built-up area that is found in in *large cities*, *urban landscapes*, and *sub-urban landscapes* more than doubled, from 12.29% in 1990 to 27.91% in 2010.

Changes in settlement systems between 1990 and 2000, as well as between 2000 and 2010 show a development towards denser settlement systems, although this development mostly comes in small incremental steps, rather than sudden large-scale changes. Figure 2.7 shows the transformation matrices of settlement systems. Between 1990 and 2000, nearly all *densely clustered towns* developed into *sub-urban landscapes* which is mainly a result of the fast increase of *large cities* and *urban landscapes* in the surroundings. Another prominent change trajectory is from village landscapes into *sub-urban landscapes,* also reflecting a process of urban sprawl in which the former villages are embedded in the urban landscape of growing cities nearby. At the same time, a large share of the *sub-urban landscapes* developed into *large cities* or *urban landscapes,* especially in the period 2000-2010. Consistently, *large cities* and *urban landscapes* basically gain from *sub-urban landscapes*, indicating a continuous process of urban expansion. In addition to this typical pattern of urban growth, this transition matrix reveals the important dynamics in the more rural parts of the spectrum. Large portions of *deep rural* change into village landscapes as well as the conversion of *isolated villages* to *sparse villages*, and *sparse villages* into *dense villages*, all indicating the appearance of new village clusters over time in these areas. This indicates that not only large settlement change is taking place near the cities, but even in the more remote areas which initially only had marginal fractions of build-up area.

Selecting from the transformation matrices the most important changes per settlement system reveals typical change trajectories. Figure 2.8 shows these change trajectories for the period 2000-2010. Settlement changes mainly follow small incremental changes between increasingly urban landscapes rather than sudden transformation from rural to urban landscapes. Moreover, this figure also shows that multiple settlement change trajectories occur at the same time. For example, *sparse villages* changed into *sparsely clustered towns*, *dense villages* and *sub-urban landscapes*, reflecting respectively an increase in ABDI, an increase in cluster density, and an increase in large urban areas in the neighbourhood.

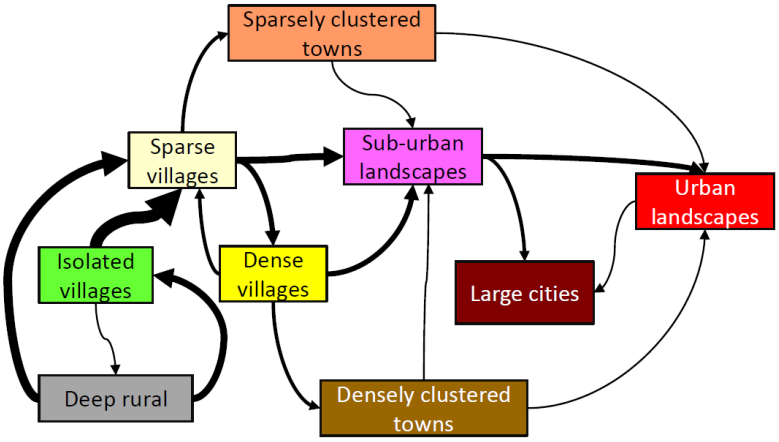


Figure 2.8: Main transitions between settlement systems observed in China between 2000 and 2010. Depicted transitions are selected to show at least 70% of all observed changes from any particular land system starting from the largest transition category, and to have at least one arrow feed into each class. The width of the arrows is proportionally related to the land area of the corresponding settlement system change.

2.4 Discussion

2.4.1 Settlement systems and their change trajectories in China

The map of settlement systems in China shows a distribution ranging from completely built-up to completely non-built-up, with the vast majority of built-up area distributed as small fractions in otherwise rural landscapes. Overall, more than 20% of all land in China is classified as one of the settlement systems in 2010. Changes in built-up area are found throughout this area, but the majority of new built-up area is added to areas with an initial low urban fraction (i.e., village growth). For instance, between 1990 and 2000 more than 65% of all new built-up area emerged in village landscapes and deep rural areas. This percentage drops for the period 2000-2010, but still remains above 50% of all new built-up land. Hence, while locally the changes in metropolises can be dramatic, the aggregated effect of increases in smaller settlements is at least as large. These findings resemble those found for Europe (van Vliet et al., 2019). This result suggests that analyses of urban growth in large metropolitan areas that have received much attention in recent years (Gong et al., 2018; Xu et al., 2007; You and Yang, 2017) are not necessarily representative for built-up area expansion in China. Instead, it suggests that there is a need to analyse settlement systems more comprehensively by also including subtler change processes in village landscapes, towns, and peri-urban areas.

Large metropolitan areas, particularly Beijing-Tianjin-Hebei region, Yangtze River Delta, and Pearl River Delta, are identified as *large cities*, *urban landscapes*, and *sub-urban landscapes*. Between 1990 and 2010 each of these three regions is characterized by typical urban growth processes, as reflected through an increase in built-up area, and an increase in cluster size in many cells (also see Xu et al. (2007), You and Yang (2017) and Gong et al. (2018)). As a result of their rapid growth, many *large cities* in these areas are facing a scarcity in land and a pressure on existing infrastructure. In order to deal with these constraints, the Beijing municipal government is gradually moved to Tongzhou district, which is quite distant from central Beijing (CGTN, 2017). At the same time, Tongzhou is now connected to downtown Beijing, leading to a continuous area characterized as *large cities*. The development in the more densely populated areas elsewhere in China are also fuelled by national level policies. For example in Xiong'an New District in northern China, which was initiated by the Central Committee of the Communist Party of China and the State Council in 2017 with a long-term control area of about 2000 km^2 (Kuang et al., 2017; Xu et al., 2018a).

The most productive agricultural areas of China, notably the North China Plain, Northeast China Plain, and Guanzhong Plain, are characterized by sparse and dense village landscapes. These settlement systems reflect the patterns that emerged exactly because of the agricultural character of the region, as they are relatively densely clustered, but the clusters themselves are smaller. However, due to the economic development in all of eastern China, as well as population growth and migration from more remote parts of the country (Cao et al., 2018; Kuang et al., 2014; Li et al., 2018a), there is an increase in urban and sub-urban landscapes in these prime agricultural regions. For example, Jiangsu, a typical developed coastal province, experienced a dramatic development towards the more urban settlement systems, especially for regions closer to the economic hub of Shanghai. The growth of urban areas is often attributed to migration from rural areas (Henderson et al., 2009). At the same time, our results show that there is no decrease in built-up land in the agricultural areas in between, nor in the more remote areas elsewhere in China. The south-eastern part of China, for example, is characterized by a rugged landscape, providing a natural constraint for both agricultural activities underlying the development of villages elsewhere as well as the emergence of large metropolitan areas. These mountain valleys are mainly filled with isolated villages, but these villages did not change much in recent decades or even showed small increases in ABDI.

The development of settlements in rural areas mostly consists of low-density developments, thus leading to relatively large amounts of land take and soil sealing for the benefit of few people. As these developments often come at a cost of

agricultural lands, it further contributes to competition for land, thus affecting food security and biodiversity habitat (van Vliet et al., 2017). This problem is not unique to China, as similar challenges have been observed elsewhere (Gardi et al., 2015; Tóth, 2012). There is an opportunity for land use planning and national policies to reduce such conflicts by restricting built-up expansion and fostering more compact settlement types. In China, there is a long history through compensation policies and densification incentives (Liu et al., 2014b; Long et al., 2012). For example, the lost villages in Shandong Province are likely a policy effect (Li et al., 2018b), to cater for the national policy termed "increasing vs. decreasing balance of urban-rural construction land", which aims to balance the total construction land and hence compensate for urban development elsewhere. Yet, our results show that these initiatives have not effectively controlled the process of increasing built-up area in these regions. On top of that, such policies can have trade-offs on social conditions (Howley, 2009; Schindler and Caruso, 2014) and may challenge rural cultural heritage (Yu et al., 2016), indicating that plans need to carefully consider local conditions and impacts.

The need for a more nuanced understanding of settlements and settlement changes is relevant beyond Chinese territory. As in China, analyses of settlement changes elsewhere have focused mainly on few large urban areas (Bagan and Yamagata, 2012; Georg et al., 2018; Taubenböck et al., 2012). Meanwhile, a large part of the built-up land in the USA was found outside urban and sub-urban areas (Theobald, 2001), suggesting the necessity of land use analysis through the whole urban-to-rural continuum. Consistently, most built-up expansion in South America occurs in small cities and rural areas, not only concentrated around major cities (Andrade-Núñez and Aide, 2018). A study on selected European countries has also shown that most built-up land is embedded in predominantly rural landscapes (van Vliet et al., 2019). As a result of the typical mosaics of settlements and other land uses, large parts of Europe have been characterized as peri-urban, or "territories-in-between" (Alexander Wandl et al., 2014). Global simulations of future land use show that such peri-urban and densified village landscapes will cover increasingly large parts of the earth and therefore need more attention (van Vliet et al., 2017; Wolff et al., 2018b).

2.4.2 From built-up area to settlement systems

Urban growth is often analysed or simulated based on the conversion of non-built-up to built-up land, without further consideration of the structure of the landscape within which this conversion takes place (Yue et al., 2013). Otherwise, human settlements in China are usually indicated as cities, towns, and villages (Tian et al.,

2007). However, these classes reflect an administrative classification, rather than a landscape description. In addition, a breath of spatial indices have been used to study patterns of urban land cover (Liu and Yang, 2015). This study, in contrast, uses spatial pattern indices for the characterization of settlement systems, a typology within which specific settlement change trajectories can be studied in their specific contexts. These processes include processes such as village growth, urban expansion, and sub-urbanization. Such in-depth understanding of settlement change trajectories is an essential prerequisite for designing policies dedicated to a more efficient land use (Hersperger et al., 2018; Mustafa et al., 2018).

The identification of settlement systems resembles recent developments in land use science leading to the identification of *anthromes* (Ellis and Ramankutty, 2008), land systems (van Asselen and Verburg, 2012) and land system archetypes (Václavík et al., 2013). These classification approaches move beyond the characterization of the terrestrial biosphere based on the predominant land cover or its related land use, and instead acknowledge the spatial distribution and spatial patterns that characterize landscapes. While urban areas are included in these approaches, their focus is essentially on the agricultural and natural parts of the landscape, leading to relatively little differentiation in settlement systems (Ornetsmüller et al., 2018; van Vliet et al., 2019). Settlement systems, therefore, complement these approaches and allow for a more in-depth investigation of settlement change processes and a more nuanced view on urban systems for land use policies and planning.

Settlement systems could also provide a starting point for a more nuanced representation of settlement change processes in land-use models. As of yet, many land-use models are limited to the simulation of land cover conversions including a conversion from non-urban to urban land (van Vliet et al., 2019). For a large number of urban growth models, this is even the only type of change that is simulated (Kamusoko and Gamba, 2015; Mahiny and Clarke, 2012). The results of this study show the wide variety in settlement systems, as well as the different change trajectories that take place in parallel. Hence the representation of built-up land constrains the extent to which we can learn from such models or use them to explore solutions for sustainability challenges. To facilitate further research in this direction, we make the datasets generated in this study freely available (https://cscproject.github.io).

2.5 Conclusion

Settlement systems exist in a wide variety ranging from isolated villages to larger metropolitan areas. This study analyses changes in settlements beyond merely

assessing the conversion of non-built-up land into built-up land. Results show that settlement systems typically change gradually and incrementally, from villages to towns to peri-urban and urban areas. Moreover, results of this study also show that the combined increase in built-up land in smaller cities, towns and villages exceeds that of large urban areas. This suggests that there is a need to analyse settlement systems more comprehensively, and beyond the increase of a few mega-cities only.

Acknowledgments

Mengmeng Li would like to acknowledge funding from China Scholarship Council (Grant No. 201706510011). This research contributes to the Global Land Programme (www.glp.earth).


Chapter 3

3 | Continental-scale mapping and analysis of 3D building structure

Urban land use is often characterized based on the presence of built-up land, while the land use intensity of different locations is often ignored. This narrow focus is at least partially due to a lack of data on the vertical dimension of urban development. The potential of Earth observation data filling this gap has already been shown in literature, while large scale applications are still absent. This study aims to map urban land use intensity (expressed as 3D building structure, i.e. footprint, height, and volume) using multiple datasets such as satellite imagery and socioeconomic data for Europe, the US, and China. Our models perform well, as indicated by median R^2 values of 0.90 for building footprint, 0.82 for building height, and 0.88 for building volume, for all three case regions combined. In our multidimensional input variables, we find the built-up density derived from Global Urban Footprint (GUF) is the most important variable for estimating building footprint, while backscatter intensity of Synthetic Aperture Radar (SAR) is the most important variable for estimating building height. A combination of the two is essential to predict building volume. Our analysis further highlights the heterogeneity of 3D building structure across space. Specifically, buildings in China tend to be taller on average (10.35 m) compared to Europe (7.37 m) and the US (6.69 m). At the same time, the building volume per capita in China is lowest, with 302.3 m^3 per capita, while Europe and the US show estimates of 404.6 m^3 and 565.4 m^3, respectively. The results of this study (3D building structure datasets for Europe, the US, and China) can be used for further analysis of the urban environment, spatial planning and land use projections.

This chapter is published as:

Li, M., Koks, E., Taubenböck, H., & van Vliet, J. (2020). Continental-scale mapping and analysis of 3D building structure. *Remote Sensing of Environment*, 245, 111859. DOI: 10.1016/j.rse.2020.111859

3.1 Introduction

Monitoring urban extent has been prominent in earth-observation studies for decades, resulting in various products available from local to global scales (Carlson and Sanchez-Azofeifa, 1999; Gong et al., 2020b; Mertes et al., 2015; Schneider et al., 2009; Taubenböck et al., 2012). These urban extent products are crucial for environmental assessments to further address sustainability challenges such as food insecurity, biodiversity loss, and risk exposure (Angel et al., 2011; Du et al., 2018; van Vliet, 2019). Moreover, urban area has also been used for better characterization of the terrestrial biosphere, for instance using landscape mosaics, *anthromes*, and land systems (Ellis and Ramankutty, 2008; Messerli et al., 2009; van Asselen and Verburg, 2012).

Urban land use intensity can be considered as the equivalent of agricultural land use intensity, as it expresses the density or intensity of the use of agricultural land in a location. Consistently, urban land use intensity can be characterized in different ways, and it is not clear *a priori* what measure is preferable (see e.g. Kuemmerle et al. (2013) for a discussion on quantifying agricultural land use intensity, and Dovey and Pafka (2013) for a discussion on measuring urban density). Recent studies for example include population density (van Vliet et al., 2019), or a spatial characterization of urban structure (Susaki et al., 2014; Xia et al., 2020).

Urban structure involves both the horizontal and vertical configurations of urban land and infrastructure (Wentz et al., 2018), which influence both biophysical and socioeconomic conditions such as urban climate, carbon emissions, travel behaviour, and public expenditure (Connors et al., 2012; Engelfriet and Koomen, 2017; Hudeček et al., 2019). Urban structure has been extensively analysed in the literature, but based primarily on two-dimensional spatial metrics (Huang et al., 2007; Lowry and Lowry, 2014; Taubenböck et al., 2018b), apart from a few exceptions that incorporate vertical dimension for small-scale case studies (He et al., 2016; Kedron et al., 2019). Yet, 3D urban structure has important implications for sustainable urbanization. For example, compact urban structure contributes to reducing greenhouse gas (GHG) emissions on the one hand, but it could also worsen urban environment through the urban heat island effect on the other hand (Berger et al., 2017). An analysis of countries in Europe shows that built-up land is relatively concentrated in the Netherlands, while it is much more dispersed in, for example, Romania (van Vliet et al., 2019). In cities, the spatial distribution of human activity from a centre to its outskirt often follows a negatively sloped exponential curve (Bertaud, 2001). However, this pattern differs across the globe with Shanghai showing a much steeper gradient in human activity when compared to other megacities such as New York, London, and Paris (Bertaud, 2001).

Urban development across the globe follows different development trajectories, both horizontally and vertically. Singapore, for example, has built numerous high-rise and compact apartments to accommodate its growing population (Grace Wong, 2004). On the other hand, urban expansion in the global south is often characterized by the proliferation of low-rise slums (Badmos et al., 2018; Kusno, 2019; Wang et al., 2019a). A recent study on selected cities finds that urban development in the United States is dominated by decentralized-sprawl patterns, while central-compact patterns are typically found in Europe and China (Dong et al., 2019). Because the impacts of different types of urban development vary, there is a need to characterize urban development beyond two-dimensional spatial patterns, in order to assess urban sustainability. However, thus far these analyses have not included the vertical dimension of urban development, except for a number of studies analysing selected (mega)cities across the globe (e.g., Frolking et al. (2013), Straka and Sodoudi (2019), and Zhang et al. (2018)).

The significance and urgency of mapping 3D urban structure (Hereafter, we refer to as 3D building structure) are further highlighted in a recent review on urban remote sensing (Zhu et al., 2019), while compared to the identification of building extent, retrieval of an building vertical profile based on remote sensing is a more complex process. Yet, several investigations attempted to extract building height (Bagheri et al., 2018; Liasis and Stavrou, 2016; Weissgerber et al., 2017). A large number of remote sensing based data sources are available to retrieve building height, which can generally be categorized in four categories: conventional optical images, stereo optical images, Light Detection And Ranging (LiDAR), and Radar. LiDAR is widely acknowledged as the most robust source. However, applications of LiDAR-derived data are highly constrained by their coverage, as data is scarce, expensive, and scattered. Natural vegetation shows relatively little vertical variation over distance due to similar biophysical conditions in its ambience, and therefore it can be estimated based on sparse LiDAR points globally (Lefsky, 2010; Simard et al., 2011). However, buildings often vary considerably within the built-up environment. Recently, hydrid data have been used to characterize 3D building structure, for example, Geiß et al. (2019) present a multistep approach to estimate 3D building structure based on TanDEM-X and Sentinel-2 data. Evaluations of building volume estimates derived from LiDAR and Radar (both scatterometer and SAR) respectively reveal that these two source datasets are highly consistent (Bagheri et al., 2018; Mathews et al., 2019), suggesting that current fine-resolution Radar data can contribute to the estimation of 3D building structure at a larger scale.

As shown in a growing body of literature, 3D building structure is the basis for many analyses, such as landscape aesthetics, urban climate, and energy consumption (Güneralp et al., 2017b; Lin et al., 2018; Stewart and Oke, 2012), while such data are still scarce, especially at a large scale. This study aims to fill in this gap by mapping the 3D building structure, i.e. building footprint, height, and volume, for Europe, the conterminous United States (the US), and mainland China. On condition of the reference data collected from various sources, we use a large number of explanatory variables to map the 3D building structure using random forest models. To the best of our knowledge, this is the first study on such a large scale mapping of 3D building structure, especially from a land use perspective. Section 3.2 describes the spatial datasets and random forest model used for mapping 3D building structure in more details. Section 3.3 presents the results of this model, as well as an analysis of how building structure differs between our study regions. In Section 3.4 we further discuss these results, and reflect on the contribution of these data for sustainable settlement development.

3.2 Materials and method

3.2.1 Overview

In this study, we estimate building footprint, building height, and building volume at a 1-km^2 resolution for Europe, the US, and China. The US and China refer to the conterminous United States and mainland China (including Hong Kong and Macao), respectively. We choose a 1-km^2 resolution because the aim of this study is to characterize urban areas as a land use type, which can be used for further analysis of land use changes as well as their impacts. For these analysis individual buildings are of little interest as the related phenomena act at a coarser scale (e.g. van Vliet et al. (2019), Stewart and Oke (2012), and Wang et al. (2019b)). Building footprint denotes the share of each 1 km^2 pixel that is occupied by buildings (therefore expressed as m^2 per m^2). Building height denotes the average height of all buildings in a pixel, weighted by the area of each building. Building volume is the total volume within each pixel taken by buildings. Conceptually, building volume is the building footprint multiplied by the average building height in a pixel, although all three are predicted independently in our study.

We train random forest models to estimate building footprint, height, and volume using reference data for different locations of the study areas, and subsequently use these trained models to estimate the same variables for all the other locations within our study areas. These study areas are mainland Europe, the US, and mainland China

(including Hong Kong and Macao, hereafter referred to as China). Figure 3.1 illustrates the overall approach of our study. This approach consists of four parts: 1) the collection and pre-processing of *spatially explanatory variable* to produce ready-to-use inputs for the model using the Google Earth Engine (GEE). GEE is a cloud-based platform for geospatial analysis at a planetary scale, which also consists of various ready-to-use datasets, co-located within a high-performance, intrinsically parallel computation service (Gorelick et al., 2017); 2) collection and pre-processing of the *reference data*, including both readily available 3D building data and manual interpretation of 3D building structure based on Very High Resolution (VHR) satellite/aerial imagery and street view; 3) training, optimizing and validating the *random forest models* to produce 3D building structure maps; 4) *spatial analysis of building properties* in the three study regions and the differences between these regions.

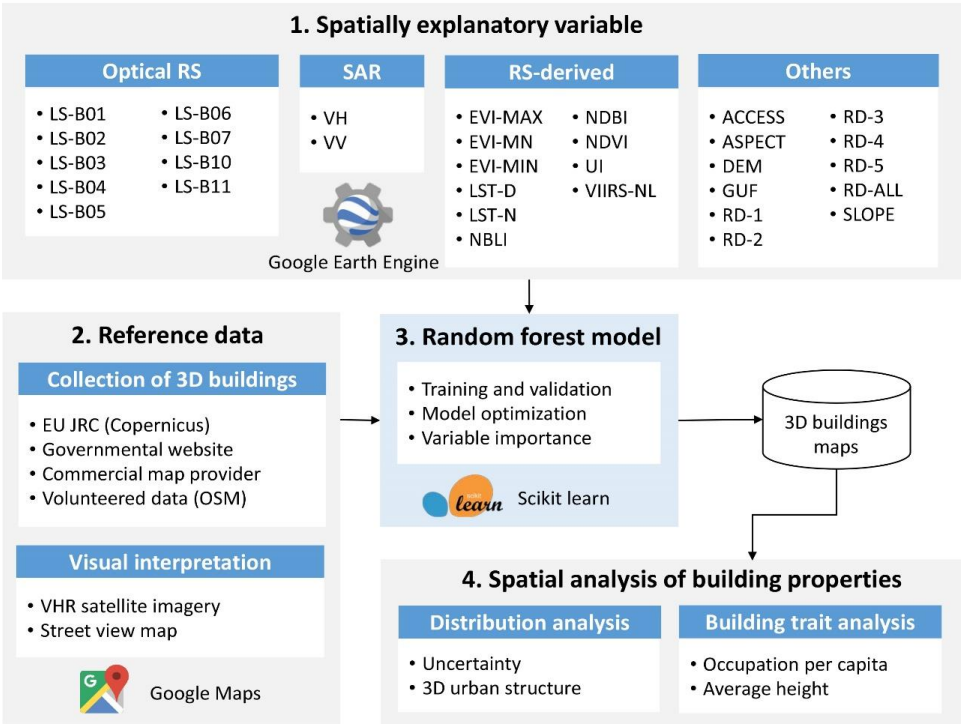


Figure 3.1: Methodological approach for the estimation of 3D building structure at a continental scale. The abbreviations for spatially explanatory variable are referred to in Table 3.1.

Table 3.1: Spatially explanatory variables used in this study.

Datasets		Original resolution	Time	Source	Abbreviation
Optical RS	Landsat Band 1	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B01
	Landsat Band 2	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B02
	Landsat Band 3	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B03
	Landsat Band 4	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B04
	Landsat Band 5	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B05
	Landsat Band 6	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B06
	Landsat Band 7	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B07
	Landsat Band 10	100 m	2015	https://landsat.gsfc.nasa.gov/	LS-B10
	Landsat Band 11	100 m	2015	https://landsat.gsfc.nasa.gov/	LS-B11
SAR	Sentinel-1 VH	10 m	2015	https://sentinel.esa.int/	VH
	Sentinel-1 VV	10 m	2015	https://sentinel.esa.int/	VV
RS-derived	EVI max	1 km	2015	MODIS/006/MYD13A2	EVI-MAX
	EVI mean	1 km	2015	MODIS/006/MYD13A2	EVI-MN
	EVI min	1 km	2015	MODIS/006/MYD13A2	EVI-MIN
	LST day	1 km	2015	MOD11A2	LST-D
	LST night	1 km	2015	MOD11A2	LST-N
	NBLI	30 m	2015	Landsat	NBLI
	NDBI	30 m	2015	Landsat	NDBI
	NDVI	30 m	2015	Landsat	NDVI
	UI	30 m	2015	Landsat	UI
	Night time light	1 km	2015	VIIRS	VIIRS-NL

Others	Accessibility	1 km	2015	Weiss et al. (2018)	ACCESS
	Aspect	~250 m	2010	GMTED2010	ASPECT
	DEM	~250 m	2010	GMTED2010	DEM
	GUF	~12 m	~2012	German Aerospace Center (DLR)	GUF
	Highways	Vector	~2015	Meijer et al. (2018)	RD-1
	Primary roads	Vector	~2015	Meijer et al. (2018)	RD-2
	Secondary roads	Vector	~2015	Meijer et al. (2018)	RD-3
	Tertiary roads	Vector	~2015	Meijer et al. (2018)	RD-4
	Local roads	Vector	~2015	Meijer et al. (2018)	RD-5
	All (+ unclassified)	Vector	~2015	Meijer et al. (2018)	RD-ALL
	Slope	~250 m	2010	GMTED2010	SLOPE

3.2.2 Spatially explanatory variable

We estimate 3D building structure using spatial data of various types and spatial resolutions. For the selection of these explanatory dataset, we use three different criteria: *First*, the data should be close to the year 2015, to keep temporal consistency. *Second*, each dataset for the three regions should be of the same source, thus allowing cross-region comparison. In practice, this means we used datasets with a global coverage. *Finally*, the data should be based on direct measurements rather than downscaled, to ensure independence. We further group explanatory data into four classes: *optical RS*, *SAR (*Synthetic Aperture Radar*)*, *RS-derived* and *others*.

Optical RS data are represented by all available bands from Landsat 8 for the year 2015, covering Europe and the US under cloud-free conditions. For China, the whole territory is not fully covered, thus we extend our data to the period 2014-2016. As explained in Figure 3.2, for each Landsat band we first computed the median of all cloud-free and shadow-free images for each pixel at the original resolution, to generate the representative values for this period and to exclude extreme values. Consequently, we made a spatial aggregation of corresponding representative values into 1-km cells using a mean function.

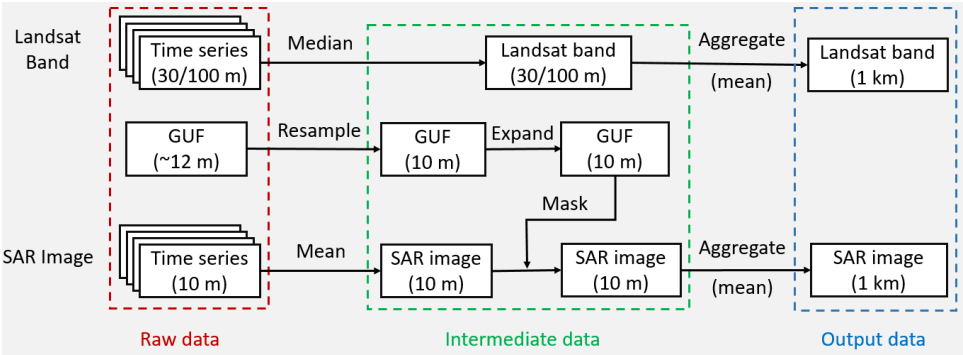


Figure 3.2: Reduction and aggregation of time-series Landsat and Sentinel-1 SAR data. Note: SAR data provided in GEE is log-scaled, we transform the scaled SAR into backscatter coefficient before further operations are applied. Algorithms used to reduce time-series Landsat and SAR collections are suggested by GEE officials (https://developers.google.com/earth-engine/), but in further aggregation operation we additionally filter SAR cell values based on the expanded GUF.

We use 10-meter resolution Sentinel-1 *SAR* images, which have a global coverage for every 12 days (Malenovský et al., 2012). Besides buildings, other objects such as trees are also sensitive to backscatter coefficients (x_{bc}) of SAR. Therefore, we selected SAR

images during two winter seasons around the year 2015, i.e. 1st December, 2014 -31st March, 2015, and 1st December, 2015 -31st March, 2016. However, we added information from adjacent years in areas that were not fully covered by the data from the winters in 2015. Besides, we do not differentiate orbit directions, i.e. ascending or descending, as exploratory data analysis reveals that our case study regions are not fully covered within one single orbit direction. All available SAR images are processed, calibrated, and geo-rectified with the Sentinel-1 Toolbox (ESA, 2019). As illustrated also in Figure 3.2, for each 10-m pixel we firstly average all backscatter coefficients (x_{bc}) available in the study period for VV (vertical transmission and vertical reception) and VH (vertical transmission and horizontal reception) polarization modes separately, and then the averaged x_{bc} for each image is aggregated to a 1×1 km resolution using the mean of only the values within the built environment. For this spatial restriction we use the built environment as mapped by the Global Urban Footprint (GUF) (Esch et al., 2017). We use this as a mask to reduce the influence of objects such as trees and topographic relief outside the built environment. In addition, our exploratory data analysis shows that a large number of buildings (especially the higher ones) are displaced due to the side-looking SAR measurement, thus the GUF mask is buffered with a distance of 2 pixels, i.e., 20 m.

RS-derived data consist of Enhanced Vegetation Index (EVI), land surface temperature (LST), relevant built-up indexes derived from Landsat, and nighttime light intensity (VIIRS). EVI is available for every 16-day period from MODIS product. We process all these data throughout the year 2015 into three variables using maximum, mean and minimum functions separately. LST data are provided by the MOD11A2 V6 product, which is a simple average of all the corresponding MOD11A1 LST cells collected within every 8-day period, and is independently stored as at daytime and nighttime (Wan et al., 2015). We average all LST data throughout the study period for daytime and nighttime respectively. Normalized Difference Built-up Index (NDBI), Normalized Difference Bare Land index (NBLI), Normalized Difference Vegetation Index (NDVI) and Urban Index (UI) are also used as the explanatory variables, which are all derived from Landsat images. To have a systematic understanding of these indices, readers are referred to Mushore et al. (2017). Nighttime light intensity data are derived from stray-light corrected VIIRS nighttime light (Mills et al., 2013), which are provided as monthly composites at 500-m scale. We combined all these monthly data available in GEE for the year 2015 into annual nighttime light intensity using a maximum function, and spatially aggregate them into 1-km data using an average function. The maximum function was used to remove cloud shadow effects in night light images. Given that other light sources (e.g., wild fire and water bodies) reflecting

moonlight or anthropogenic light can appear in non-built-up area, we also applied the GUF mask in order to exclude these areas.

In addition to remote sensing imagery, we use a series of *other data* including the mentioned GUF, accessibility, roads and topography (see Table 3.1). Built-up density is calculated based on the GUF, which is a global binary settlement layer created by the German Aerospace Centre using satellite images from TerraSAR-X and TanDEM-X (Esch et al., 2013). Based on an comparison of estimates for Central Europe, GUF comes out as the most reliable map of urban extent datasets in terms of resolution and accuracy (Klotz et al., 2016). However, it is generated using images during 2011-2013. We assume that other explanatory datasets for the year 2015 and short time interval are sufficient to compensate such defect. Accessibility-to-cities data by Weiss et al. (2018) represent land-based travel time to the nearest densely-populated area for the nominal year 2015. Vector road data from Meijer et al. (2018) are used to generate five hierarchal road density maps including highways, primary roads, secondary roads, tertiary roads, and local roads. In addition, we also add a density map for all roads, which embodies unclassified roads. DEM, slope and aspect are all derived from Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010).

3.2.3 Reference data

Reference data are collected using publicly available datasets from various sources (Table 3.2) for the three case study areas. Specifically, for Europe we use gridded building height data of 25 cities (data supported by European Copernicus Land Monitoring Service), complemented by building footprint layers from OpenStreetMap (OSM, access date: 11 January 2019). To reduce the negative effects caused by null values in building height data, we only consider areas where the proportion of valid buildings exceeds 80% for the footprint area. This threshold is set to exclude locations where a large share of buildings has been built after the gridded building height data have been produced. For the US, we employ data that is publicly available from websites of local governments (including occasional updates published in the ArcGIS Hub (http://hub.arcgis.com/). This data includes building footprints with vertical properties for 27 urban areas. These datasets include areas ranging from megacities like New York and Los Angeles to counties that only include small villages in remote areas. Thus, these datasets include the full variability with respect to the combination of building footprint and building height. Building height data for China, expressed as floor number, are available for 24 selected large cities only. In this paper, for all building height expressed as floor numbers, we assume that

each floor is 3-meter high (Leichtle et al., 2019; Zhou et al., 2014). A relative low model performance was observed for China in our preliminary evaluation, which was ultimately found due to a substantial number of missing buildings in some areas of Chinese cities when compared with VHR satellite images from Google Maps. Therefore, we removed all the data points (i.e., pixels) that were visually incorrect.

Table 3.2: Reference datasets collected for model training and validation.

Datasets	Resolution	Time	Source
Building height for Europe	10 m	2012	https://land.copernicus.eu
Building layers for Europe	Vector	~2018	https://www.openstreetmap.org
Building height for the US	Vector	~2015	See supplementary Table B2
Building height for China	Vector	~2015	https://www.amap.com

Available reference data is biased towards large urban regions. We therefore complement these data with empirical data for smaller settlements, which are classified manually. For this, we use Google APIs to randomly download VHR satellite images outside large urban regions (travel time to cities >10 min, built-up density > 0). Each imagery represents a 1×1 km landscape at a 0.25 m resolution, which we assume is sufficiently detailed for building footprint detection. During the visual interpretation process, a fishnet layer with 50×50 regular squared grids is used for specifying grid numbers, as well as Google Street View for the estimation of building height. These estimations are all based on visual interpretation of VHR satellite images and streetscapes provided by Google Maps. Hence, within all 1×1 km grid cells where built-up land exists as identified by GUF, we randomly selected 1146 images from the US, 2573 images from Europe, and 2445 images from China to complement our data set. Because of the scarcity of street view maps in mainland China, building height is not estimated manually there. We exclude invalid images due to high cloud coverage or image inaccessibility. See supplementary Figure B3 for an example of valid imageries. The methodology for estimating building height is further illustrated in supplementary Figure B4 and Table B2. For locations where no street view map exists, we specify building height by interpreting similar adjacent places where street view maps are available. Finally, for all three case regions we have 55656, 47639, and 47553 1×1 km reference cells for the estimations of building footprint, height, and volume, respectively (Table 3.3), which correspondingly account for 1.17%, 1.00%, and 1.00% of the total 1×1 km cells to be estimated.

Table 3.3: Number of 1×1 km cells in the reference data as well as in the predicted data. Collected data refers to data that was collected from multiple sources except for the interpreted results from Google Maps (see Table 3.2), while Interpreted refers to data that was manually interpreted from Google Maps.

	Predicted data points	**Reference data points**					
		Footprint		Height		Volume	
		Collected	Interpreted	Collected	Interpreted	Collected	Interpreted
Europe	1681014	13728	2469	13731	2466	13731	2466
The US	1447489	30350	1091	24551	1091	24465	1091
China	1632283	5814	2204	5800	0	5800	0
Total	4760786	49892	5764	44082	3557	43996	3557
		55656 (1.17%)		47639 (1.00%)		47553 (1.00%)	

To examine the reliability of our visual interpretation approach, we digitized building footprints based on 100 randomly selected VHR images. Because of the high amount of detail in this VHR imagery relative to the information that was coded, and because this data was collected independently from the RF model, it was found acceptable for generating ground truth data. The comparison shows a very high reliability (see supplementary Figure B1). Abandoned buildings and temporary structures are all included, due to the fact that we are not able to differentiate building types for specific purposes from Google Earth images. As a consequence, the total building footprint area provided here could exceed the actual footprint of 'permanent buildings' or 'under roof' measurements published elsewhere. As shown in studies testing positional accuracy of Google Earth images (Mohammed et al., 2013; Pulighe et al., 2015), error in the horizontal planimetric accuracy (the correct longitudinal and latitudinal placement of a feature on the Earth's surface) is expected to be less than 1.6 m, which we consider sufficiently accurate for our 1-km resolution analysis.

We combine the available reference data with the manually classified data derived from Google maps to obtain the full set of reference data for training the model. Figure 3.3 shows the distribution of reference data as a function of footprint and height, in which only reference data where both footprint and height are valid are shown. Reference data points (i.e., 1×1 km cells) are unevenly distributed within one region, but show complementarity across the three case regions. Specifically, the US has more reference data in medium-footprint and low-height compared to Europe and China, while China has more reference data in medium-footprint and medium-height than the other regions.

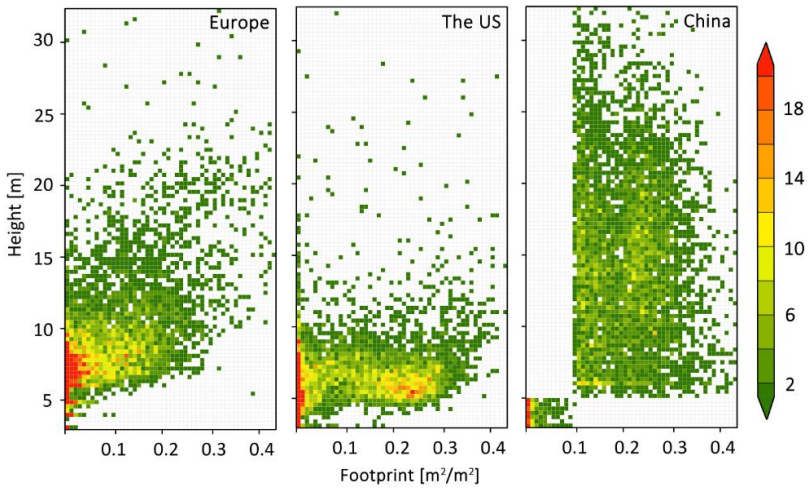


Figure 3.3: Distribution of reference data points (1×1 km cells) for mapping 3D building structure. For China, we manually removed uncertain data points from the reference data. These removals include all locations in the reference data where "footprint < 0.1 & height > 5m" or "footprint > 0.1 & height < 5 m" because our preliminary evaluation of these data showed large inaccuracies.

3.2.4 Model development

We predict three parameters for each pixel: 1) building footprint (m^2/m^2), 2) building height (m), and 3) building volume (m^3/km^2). A 1×1 km mask of settlement is firstly developed based on the absence of built-up land as provided in the GUF. Subsequently, we exclude non-settlement from further operations to save computation resources. Therefore, valid reference data described in Section 3.2.3 account for 1.17%, 1.00%, and 1.00% of the total masked areas for building footprint, building height, and building volume, respectively (also see Table 3.3).

The ensemble regression random forest (RF) approach is used for predicting building footprint, building height, and building volume. This is an efficient prediction method, especially when observations are much scarcer compared to the predictors (Svetnik et al., 2003). The RF model is trained and applied for each of the three variables and for each case region separately, as well as for all regions combined. RF combines several decision trees, built on different combinations of input explanatory variables, and produces the mean prediction of the individual trees. This strategy is beneficial to alleviate the overfitting problem of simple decision trees (Pelletier et al., 2016; Tramontana et al., 2015). The primary property of tree-models is a partitioning of space into smaller regions to manage phenomena characterized by very complex

interactions among variables. In particular, in tree models, partitioning is recursive. The phenomena occur when the subdivisions are divided again until the partitioning reduces the appropriate cost function. Recursive partitioning is terminated when the cost function cannot be further minimized. Hence, a simple model, usable only for the partitioned sub-region, can be estimated. For each observation, the output of a RF model is the average of the outputs of the trees. Therefore, RF models typically yield a reduced bias in the estimations and in general good accuracies (Tramontana et al., 2015). More technical details on the applied RF algorithm can be found in Breiman (2001).

We perform the RF using scikit-learn, a machine-learning package in Python (Pedregosa et al., 2011). To some extent, more trees yield better results. However, the improvement decreases as the number of trees increases, and at a certain point the benefit in prediction performance from including more trees will not be worth the extra computation resources. Therefore, after initial tuning experiments we maximize the number of trees to 150, whereas the minimum number of samples required at a leaf node is fixed to 5. The importance of predictor variables is measured by the Gini decrease in node impurity measure, which is computed by permuting the predictor variables with the out-of-bag data in the RF validation approach (Breiman, 2001).

For each of the three building properties, the reliability of our model is evaluated by a 10-fold cross-validation method as well as an uncertainty analysis. The validation dataset is built by a random selection of 10% of the reference data in these three regions, while the other reference data (90%) are used as training data. This process is repeated 100 times, and for each run we calculate the Pearson's correlation coefficients (R^2), followed by the prediction of corresponding building property based on the explanatory data. Uncertainty can be propagated from the multiple data sources (Heuvelink et al., 1989), while in this study we interpret uncertainty as the stability of our model itself given a certain number of random implementations. Specifically, for the 100 predicted values in each cell, we calculate its coefficient of variation (CV) as the indicator for uncertainty, see Eq. (3.1):

$$\mathrm{CV} = \frac{\sigma}{\mu} \quad (3.1)$$

Where σ and μ respectively refer to the standard deviation and mean value of a corresponding cell for these 100 runs, of which μ is also the predicted value as defined here.

Furthermore, we calculate Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) to assess model performance. We do not merely use the whole reference collection and μ to calculate RMSE, since errors will be under-reported due to the overlap of training and test collections throughout the whole process. For each run, the division of collections for training and validation is random in nature, we therefore calculate RMSE and MAE based on the 100 test collections, see Eq. (3.2), Eq. (3.3), and Eq. (3.4).

$$\mathrm{B}_{pred,j} = \frac{\sum_{i=1}^{t} B_{pred,j,i}}{t} \qquad (3.2)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{j=1}^{s} (B_{pred,j} - B_{test,j})^2}{s}} \qquad (3.3)$$

$$\mathrm{MAE} = \frac{\sum_{j=1}^{s} |B_{pred,j} - B_{test,j}|}{s} \qquad (3.4)$$

Where $B_{pred,j,i}$ refers to the predicted value of endmember j in the ith model, and t refers to the total number of test collections that the endmember j is included. $B_{test,j}$ is the reference value of endmember j, and s is the total number of unique endmembers in the 100 test collections.

Finally, we examine variable importance of the best-fitted runs as identified by the R^2 values. The core principle of variable importance is to calculate the degradation of model performance if such variable is permuted randomly while keeping other input variables constant, which allows for evaluating the relevance of one variable for model output (Zhao et al., 2019).

3.2.5 Analysis of the building structure

We characterize the 3D building structure for three case study regions, and compare the results to analyse structural differences. To do so, we first calculate the average as well as the distribution of all three variables in all three regions. Subsequently, for each region we randomly selected 100,000 1×1 km grid cells for which results are estimated, and make histograms to present the distribution of the building footprint, height, and volume per case region. These distributions are subsequently compared across regions. Moreover, we analyse the correlation coefficients between the building properties in the case regions as well as the combined region. Finally, we analyse regional and sub-regional differences in building structure. Specifically, we calculate building footprint per capita, average building height, and building volume per capita

for each region. In addition, we calculate these variables for each European country, American state, and Chinese province using population data from multiple sources (Table 3.4).

Table 3.4: Population data collected for the analysis of building occupation per capita across case study regions.

Region	Time	Source
Europe	2015	https://population.un.org/
The US	2015	https://www.census.gov/
Mainland China	2015	http://www.stats.gov.cn/
Hong Kong and Macao	2015	https://www.worldbank.org/

3.3 Results

3.3.1 Characterization of 3D building structure

The distribution of building footprint, height and volume is presented in Figure 3.4. In general, the spatial patterns correspond to each other: high values for all three variables are, as expected, concentrated around the larger urban areas of the three regions, such as Paris, New York and Shanghai. Yet, there are notable differences across the three study regions, which are visible from the distribution of all values per continent on the right graphs in Figure 3.4. For example, China has more pixels with a relatively large building footprint as well as a high building height, while the US has more pixels with a low building height, typical for suburban sprawl. Specifically, China has the highest average building height at 10.35 m. In Europe, the average building height is 7.37 m, and in the US this is 6.69 m. Consistently, China has more areas that have a very high building volume, while the opposite is true for the US.

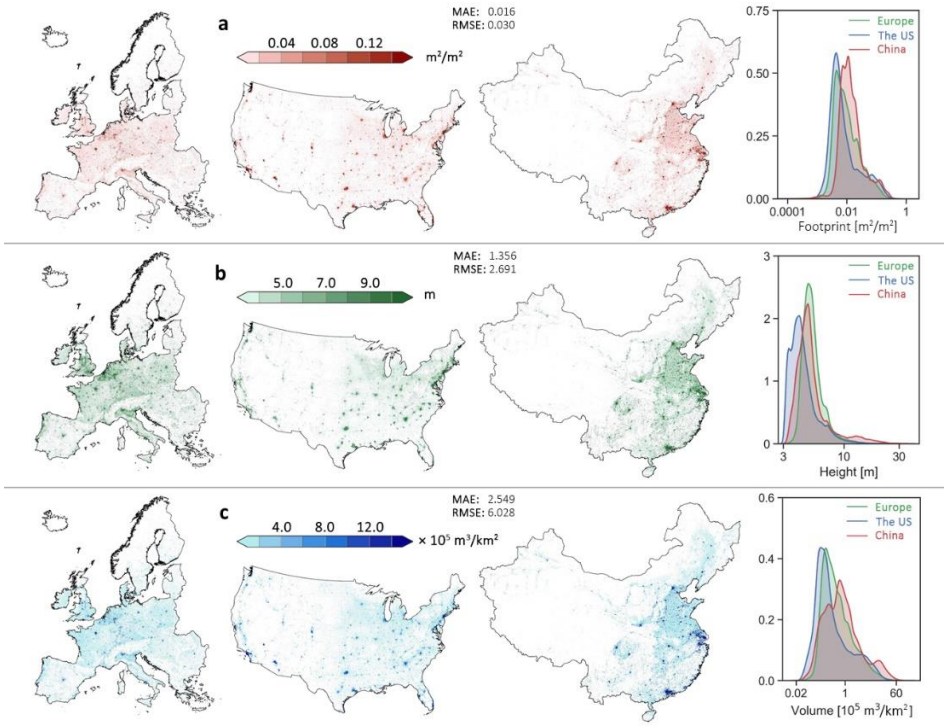


Figure 3.4: Distribution of building footprint, height and volume in three study regions. (a) building footprint; (b) building height; (c) building volume. For each of the three building properties, error statistics are shown in the upper-middle. The graphs on the right, i.e. the kernel density estimations, are plotted based on 100,000 randomly selected points for each region, of which the *x*-axis is scaled using a logarithmic function. The area under the curves is normalized to 1 to facilitate the comparison of distributions across continents if using the logarithm transformed value of x-axis.

A more detailed inspection of 3D building structure highlights the different spatial configurations of buildings in different regions (Figure 3.5). For example, building footprints in the Chinese agricultural plain (Figure 3.5c) are rather dense, as compared to Europe and the US. A large urban footprint is often associated with high-rise buildings, especially for China. Yet, this seems not appropriate for many locations in the US, as is illustrated in the area encircled in Figure 3.5b. Conversely, we also find some areas with a relatively sparse footprint value and a large height value (e.g., in Europe in the encircled area in Figure 3.5a). The detailed results in Figure 3.5c also highlight a particular phenomenon in China, where buildings tend to be taller along main roads that connect large cities, much more than these in Europe and the US.

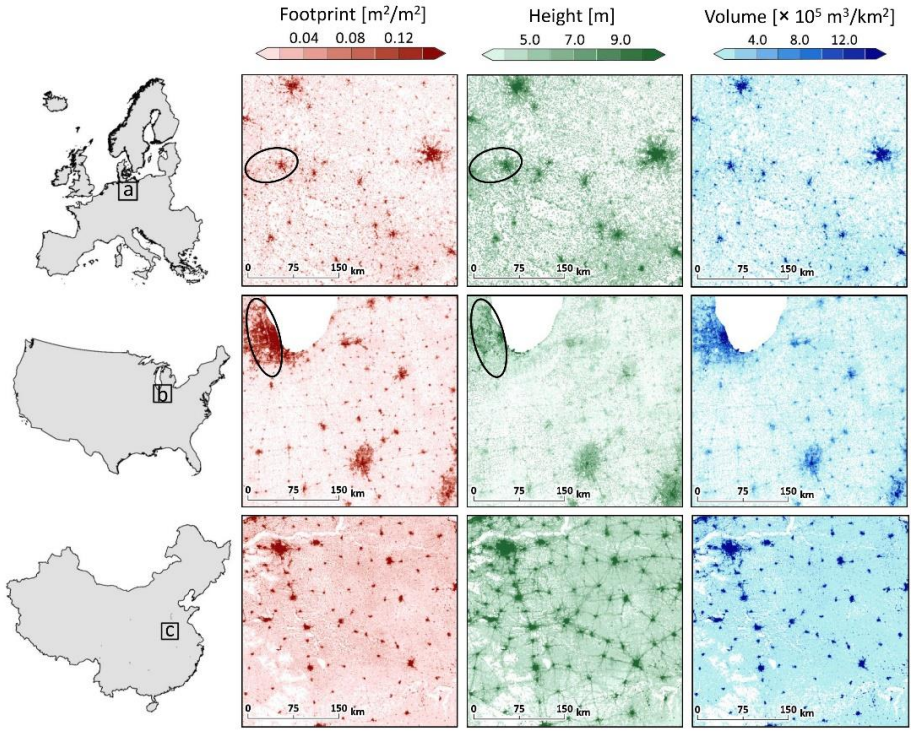


Figure 3.5: Settlement pattern in three densely populated areas embodied in large agricultural plains. (a) around Berlin; (b) around Chicago; (c) around Zhengzhou of the Henan province. Encircles areas indicate regions with smaller footprints but higher buildings (around Hannover), and higher footprint but lower buildings (around Chicago). These three snapshots of densely populated agricultural areas are shown, because they reflect typical settlement patterns that are dominantly shaped by human activities, rather than natural or biophysical constraints such as topography.

Further quantitive analysis shows that building footprint, height, and volume are correlated, but that this correlation is well below 1 (Figure 3.6). This indicates the need to analyse the three different properties independently. The correlation coefficient between footprint and height ranges from from 0.55 in the US to 0.74 in Europe. The correlation coefficients between volume and height as well as volume and footprint are higher ranging from 0.69 in the US to 0.93 in Europe. It is not unexpected that the correlation between footprint and height is lower than the other two correlation coefficients, as volume is by definition the product of footprint and height, and thus at least partially related. Nonetheless, all three properties are estimated independently in this study, and therefore this is not trivial from the setup of the study.

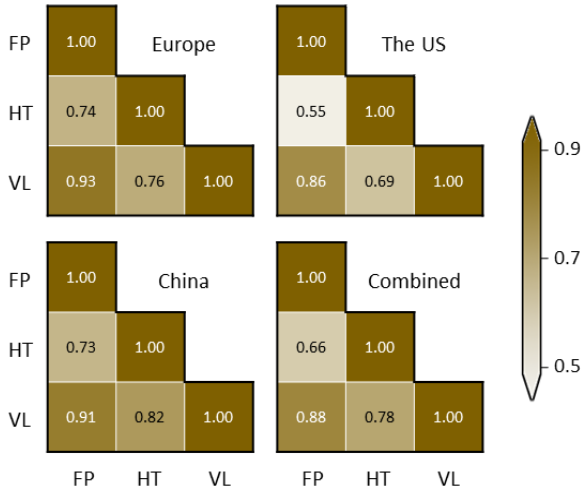


Figure 3.6: Correlation coefficients for building footprint, height and volume, which are referred to as FP, HT and VL, respectively.

The average building footprint per capita is only 29.2 m^2 in China, which is about one third of that in the US (84.5 m^2), and about a half of that in Europe (54.9 m^2). Building volume per capita in China is 302.3 m^3, which is 565.4 m^3 for the US, and 404.6 m^3 for Europe. These results indicate that settlements in the US have a higher land take per person as well as a higher space consumption per person, in comparison to the other regions.

The spatial distribution of 3D building structure also differs between sub-regions (Figure 3.7). For example, the values for building footprint per capita vary much more across US states than across European countries and Chinese provinces, and especially high values are observed in several predominantly rural states such as North/South Dakota, Wyoming, Iowa, and Montana (Figure 3.7a). The high values in these areas could be related to the abundance of agricultural buildings such as barns for livestock (Harun and Ogneva-Himmelberger, 2013), officially termed as '*concentrated animal feeding operation*' in the US. At the same time, in Sichuan and Guizhou, two rural sub-regions of China, building footprint per capita is lower than most other equally-developed sub-regions (Figure 3.7a). Building height, on the other hand, varies most across Chinese provinces and much less across EU countries and US states (Figure 3.7b). Buildings tend to be lower in the inland rural states of the US. Conversely, buildings are much higher in developed sub-regions of China, most of which are coastal sub-regions. The distribution of building volume per person is mostly consistent with the distribution of building footprint per person, with relatively high variation in the US and relatively low variation in Europe and China

(Figure 3.7c). In the US, sub-regions of which have large building volume per capita are mostly located in rural inland states with high values in building footprint per capita, despite their moderate height. In China, regions of which have large building volume per capita are mostly located in urbanized coastal sub-regions such as Jiangsu and Zhejiang, characterized by high buildings but not necessarily by a large building footprint per capita.

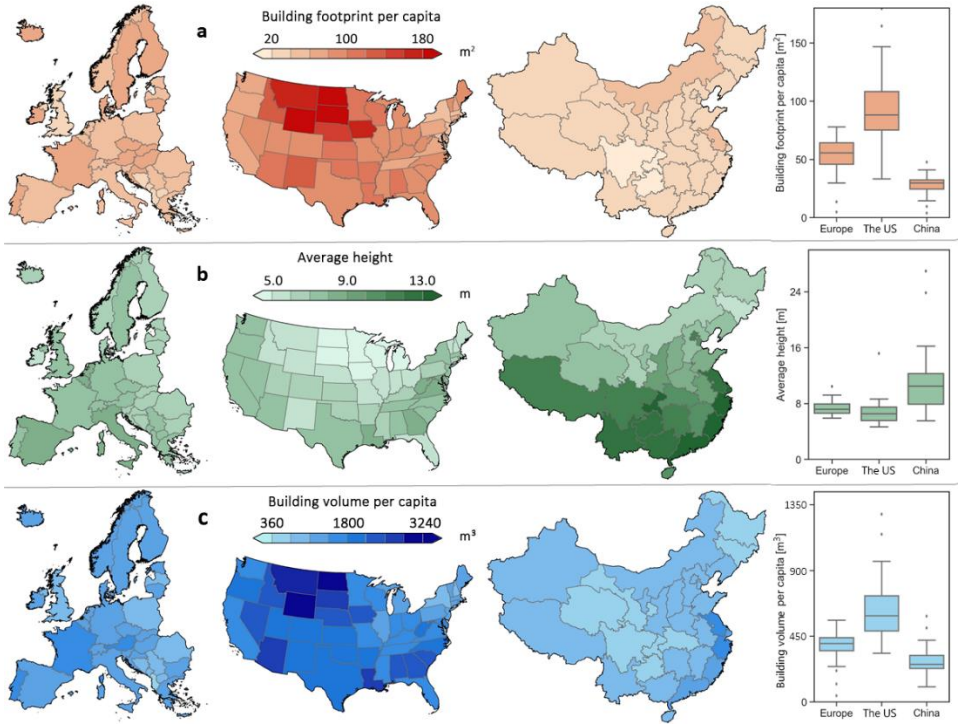


Figure 3.7: Analysis of 3D building structure at sub-regional scale. The boxplots on the right are plotted based on all sub-regions for each study area, of which y-axes are capped to enhance interpretation. The boxes represent the interquartile ranges (25%–75%) and the lines the ensemble-median values.

3.3.2 Model performance and uncertainty

The RF models yield high accuracies for building footprint, height, and volume, as indicated by R^2 values for the three regions combined all larger than 0.80, either for the separated models or the combined models (Figure 3.8). When models are run for each case region separately, building footprint is most accurately predicted for the US. As for the building volume, results for Europe and the US are more accurate than for China. When models are run for all case regions combined, there is no significant

improvement compared with the separated models. Among the three properties, building footprint is most accurately predicted, especially for the US. As for the building volume, results for Europe and the US are more accurate than for China.

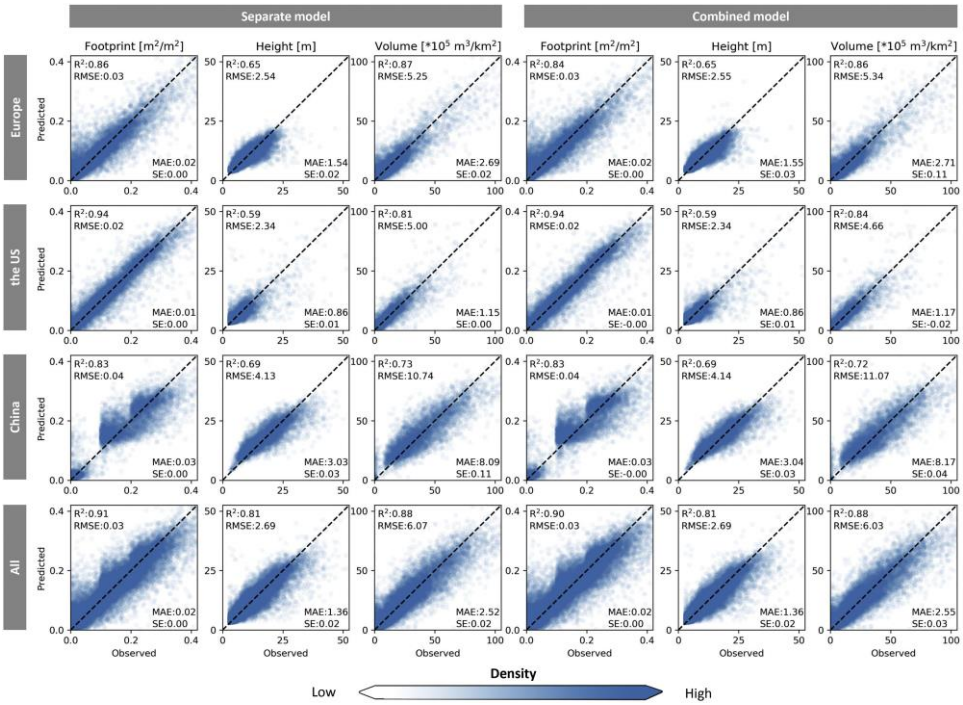


Figure 3.8: Scatterplots of the observed values and predicted values for building footprint, height, and volume. Predicted values represent the mean of all independently predicted values for each location in the reference data.

To further characterize the accuracy of our estimates, we assessed the RMSE, MAE, and SE, based on the independent validation data for each model. It should be noted that the training data have on average higher values of building footprint, height, and volume, thus also leading to higher values for RMSE, MAE, and SE than can be expected for the complete estimated data set. For the combined models, the RMSE values of building footprint, height, and volume for the three regions combined is 0.03 m^2/m^2, 2.69 m, and 6.03 × 10^5 m^3/km^2, respectively. Correspondingly, MAE values of the three building properties are only 0.02 m^2/m^2, 1.36 m, and 2.55 × 10^5 m^3/km^2. While SE values for the three building properties are all close to 0, suggesting that there is no lateral overestimation or underestimation in general.

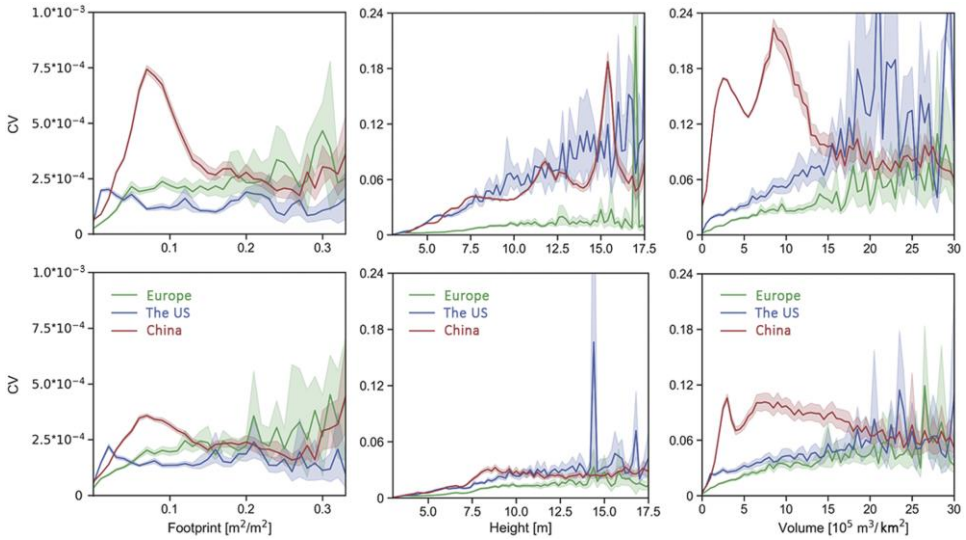


Figure 3.9: Uncertainty, as expressed in the CV values of three predicted properties as a function of their values for these corresponding building properties. Upper row: model ran separately in each of the three regions; bottom row: model ran for all three regions combined. All these figures are plotted based on 100,000 randomly selected points within each region. The shaded areas represent the interquartile ranges (25%–75%) and the lines the ensemble-median values.

The accuracy of separate models and one combined model for all regions is somehow comparable, but combining reference data for all case regions into one model yields a decrease in uncertainty, relative to models trained on one region only (Figure 3.9). Therefore, analyses in the above section are based on the results generated by the "combined" model. Specifically, when trained with data from all regions together, the model for building footprint shows a large decrease in uncertainty in areas with a low building footprint (<0.1 m^2/m^2), which accounts for a large proportion of the study area. Uncertainty of predicted building height shows a decreasing trend over a wider range of values compared to building footprint, especially for the US and China. Moreover, uncertainty is distributed unevenly over different combinations of building footprint and building height (Figure 3.10). Notably, the uncertainty in building footprint was found mostly in areas with either a rather sparse footprint (around 0.04–0.1 m^2/m^2), or at a rather dense building footprint (around 0.3 m^2/m^2). Overall, uncertainty of building height is largest in areas with small values for building footprint and high values for building height. In particular, we find some scattered pixels with large uncertainty in some mountainous areas of southern China, which

can potentially be explained by interference from other landscape elements, such as trees and rocks. We also find that building height is prone to large uncertainty in areas that are not covered by SAR data, for example, a diamond-shaped area in Sichuan province of China, and a small square area in Milwaukee city of the US. The largest uncertainty in building volume is found in areas with a low building footprint and a medium-high building height, as well as in some locations with a high building footprint. Possibly, this uncertainty is also explained by disturbance of other objects, especially in places with a low building footprint.

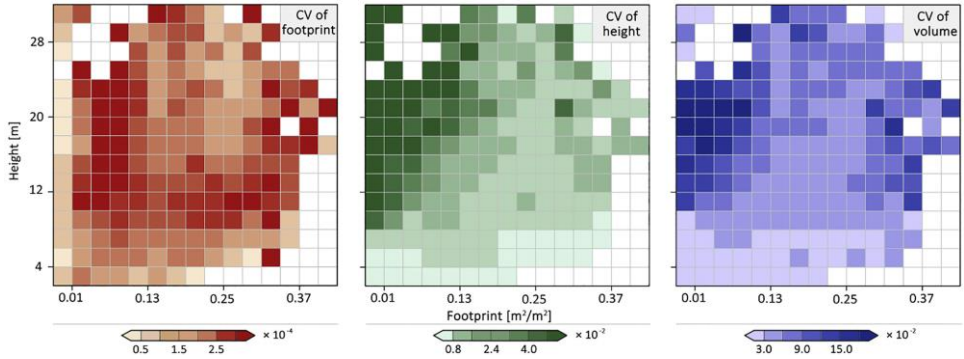


Figure 3.10: Coefficient of variation (CV) plotted as a function of the combination of building footprint and building height, for the combined reference data set of all three study regions, based on results in a sample of 300,000 pixels in the three regions combined.

Zooming in on individual cities further demonstrates the superior performance of the combined model over the models for separate regions (Figure 3.11). For building footprint there is no visible difference between the separate models and the combined model. However, the separate models generally overestimate building height for Paris, Los Angeles, and Shanghai.

The best performing models for each of the three characteristics of building structure are selected for further analysis of the variable importance. This analysis reveals that built-up density derived from GUF, in general, is most valuable for estimating building footprint, while backscatter intensity of SAR has little influence (Figure 3.12). The opposite is true for building height estimation, as backscatter intensity has the largest importance, while built-up density is of course of little influence. Compared to other variables, both built-up density and backscatter intensity are important to predict building volume. In addition, we find a trend shift of VH/VV variable importance when reference data in the three regions are combined. A further analysis

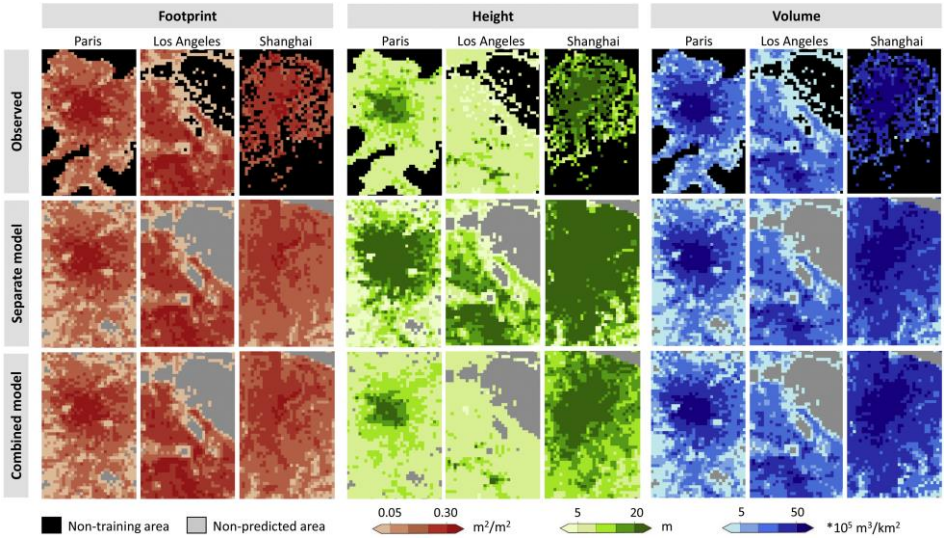


Figure 3.11: Comparison of observed and predicted results for building structure in Paris, Los Angeles, and Shanghai. Each map is 30 × 50 km in size.

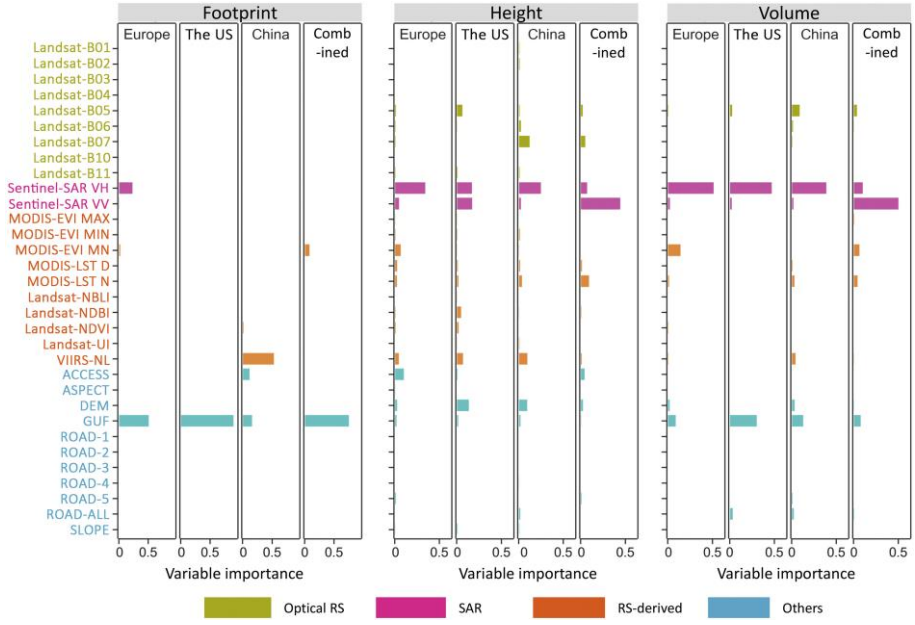


Figure 3.12: Normalized variable importance based on the best-fitted run as identified by their R^2. We run the model separately for Europe, the US, and China, as well as for all regions combined. For each property in one model, the summed variable importance equals 1.

indicates that VH and VV are complements when explaining height and volume (see supplementary Figures B4 and B5).

3.4 Discussion

3.4.1 Mapping 3D building structure at a continental scale

This study shows that the combination of various remote sensing data sets and other spatial data allows deriving the 3D building structure at a continental scale with a high accuracy. The proposed RF models perform better than other models for estimating 3D building structure at large spatial scales, such as the Bayesian Network-based model developed by Paprotny et al. (2020). Hence our models make it possible to map the built landscape in three dimensions and to analyse its geographical characteristics.

We find that buildings in China are the highest on average (10.35 m), compared with the other regions (7.37 m for Europe, and 6.69 m for the US). At the same time both building footprint per capita and building volume per capita are the smallest for China. These data open up a range of new thematic applications. For example, some studies have indicated that urban land per capita is driven by biophysical and socioeconomic conditions such as terrain characteristics, wealth, price of gasoline, and planning strictness (Angel et al., 2011; Taubenböck et al., 2018a). Yet, these analyses have mainly focused on the two-dimensional urban footprint, while the drivers of 3D building structure remain to be explored in more detail. Similarly, differences in building structure could also lead to different energy consumption typologies (Kennedy et al., 2015; Zambon et al., 2019), for which this dataset provides a continental-scale source for further analysis.

Our 3D building structure datasets can facilitate the classification of different settlement types (e.g. suburbs, slums, and business districts) based on *a priori* knowledge of these settlement types (Taubenböck et al., 2018a), to further investigate social or environmental impacts of urban areas. By accounting for urban vertical dimension, such data could also for example improve the risk analysis of urban infrastructure to natural hazards, as this risk depends directly in the building stock, which is increasingly important in the context of climate change (Du et al., 2018; Koks et al., 2019). Another application area is the impact of urban form on environmental conditions (Seto and Shepherd, 2009). To what extent urban climate is affected by building form and their mechanisms remain unclear, as conclusions vary across cases (Manoli et al., 2019; Yue et al., 2019; Zhou et al., 2017). However, most of these

studies still focus on urban configuration in 2D dimensions such as city size and urban centricity. Stewart and Oke (2012) proposed a framework for urban climate research and standardized the worldwide exchange of urban temperature observations. In their framework, settlement systems are generically classified based on configurations of 3D buildings and related land cover types. Yet, large scale analyses based on this framework are constrained to the scarcity of 3D building information. The 3D building structure data set produced in this study could fill in this gap.

Mapping 3D building structure at large spatial scales could also benefit from the accelerated developments of Artificial Intelligence (AI), which increasingly serves as a powerful tool for addressing complex problems (LeCun et al., 2015; Reichstein et al., 2019). However, one of the most essential and challenging parts of AI is that it needs to be trained through large amounts of precisely labelled benchmark data. Currently, there are several databases available for universal objects such as the well-known ImageNet (Deng et al., 2009). Increasingly, there are some urban thematic benchmark databases such as DeepGlobe (Demir et al., 2018), BigEarthNet (Sumbul et al., 2019), and SEN12MS (Schmitt et al., 2019). Yet, these datasets focus mostly on the identification of objects, whereas they do not provide sufficient information on building height and volume. Therefore, we additionally developed a large amount of new reference data for this study specifically. In parallel, computer vision research has made great progress in detecting changes based on digital imagery (Kuehne et al., 2011; Soomro et al., 2012). These developments could greatly benefit urban scientists in characterizing changes in building structure based on time-series satellite data. Yet, as several of the data that feed into our analyses are only available for recent years, notable Sentinel-1 SAR data, change analysis of 3D building structure remains to be challenging.

3.4.2 3D building structure as an indicator for urban land use intensity

Urban land plays an increasingly important role in the global competition for land, and impacts of urban expansion have been widely reported in scientific literature. For example, urban expansion on a global scale leads to the displacement of cropland and subsequent losses in natural areas (van Vliet, 2019). Consequently, increasing urban land use intensity could be a way to reduce urban expansion and thus alleviate the global competition for land.

Population density has been used frequently for analysing urban land use intensity. However, population density maps are mostly produced by using a downscaling approach, based on a combination of census data and spatial data, such as nighttime

light and built-up area (Florczyk et al., 2019a; Wang et al., 2018). Therefore, while population data is typically rather accurate at the census level, they remain rather uncertain at the local/pixel level. Moreover, population density reflects residential activities only, while other urban activities remain unaddressed (Dovey and Pafka, 2013). A few studies have investigated population distribution based on building volume, but at a local scale (Dong et al., 2010; Tomás et al., 2015; Zhao et al., 2017), the large-scale building structure data produced in this study can therefore move population estimation forward.

Building characteristics as presented in this paper offer an alternative for characterizing urban land use intensity. Previously, such information has already been presented for selected global megacities (e.g. Bagan and Yamagata (2012), Mertes et al. (2015), and Taubenböck et al. (2012)). Yet, a large proportion of the built-up area, and thus of urban activities, is located outsides these megacities (Li et al., 2019). The approach presented in this paper therefore complements population density as a measure for urban land use intensity, and covers all types of human settlements regardless of their size.

The comparison between building footprint and height shows that they are only partly correlated (0.55 for all three case regions combined). In other words, there is a considerable amount of variation in building height within locations with a comparable building footprint, thus justifying the mapping of these properties separately. This also implies that the analysis of 2D urban density as a proxy for urban intensity hides a significant part of the variation in actual building structure. Local patterns in the relations between building footprint and building height also differ across the three regions. The particular phenomenon in China, i.e. buildings along main roads tend to be higher, suggests that local conditions largely affect building structure. However, evidence for these differences as well as explanations for their causes are still sparse in literature. For example, this particular phenomenon in China could be attributed to the mobility requirements of population (Wang et al., 2016), which facilitates the development of retail and service industries, resulting in higher buildings for mixed uses along main roads. Yet, this push-pull theory behind the spatial differences in 3D building structure is rather anecdotal.

Our study also reveals the potential to guide settlement development towards sustainable land use patterns for the benefit of human well-being. In the sustainability community, consensus has not been reached on whether urbanization is part of the problem or a solution to sustainability challenges (McFarlane, 2019; Seto et al., 2010). Either way, urban densification, both horizontally and vertically, is acknowledged as

one of the tangible solutions to satisfy the increased urban land demand while conserve other land (Wang et al., 2019b). However, we also notice that local settlement trajectories should be guided in a large-scale context with broad considerations, including quality of live for inhabitants of human settlements, while these trade-offs and synergies remain largely unexplored.

3.5 Conclusion

This study presents the first consistent continental-scale dataset on 3D building structure for Europe, the US and China. The presented data was generated using random forest (RF) models fed with optical remote sensing imagery, SAR imagery, remote sensing derived indices, and spatial socio-economic data. The RF models yield R^2 values of 0.90, 0.82, and 0.88 for building footprint, height, and volume, respectively, for all three continents combined. Our results show that building height is to a large extent independent from building footprint, emphasizing the importance of mapping these properties independently. The average building footprint per capita is only 29.2 m^2 in China, which is about one third of that in the US (84.5 m^2), and about a half of that in Europe (54.9 m^2). Building volume per capita in China is 302.3 m^3, which is 565.4 m^3 for the US, and 404.6 m^3 for Europe.

The 3D building structure data produced in this study provide a nuanced representation of settlement pattern, which can be used for urban environment analysis, spatial planning and land use modelling that aim to guide settlement development towards sustainability. At the same time, it already now reveals geographic peculiarities across different regions in the globe.

Data availability

All datasets used in our analysis, as well as the codes for model algorithm and statistical visualization are also available (https://cscproject.github.io).

Acknowledgements

Mengmeng Li would like to acknowledge funding from China Scholarship Council (Grant No 201706510011). Jasper van Vliet was further supported by the Netherlands Organization for Scientific Research NWO (Grant No VI.Vidi.198.008). All the

authors thank the four anonymous reviewers for their insightful comments and critical reviews, which helped to improve our manuscript. This research contributes to the Global Land Programme (www.glp.earth).


Chapter 4

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

Horizontal and vertical patterns of built-up land are essential to analyze 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^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^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^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^2 in both Oceania and North America, while it is only 29 m^2 in South Asia and 38 m^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.

This chapter is published as:

Li, M., Wang, Y, Roiser, J.F., Verburg, P. H., & van Vliet, J (2022). Global maps of 3D built-up patterns for urban morphological analysis. *International Journal of Applied Earth Observation and Geoinformation*. 114, 103048. DOI: 10.1016/j.jag.2022.103048

4.1 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., 2012a). 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 building 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 Pafka, 2013; Li et al., 2022).

Large-scale gridded datasets of 3D built-up patterns are derived using both indirect measurements and estimation approaches (Esch et al., 2022; Frantz et al., 2021; Huang et al., 2022; Lao et al., 2021; Liasis 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 near 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; Frantz et al., 2021; Li et al., 2020). The advantage of this approach is the relatively small computation 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 building structure 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 3D 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 79186, 71163, and 71079 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.

4.2 Methods and data

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 71079 and 79186 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 Figure 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.

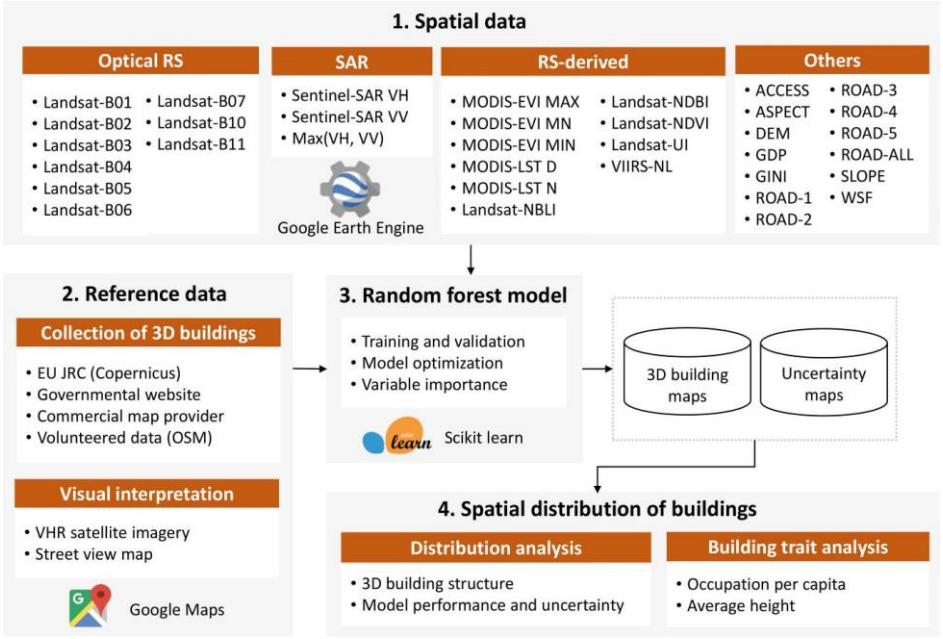


Figure 4.1: An overview of the methodology applied and data sources included in this study. Please refer to supplementary Table C1 for the abbreviation of spatial data.

4.2.1 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 downscaled, in order to ensure independency of these variables. These three criteria yielded 35 datasets (see supplementary Table C1), 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-km^2 resolution using a mean function (see supplementary Figure C1).

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., 2020b). 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 displaced 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 supplementary Figure C1 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-km^2 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 use 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. LST is derived from the MOD11A2 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 NDBI, NBLI, NDVI, and UI as explanatory variables. The variable NTL was derived from VIIRS 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-km^2 grids using a mean function (also see supplementary Figure C1 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, Gennaioli et al. (2013), and (Solt, 2020)) was included. These variables were not included in Li et al (2020) in which this analysis built. Yet, Specifically, 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 large on a global scale than for the countries included in Li et al (2020). We did not use the downscaled version (Kummu et al., 2018), as downscaling is based on other datasets similar to those used here, 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.

4.2.2 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.

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 (http://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 meter high (Leichtle et al., 2019). Our preliminary evaluation suggests 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 under-represented buildings on the periphery of cities. Therefore, we manually removed the 1-km² grids that contain 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 10000 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 71079 to 79186 for building footprint, building height, and building volume, distributed over all continents except Antarctica (Figure 4.2, and supplementary Table C2).

4.2.3 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.

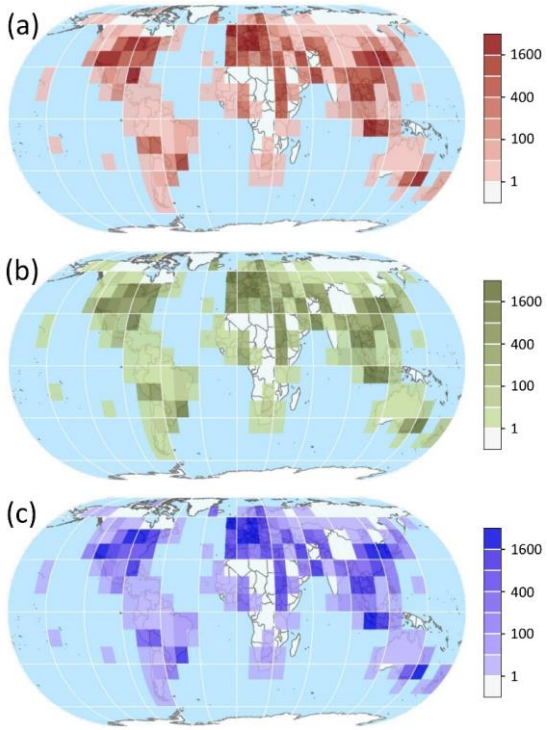


Figure 4.2: Distribution of reference data over 10° × 10° tiles. Numbers refer to the number of 1-km² pixels for which reference data was included in our model. (a) building footprint; (b) building height; and (c) building volume.

In general, tree number in the RF model is positively associated with model performance. Yet, the improvement becomes less prominent as the tree number increases. Therefore, after primarily experiments, we set the tree number as 150, and set the minimum samples 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 $\geq 0.5\%$.

Consequently, 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 Figure 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.

4.2.4 Analysing distribution of building footprint, height, and volume

Following the Eq. (4.1) – (4.3), we calculate the global sum of building footprint (F_{sum}), average of building height (H_{ave}), and global sum of building volume (V_{sum}), respectively.

$$\mathrm{F_{sum}} = \sum_{i=1}^{\mathrm{N}} \mathrm{f}_i \qquad (4.1)$$

$$\mathrm{H_{ave}} = \frac{\sum_{i=1}^{\mathrm{N}} (\mathrm{f}_i \times \mathrm{h}_i)}{\sum_{i=1}^{\mathrm{N}} \mathrm{f}_i} \qquad (4.2)$$

$$\mathrm{V_{sum}} = \sum_{i=1}^{\mathrm{N}} \mathrm{v}_i \qquad (4.3)$$

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.

Consequently, we analysed 3D built-up patterns along urban-rural gradient in selected larger cities, as well as 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 (Florczyk et al., 2019b). 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 (Figure C2). 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.

4.3 Result

4.3.1 Global distribution of 3D built-up patterns

Our results show a total building footprint area of 264 thousand km², a total building volume of 1645 km³, 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.

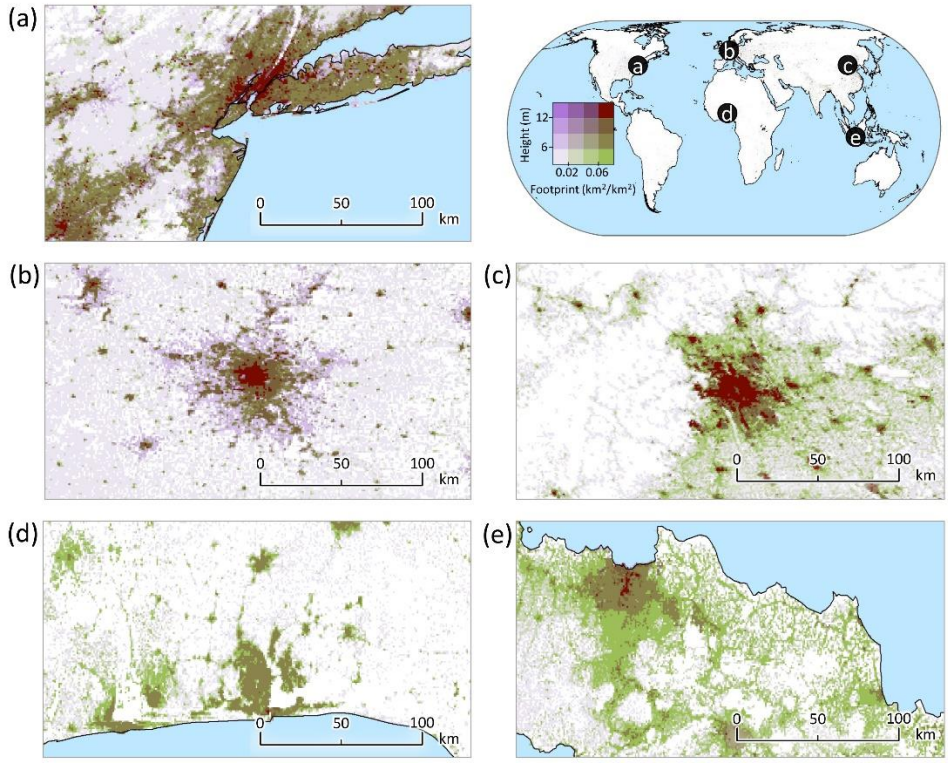


Figure 4.3: Maps of built-up structure at a 1-km² resolution for selected areas. Maps show the differences in building height and building footprint across the globe, with dense and high buildings in New York (a) and Paris (b), mixed patterns in Beijing (c), and relatively dense but low urban in Lagos (d) and Java (e).

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 Figure 4.3, 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 Southern 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 partly attribute to the sufficient 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 (Figure 4.4). Yet, exceptions exist, such as Kano City in Nigeria, where buildings in the centre are low as compared to other cities and this height barely decreases with the distance from the centre (Figure 4.4b). 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.

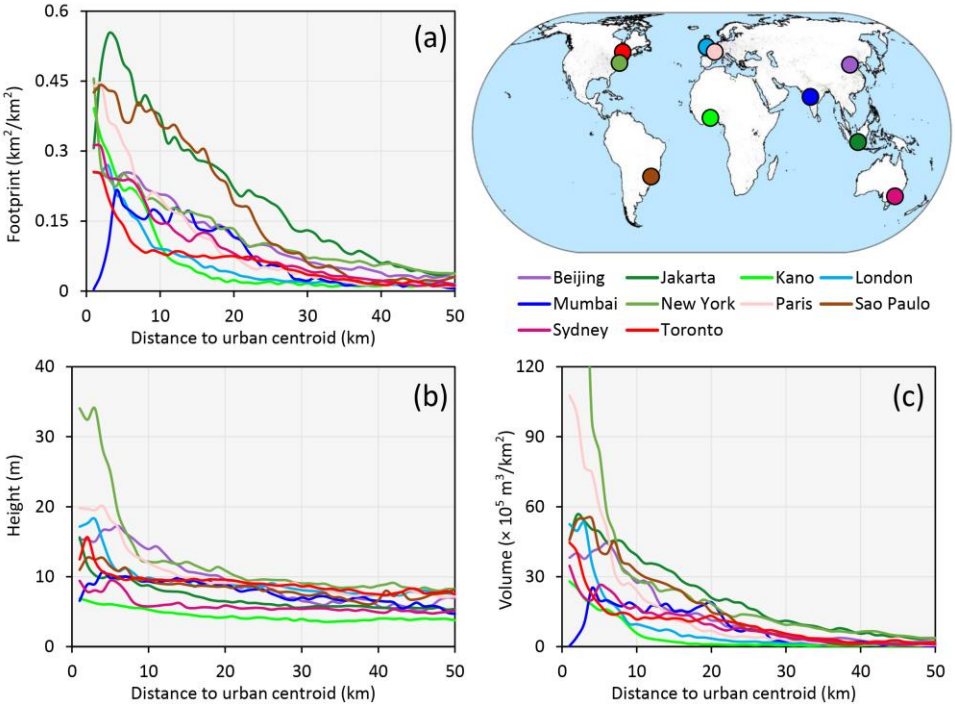


Figure 4.4: Distance decay curves of (a) building footprint, (b) building height, and (c) building volume as a function of the distance from the city centre for 10 large cities across the world.

Of the ten world regions, building footprint per person is the highest in Oceania, followed by Canada and USA (Table 4.1). In contrast, the building footprint per person in Sub-Saharan Africa is only 23 m^2, and in Southern 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 Southern 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.

Table 4.1: Total, average, and per person building footprint, volume, and height for ten major world regions. World regions are shown in supplementary Figure C2.

World region	Total population	Total footprint	Average	Total volume	Footprint per	Volume per	Floor area per
Canada and USA	0.36	36	5.97	217	100	605	202
China	1.41	46	6.61	330	33	234	78
Europe	0.68	39	7.05	264	58	390	130
South Asia	1.74	31	5.08	151	18	86	29
Latin America	0.63	24	6.28	145	38	229	76
Middle-East and Northern	0.42	13	7.36	93	31	220	73
Oceania	0.03	37	5.12	20	122	682	227
Russia and Central Asia	0.23	13	6.30	87	58	384	128
Southeast Asia	0.84	35	6.15	221	41	262	87
Sub-Saharan Africa	1.00	23	4.77	116	23	115	38
World	**7.35**	**264**	**6.16**	**1645**	**36**	**224**	**75**

Although sparse and low-rise buildings dominate rural areas globally, we see different patterns across the ten world regions (Figure 4.5). 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 Southern Asia and Southeast Asia, whereas the phenomena in Sub-Saharan Africa and Europe co-exist.

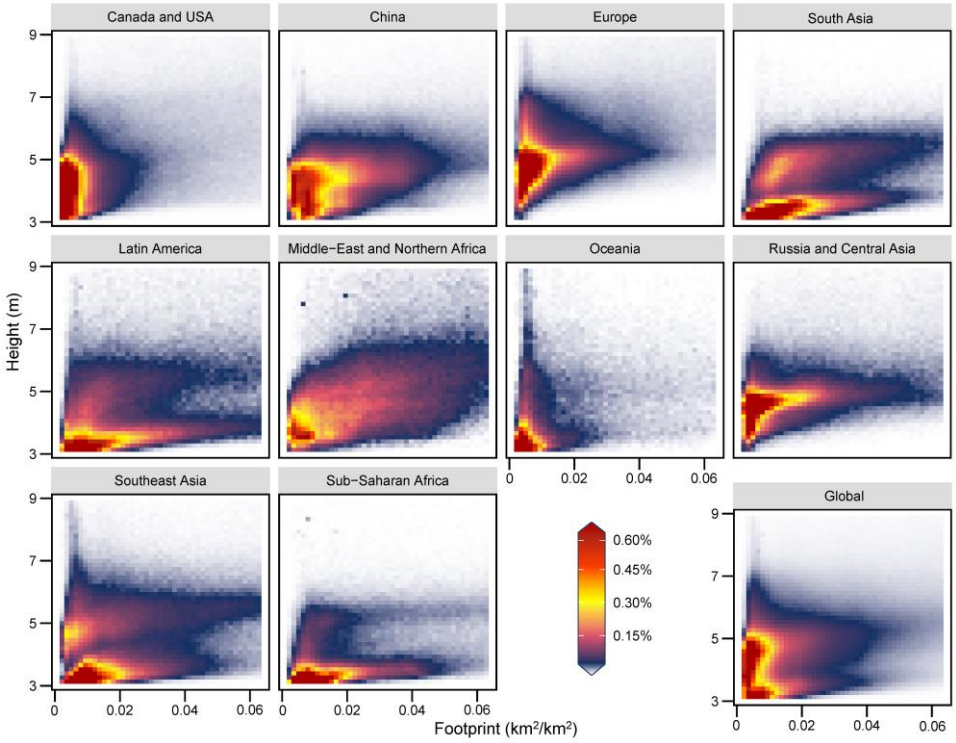


Figure 4.5: Relative frequency of the 1-km² pixels as the joint occurrence of footprint and height for ten world regions as well as for the globe. The delineation of world regions is shown in supplementary Figure C2.

4.3.2 Model performance and uncertainty

The exclusion of variables variable importance values $\leqslant$ 0.5% process led to 22, 31, and 25 variables for estimating building footprint, height, and volume, respectively (supplementary Table C3). 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 (supplementary Figure C3).

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 (Figure 4.6, and supplementary Figures C4 and C5).

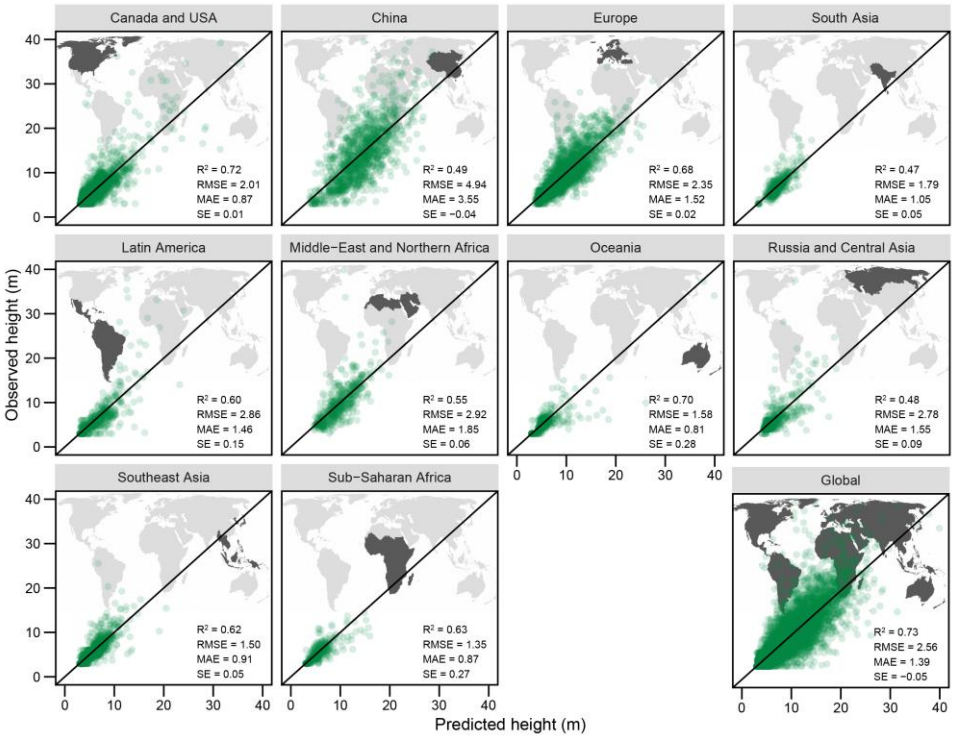


Figure 4.6: Scatterplots the predicted values v.s. observed values for building height in the independent 20% test subset. Predicted values represent the mean of 100 predicted values for each location in the test subset.

Overall, model uncertainty reflected by CV values for footprint shows an inverse U-shape, whereas model uncertainties for height and volume constantly increase (Figure 4.7). 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.

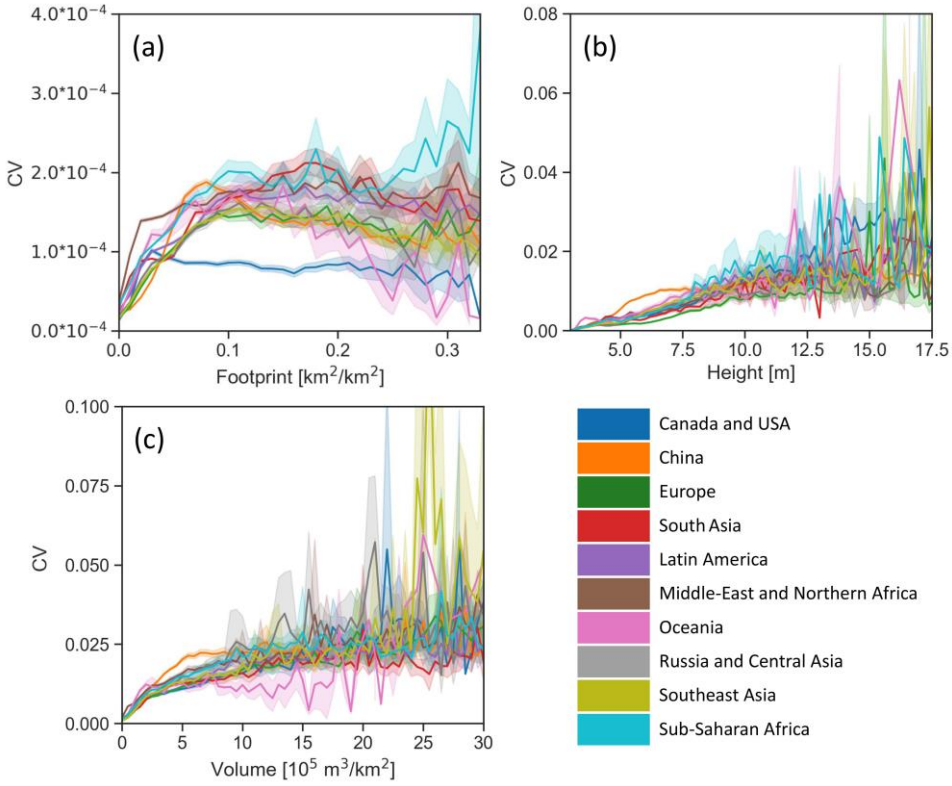


Figure 4.7: Distribution of model uncertainty, expressed in the CV values of three predicted properties, as a function of their values for these corresponding building properties. (a) Building footprint; (b) Building height; and (c) Building volume. Each of the three graphs is generated using a random subset of one million points across the world.

4.4 Discussion

The global maps of 3D built-up patterns provide an image of the heterogeneity within built-up areas, thereby complementing global datasets of built-up land (Gong et al., 2020b; Schneider et al., 2010; Zhang et al., 2022) and gridded population density (Leyk et al., 2019). 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 (supplementary Table C4) 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 these maps, as errors within a single pixel cancel out. However, this is not the case for systematic errors (SE). The SE of -0.00 km^2/km^2, -0.05 m, and -0.03×10^5 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 for 2013. 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. Whilst WSF-2015 is proved to be one of the most reliable global built-up layer (Marconcini et al., 2020), it still can 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 critical input for future studies on urban sustainability, including analyses of urban form (Taubenböck et al., 2020), exposure to natural hazards (Paprotny 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.

Data and code availability

The generated maps, as well as the input data and algorithms used for training and classification are publicly available at https://landbigdata.github.io. Original data can be accessed using links provided in the methods and data section of this article.

Acknowledgments

We would like to acknowledge Elco Koks and Hannes Taubenböck for the discussion on methodological improvement in earlier stages of the project. This research was jointly supported by China Scholarship Council (201706510011) and Netherlands Organization for Scientific Research NWO (VI.Vidi.198.008).


Chapter 5

5 | Global trends and local variations in land take per person

Globally, urban areas are growing at a faster rate than their population, potentially reducing environmental sustainability due to undesirable land take in (semi)natural and agricultural lands. However, it is unclear to what extent this trend varies locally, which may hamper the formulation and implementation of local-scale policies in the context of global competition for land. Here, we attribute built-up land change to population dynamics and changes in land take per person, for more than 75,000 administrative regions worldwide, typically representing municipalities or counties. Results show that changes in land take per person, expressed as the area of built-up land per capita, relate to 38.3%, 49.6%, and 37.5% of the total increase in built-up land during the periods 1975-1990, 1990-2000, and 2000-2015, respectively, but with large local variations. Interestingly, we find that centers of large cities intensify in all three periods, while their rural areas show an opposite development, suggesting an urban polarization effect. We also find densification in many regions in the Global South that already have a high population density, leading to potential trade-offs in terms of human wellbeing. Therefore, our work provides novel insights into the debate on sustainable urban development at a global scale.

This chapter is published as:

Li, M., Verburg, P.H., & van Vliet, J. (2022). Global trends and local variations in land take per person. *Landscape and Urban Planning,* 218, 104308. DOI: 10.1016/j.landurbplan.2021.104308

5.1 Introduction

Built-up land covers only a small fraction of the global land area (Klotz et al., 2016; Liu et al., 2014c), and it has thus far received little attention in the context of the global competition for land (Lambin and Meyfroidt, 2011). However, in recent years, built-up land is expanding about as fast as agricultural land, suggesting the need to consider this process in more detail at a global level (van Vliet, 2019). Since human settlements are often surrounded by fertile cropland, built-up land expansion has taken place predominantly at the expense of croplands, and this process is expected to continue in the near future (Bren d'Amour et al., 2017; van Vliet et al., 2017). Studies also show that unless we change our lifestyle drastically, increases in human population and especially increases in wealth will stimulate the demand for both food and housing over the next decades (Gao and O'Neill, 2020; Laroche et al., 2020). However, in contrast with agricultural land (Foley et al., 2011; Zabel et al., 2019), the intensity with which built-up land is used, measured by the amount of people per area of built-up land, has decreased continuously in recent decades, and this trend is expected to continue in the coming decades (Angel et al., 2011; Gao and O'Neill, 2020).

The observed increase in land take per person is exhibited in multiple ways, including urban sprawl in the United States (Barrington-Leigh and Millard-Ball, 2015), peri-urbanization in Europe (Shaw et al., 2020), and rural hollowing fuelled by industrialization in China (Fan et al., 2020; Liu et al., 2010; Long et al., 2009). These processes have led to rapid built-up land expansion in recent decades, globally, despite a decrease in population growth (Angel et al., 2011; Seto et al., 2011). Conversely, at smaller scales, there are also some reports on densification in urban centres (Broitman and Koomen, 2020; Chen et al., 2020a). The development of higher density in urban centres and lower density in their surrounding regions can be partially explained by the classic land-rent theories (Alonso, 1964; Park, 2014). These theories explain how areas close to the city centre, assumed to contain the central business district, are more attractive. As a result, population increases will lead to densification in the city centre. Consistently, the same increase in population makes the locations at the edge of a city also sufficiently attractive for built-up land expansion. As land prices are relatively low here, these can be constructed with a low density. This expansion process is further amplified by the decrease in transportation costs, which has been associated with urban sprawl and suburbanization in different locations across the worlds (Anas et al., 1998; Colsaet et al., 2018).

Dynamics in land take per person are represented in SDG target 11.3, which aims at achieving a rate of increase in built-up land that does not exceed the rate of increase in population. To restrict land take, i.e., the conversion of agricultural, forest, and other (semi-)natural lands into built-up area, many local and regional policies are introduced and implemented to promote high-density urban development, often indicated with terms like compact cities and smart cities (Batty et al., 2012; Tappert et al., 2018). *Plan Melbourne 2017-2050*, for instance, highlights the priority to facilitate high-density residential development in the coming decades (Nethercote, 2019). Similarly, a recent review identified six EU-level spatial strategies, which are designated to guide urban development towards compact cities, and ultimately promoting urban sustainability (Cortinovis et al., 2019).

In the context of the global trend that built-up land increases faster than the population, there is a need to better understand built-up land change trajectories, in order to identify leverage points to reduce the global competition for land (Abson et al., 2017). Urban expansion has been investigated for a few large cities and metropolitan areas (e.g. Schneider and Mertes (2014), Seto and Fragkias (2005), and Xu et al. (2019)), as well as on a global scale (e.g. Angel et al. (2011), Gao and O'Neill (2020), Hu et al. (2021), and Xu et al. (2020)). Yet, it is not clear to what extent countries and major world regions follow this global trend, or whether there is any difference between different cities, their peripheries, and relatively rural regions, which could potentially be related to land-rent theories. Moreover, the importance of small and medium-sized settlements are increasingly apparent in the context of global built-up land expansion (Chai and Seto, 2019; Li et al., 2019), suggesting the need for a comprehensive analysis of built-up land dynamics covering the full range of human settlements, i.e., from hinterlands to metropolises.

In this study, we analyse for the first time to what extent built-up land change is related to population change and to what extent it is related to changes in the area of built-up land per capita, for 75102 populated regions worldwide, typically reflecting municipalities or counties. Subsequently, we analyse whether large city centres, small and medium city centres, their affiliated peripheral regions, and other regions follow different trends. Finally, we discuss the implications of these trajectories for sustainable urbanization from a global perspective.

5.2 Materials and methods

5.2.1 Conceptualization

In this study we attribute the expansion of built-up land to changes in population and changes in land take per person for 75102 regions, globally, similar to the conceptualization applied previously at a regional scale (van Vliet et al., 2019). Land take per person can be interpreted as a measure for land-use intensity. Land-use intensity of built-up land, sometimes also referred to as density (we use both terms interchangeably in this paper), has been operationalized in different ways (Dovey and Pafka, 2013; McFarlane, 2016). In this study, we operationalize land-use intensity as population per unit built-up land, which is similar to the interpretation of agricultural land-use intensity expressed as production per unit agricultural land area (Kuemmerle et al., 2013). The inverse of population per unit built-up land, i.e., built-up land area per capita (BPC), is used in our calculations. An increase in BPC thus indicates a decrease in the intensity with which built-up land is used, and vice-versa.

For this study we attribute changes in built-up land to changes in population and changes in BPC, following Eq. (5.1) and (5.2), respectively. These equations ensure that the products of the change in population and the change in BPC are equal to the change in built-up land. For example, an increase in built up land of 10% and an increase in BPC of 20% will lead to an increase in built-up land of 32%, as 110% × 120% = 132%.

$$\mathrm{A_{POP}} = (\mathrm{A_1} - \mathrm{A_0}) \times (\ln \frac{\mathrm{POP_1}}{\mathrm{POP_0}}) / (\ln \frac{\mathrm{A_1}}{\mathrm{A_0}}) \qquad (5.1)$$

$$\mathrm{A_{BPC}} = (\mathrm{A_1} - \mathrm{A_0}) \times (\ln \frac{\mathrm{BPC_1}}{\mathrm{BPC_0}}) / (\ln \frac{\mathrm{A_1}}{\mathrm{A_0}}) \qquad (5.2)$$

Where $\mathrm{A_{POP}}$ is the change of built-up land attributed to population change, and $\mathrm{A_{BPC}}$ is the change of built-up land attributed to BPC change. A, POP, and BPC with numerical subscripts 0 and 1 indicate the built-up area, population, and the area of built-up land per capita, at the beginning and end of a period, respectively. For $\mathrm{A_{POP}}$ and $\mathrm{A_{BPC}}$, negative values denote the area of built-up land that can be potentially saved as a result of population decrease or BPC decrease (i.e., urban densification or urban intensification). Changes in built-up land can also be attributed to changes in population in different ways. Therefore, we test sensitivity of different attributions choices by using three alternatives for Eq. (5.1) and (5.2), as explained in the supplementary method of Appendix D.

Built-up land changes owning to changes in either population or BPC are allocated for each of the 75102 administrative units, and these results are further aggregated for larger-scale representations, in which only net changes in their affiliated regions are accounted when we report percentage of built-up land change due to either population change or BPC change.

5.2.2 Datasets

We use the Global Human Settlement Layer (GHSL, https://ghsl.jrc.ec.europa.eu/) as the source for built-up land as well as population, and combine these to calculate BPC. GHSL is a suite of global gridded data sets of the human presence from 1975 to 2015, which is originally derived from 30-m Landsat satellite image collections. It provides consistent data on both built-up land and population, i.e., GHS-BUILT and GHS-POP (Ehrlich et al., 2018; Schiavina et al., 2019). These GHSL layers are available at various spatial resolutions: approximately 38 m (GHS-BUILT), 250 m (GHS-BUILT and GHS-POP), and 1 km (all the layers in the package), for the years 1975, 1990, 2000, and 2015. In this study, we use the GHS-BUILT and GHS-POP at the 1-km resolution for the three epochs between 1975 and 2015, where the built-up land for the year 2014 is combined with the population data for the year 2015 and used together to represent the situation in 2015.

We included 75102 regions in this analysis, together covering all populated land, globally. The regions represent administrative regions and are taken from the subnational divisions provided in GADM-3.6 (https://gadm.org/). GADM contains multiple levels of subdivision, ranging from level 0 (country) to level 5 (smallest unit). Not all levels are available for all countries, and similar levels in different countries do not always represent comparable spatial units in terms of area or population. We selected subdivisions that typically represent municipalities or counties, because these are the regions at which spatial plans and policies are often implemented and because these regions often correspond with what is perceived a city or a town. Specifically, we used level-2 units as our unit for analysis, unless the average population of level-2 regions in a country exceeded 200000 inhabitants in the year 2015, in which case we used level-3 subdivisions instead. In all cases, when subdivisions at specific levels were not available, we used the smallest available subdivision. We excluded countries/regions of which subdivisions are only available at level-0, because they are predominantly remote islands with little or no population (see supplementary Table D1). Regions where no built-up land/population exists in the year 2015 were also excluded, as these are of little relevance for this study.

5.2.3 Analysis of built-up land change trajectories

For each period, we categorize regions into one of nine trajectories of built-up land change, shown in Figure 5.1. To that effect, we first calculate the built-up land and population for all regions included, and subsequently calculate annual changes for both for the periods 1975-1990, 1990-2000, and 2000-2015. Changes are calculated on an annual basis, to facilitate comparison between periods of different length, and are expressed as a percentage of the start year of each period (see Eq. (5.3) and Eq. (5.4)). Second, the thresholds between change and no change categories for both variables are determined using a histogram-based approach. Specifically, we calculate Annual Growth Rates (AGRs) of population and BPC during the three periods and identify one-third of the total samples for all the three time periods combined, centred on the no-change value. Additionally, to facilitate interpretation, we applied the same threshold for both variables to identify the nine change trajectories. Following this process, we identified this threshold as ±1.00% per year changes in population and changes in BPC. In other words, only increases of more than 1% or less than -1% in these variables are indicated as growth and decline, respectively.

$$\mathrm{AGR_{POP}} = \left(\left(\frac{\mathrm{POP_{end}}}{\mathrm{POP_{start}}}\right)^{\frac{1}{n}} - 1\right) \times 100\% \qquad (5.3)$$

$$\mathrm{AGR_{BPC}} = \left(\left(\frac{\mathrm{BPC_{end}}}{\mathrm{BPC_{start}}}\right)^{\frac{1}{n}} - 1\right) \times 100\% \qquad (5.4)$$

In these equations, $\mathrm{POP_{start}}$ and $\mathrm{BPC_{start}}$ are the population and BPC in the beginning of the time period, $\mathrm{POP_{end}}$ and $\mathrm{BPC_{end}}$ are the population and BPC in the end of the time period, and n is the number of years in between.

For each of the nine trajectories, we summarize population and built-up land area in the starting year, as well as population change and built-up land area change in the corresponding period, as the sum and mean of all regions characterized by that trajectory. Subsequently, we analyse whether each of these classes is statistically different from the complete set of all classes combined in terms of their population and built-up land in the starting year and the changes in both for each study period. These differences were assessed using the two-tailed student's *t*-test.

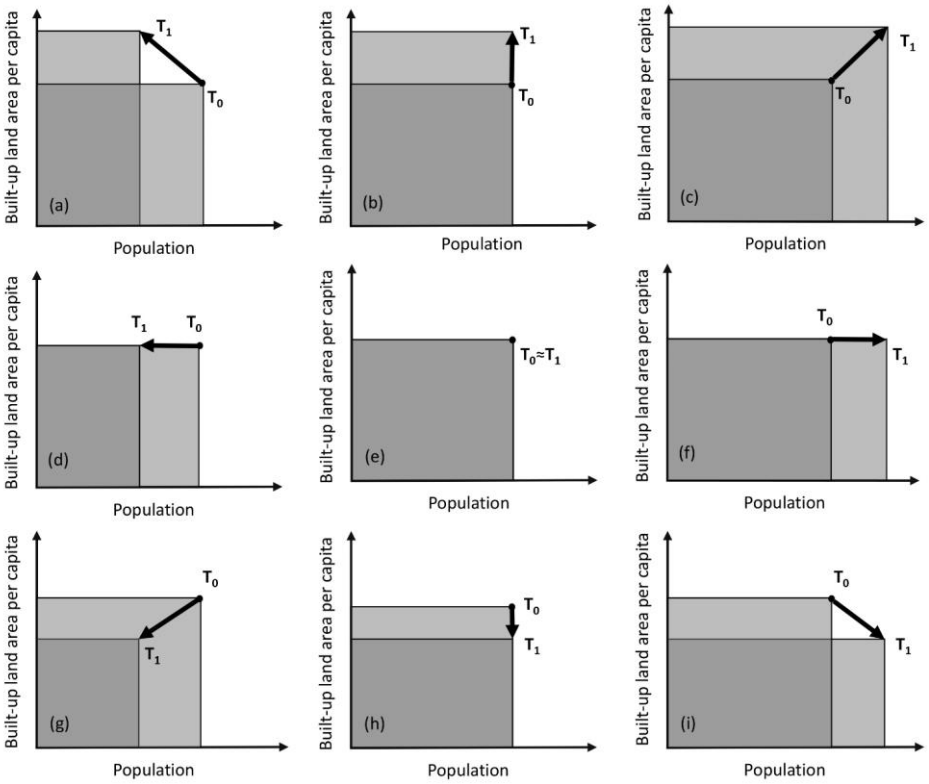


Figure 5.1: Graphical representation of different built-up land change trajectories as a function of changes in population and changes in built-up land area per capita (BPC). (a) Population decline: a decrease in population in combination with an increase in BPC; (b) Expansion: a stable population in combination with an increase in BPC, leading to an increase in total built-up land; (c) Sprawl growth: an increase in population in combination with an increase in BPC, thus representing a decrease in land-use intensity; (d) Proportional decline: a decrease in population in combination with a stable BPC; (e) Persistence: no substantial change in either population or BPC; (f) Proportional growth: an increase in population in combination with a stable BPC; (g) Shrinkage: a decrease in population in combination with a decrease in BPC; (h) Stable densification: a stable population in combination with a decrease in BPC; (i) Growth densification: an increase in population in combination with a decrease in BPC, denoting densification.

5.2.4 Analysis of built-up land change for different types of regions

For the analysis of change trajectories, we further subdivided all regions into five classes based on their population size and their vicinity to populated places. This subdivision yielded five groups: 1) *large city centres*, 2) *large city peripheries*, 3) *small and*

medium city centres, 4) *small and medium city peripheries*, and 5) *other regions*. To categorize regions accordingly, we obtained data on population numbers from Natural Earth (https://www.naturalearthdata.com/). These data are point data representing the inhabitants of the metropolitan areas of cities (see supplementary Figure D1 for their geographical distribution), and are independent from the regions and population data used in the rest of our analysis. We use these point data because we are interested in the structure of cities, and this allows to differentiate between urban centres and their peripheries. In addition, the variation in size of administrative regions used in this study constrains the identification of cities by population based on these regions strictly.

All regions that include a point (city) of more than 2M inhabitants were classified as *large city centres*, all regions that include a point (city) of less than 2M but more than 0.3M were classified as *small and medium city centres*. Accordingly, their affiliated peripheral regions identified by GADM cross-level relationship are classified as *large city peripheries,* and *small and medium city peripheries* (see Figure 5.2). The remaining regions, typically representing more rural counterparts, are classified as *other regions*. We select these threshold values to identify cities that are of global importance due to their size (> 2M) and regional to national importance (> 0.3M). We used population thresholds rather than other classifications of urban areas such as governmental functions (national and provincial capitals) or economic importance ("world cities"), because urbanization, one of the important processes underlying the expansion of built-up land, is primarily a demographic process and thus depending on population numbers. Hence city sizes are directly relevant for studying the land-use impacts of urbanization. The same thresholds have been used previously in other global studies of urban development (see, for example, Güneralp et al. (2020)). It is argued that previous studies have overwhelmingly focused on cities, but a large amount of (new) built-up land is found in smaller towns and villages (Li et al., 2019), we therefore used a subdivision that allows contrasting developments in cities of different sizes versus developments in less populated places. Previous studies have relied on the (demographic) indication of urban versus rural regions (Ehrlich et al., 2021; Wang et al., 2015). However, this subdivision is rather arbitrary and not consistently applied between countries (Easterlin et al., 2011; Lattes et al., 2017). Moreover, indications have changed over time as a result of new classification. Therefore, we rely on the number of inhabitants instead. In addition, recent literature has highlighted the importance of land-use dynamics in peripheral regions versus core cities (Broitman and Koomen, 2020; Shaw et al., 2020). These regions, including commuter towns and peri-urban areas for example, increasingly act as overflow areas of city centres that

cannot grow themselves, and are therefore included them as separate categories together with *other regions* (Chi and Marcouiller, 2013; Salem et al., 2020).

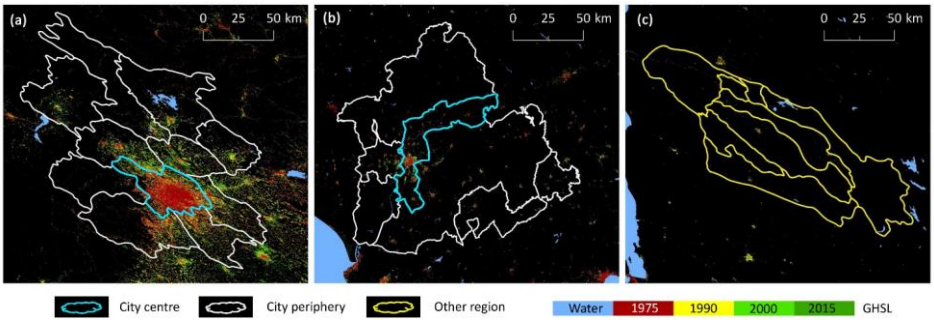


Figure 5.2: An example of the five region types in three metropolitan areas. (a) large city centre, and large city periphery: Beijing in China; (b) small and medium city centre, and small and medium city periphery: Seville in Spain; (c) Other region: Linares in Chile. The background built-up land data is from the multi-temporal layer of GHSL data at its original resolution (~38 m), boundary vector is from GADM.

For different types of regions, we summarize population and built-up land area in the starting year, as well as population change and built-up land area change in the corresponding period, as the sum and mean of all regions of that specific type. Subsequently, we analyse whether each of these classes is statistically different from the complete set of all classes combined in terms of their population and built-up land in the starting year and the changes in both for each study period. These differences were assessed using the two-tailed student's *t*-test. In addition, we also analyse the distribution of different urban development trajectories per type of region.

5.2.5 Analysis of variations in built-up land change trajectories

To analyse the variation in built-up land change trajectories, we analyse the heterogeneity or homogeneity in population changes and changes in built-up land consumption for ten major world regions, as well as for all countries separately. Heterogeneity here refers to population or BPC changes in different directions, while homogeneity refers to changes in the same direction. For example, if all regions in, say, Europe increase in population, this is interpreted as a homogenous development, while if some regions increase and others decrease in population, this is interpreted as a heterogeneous development. We conduct this analysis for ten major world regions: Canada and United States, China, Europe, India, Latin America, Middle-East and Northern Africa, Oceania, Russia and Central Asia, Southeast Asia, and Sub-Saharan Africa (see Supplementary Figure D2). These regions represent more or less

coherent socio-economic world regions, which have also been used in other studies of land use and land cover change (Eitelberg et al., 2016; van Asselen and Verburg, 2013).

For each major world region, we calculate the gross and net change in built-up land area as result of either population change or BPC change. For each unit of analysis (world and major world regions), we calculate gross increase ($A_{POP,inc}$), gross decrease ($A_{POP,dec}$), and net change ($A_{POP,net}$) in built-up land related to population dynamics and the gross increase ($A_{BPC,inc}$), gross decrease ($A_{BPC,dec}$), and net change ($A_{BPC,net}$) in built-up land related to changes in BPC. Gross increase in built-up land according to population change is thus simply the sum of all the built-up land change related to population change for all sub-regions in which the population increases (i.e., $A_{POP,inc}$). Conversely, gross decrease is the sum of the built-up land change related to population change for all sub-regions in which the population decreases (i.e., $A_{POP,dec}$). Therefore, the net change in built-up land attributed to population change (i.e., $A_{POP,net}$) is the sum of both (i.e., $A_{POP,inc}$ and $A_{POP,dec}$). The same for the analysis of built-up changes related to changes in built-up land consumption ($A_{BPC,inc}$ and $A_{BPC,dec}$, respectively). We compare the difference between the gross change and the net change globally, as well as for major world regions, to assess the heterogeneity in urban development.

At a national level, we build on these calculations to quantify heterogeneity or homogeneity in a single index. We propose a Homogenous Urban Development index (HUDI) to indicate to what extent a country has homogeneous or heterogeneous developments in either population dynamics or BPC changes. HUDI is calculated based on the changes of all study regions within one country following:

$$HUDI_{POP} = \frac{A_{POP,net}}{MAX(A_{POP,inc}, |A_{POP,dec}|)} \qquad (5.5)$$

$$HUDI_{BPC} = \frac{A_{BPC,net}}{MAX(A_{BPC,inc}, |A_{BPC,dec}|)} \qquad (5.6)$$

Where $HUDI_{POP}$ and $HUDI_{BPC}$ are the HUDI index for built-up land change due to the changes in population and BPC, respectively. By definition, HUDI values range from -1 to 1, where a positive value indicates an increase in built-up land due to changes in population and BPC, respectively, and a negative value indicates the opposite. The absolute values indicate the homogeneity or heterogeneity of the changes in a country. For example, a value of 1 for $HUDI_{POP}$ means that all regions

within a country have an increase in population leading to an increase in built-up land. Conversely, a value for $HUDI_{BPC}$ close to 0 would mean that BPC increases in some regions, while it decreases in other regions in the same country, with a very small net change relative to the gross increase and gross decrease, thus indicating heterogeneity in BPC-related changes in this country.

5.3 Results

5.3.1 Built-up land change trajectories and their characteristics

At a global level, built-up land increased by 143, 132, and 133 thousand km² during the periods 1975-1990, 1990-2000, and 2000-2015 respectively, of which 38.3%, 49.6%, and 37.5% can be attributed to an increase in built-up land area per capita (BPC), suggesting a decreasing density worldwide. Different methods for attributing built-up changes to changes in population and changes in BPC yield ranges of 29.1%-38.3% for 1975-1990, 48.0%-49.9% for 1990-2000, and 29.5%-43.6% for 2000-2015 (see supplementary Table D2, and Figures D5-D7).

Across the globe, built-up land development at a local level follows different trajectories in terms of changes in population and in land take per person (see Figure 5.3a-c and Table 5.1). Because no region saw a decline in built-up land area in any of the three periods analysed, three of the nine built-up land change trajectories (i.e., *proportional decline*, *shrinkage*, and *stable densification*) were not observed.

Globally, only 2870 out of the 75102 regions in the period 1975-1990 follow the *population decline* trajectory, which is characterized by a decrease in population and an increase in BPC. This number increased to 6474 and 8893 regions in 1990-2000 and 2000-2015, respectively. Regions following this trajectory are located mostly in Europe, Russia, and Central Asia, but in recent periods some, mostly rural, regions elsewhere also followed this trajectory. For example, population decreased across inland regions of Portugal between 2000 and 2015 (Figure 5.3e). Because no built-up land was lost, land-use intensity decreased in these regions. Regions following this trajectory represent a relatively small increase in built-up land, totalling 2.3, 4.5, and 5.5 thousand km², in the three different periods, respectively (Table 5.1).

The *expansion* trajectory, characterized by a stable population and an increase in BPC, represents the largest total increase in built-up land in all periods combined and also in the periods 1990-2000 and 2000-2015. In the latter two periods, this trajectory added 43.8 thousand km² and 45.7 thousand km² built-up land, respectively (Table

5.1). Throughout all three periods this trajectory is mostly observed in North-America, Europe, Russia, and China (Figure 5.3a-c).

Regions following the *sprawl growth* trajectory are mostly located in the Global South, including Latin America and India, but also scattered across developed regions like Europe and North America. This trajectory is characterized by an increase in both population and BPC. Regions that follow this trajectory include a large proportion of the increase in population, especially during 1975-1990 and 1990-2000 (Table 5.1).

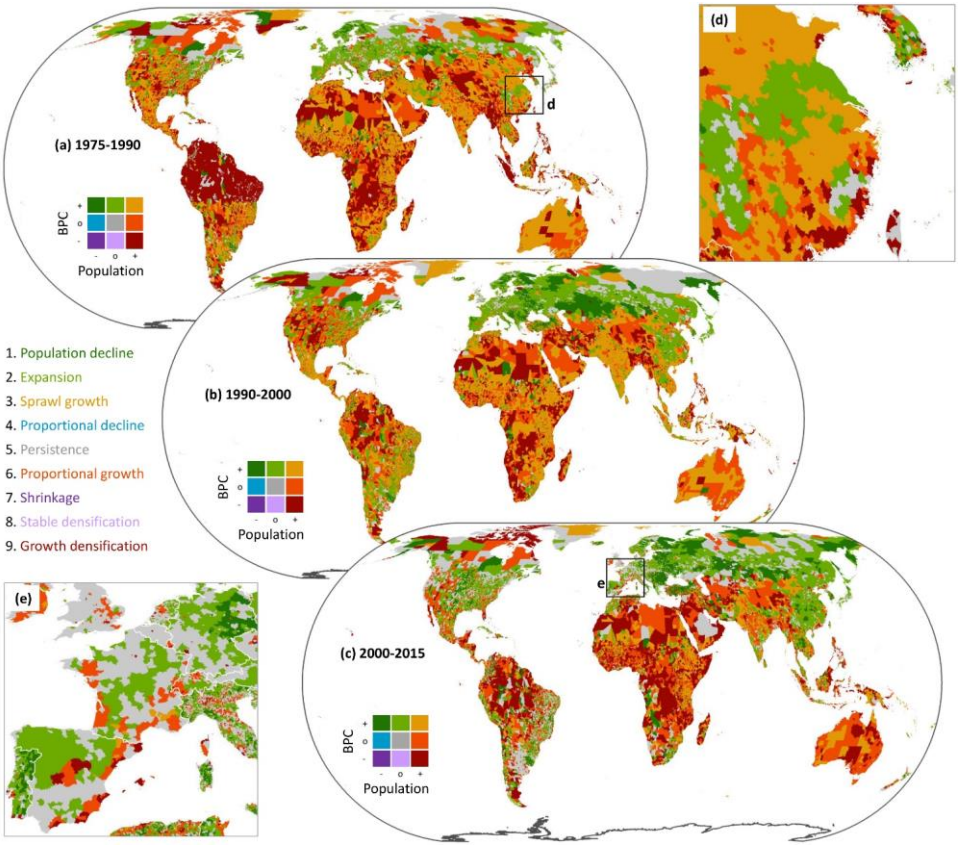


Figure 5.3: Built-up land change trajectories across the world in the periods 1975-1990, 1990-2000, and 2000-2015. BPC refers to built-up land area per capita, which is an expression of land-use intensity. Figures (a), (b), and (c) present the different change trajectories in the periods 1975-1990, 1990-2000, and 2000-2015, respectively, for 75102 regions across the world. Figures (d) and (e) show changes in eastern China and Europe in more detail, respectively, and both of them are 2000 × 2000 km in size. Note that trajectories 4, 7, and 8 are not observed and that they are only included in the legend for the sake of completeness.

Table 5.1: Built-up land and population for different change trajectories as well as their changes over time. For each trajectory, 'Sum' refers to the total built-up land area or population in the starting year, and their total changes in that period. Similarly, 'Mean' represents the average built-up land area or population per subdivision in the starting year, and their average changes per region in the corresponding trajectories. As indicated, mean values for nearly all major world regions are significantly different from the mean of the entire regions combined (two-tailed t-test, $^{**}p<0.01$, $^{*}p<0.05$).

Period	Trajectory (# regions)	Built-up land in the starting year [km²]		New built-up land [km²]		Population in the starting year [million]		Population change [million]	
		Sum	Mean	Sum	Mean	Sum	Mean	Sum	Mean
1975-1990	Population decline (2870)	4946	1.72**	2255	0.79**	161	0.06	-60	-0.02**
	Expansion (15309)	80168	5.24	38656	2.53**	793	0.05	34	0.00**
	Sprawl growth (21056)	39531	1.88**	52292	2.48**	1123	0.05	482	0.02**
	Persistence (7936)	87580	11.04**	9749	1.23**	518	0.07**	35	0.00**
	Proportional growth (10517)	91732	8.72**	30144	2.87**	844	0.08**	334	0.03**
	Growth densification (17414)	71109	4.08**	9892	0.57**	641	0.04**	386	0.02**
1990-2000	Population decline (6474)	16849	2.60**	4452	0.69**	218	0.03**	-53	-0.01**
	Expansion (22415)	155369	6.93	43821	1.95**	1356	0.06**	17	0.00**
	Sprawl growth (22739)	70156	3.09**	40322	1.77	1519	0.07*	378	0.02**
	Persistence (5528)	93351	16.89**	8070	1.46**	596	0.11**	25	0.00**
	Proportional growth (10993)	128573	11.70**	27840	2.53**	1049	0.10**	251	0.02**
	Growth densification (6953)	53755	7.73*	7091	1.02**	554	0.08**	214	0.03**
2015	Population decline (8893)	28301	3.15**	5451	0.61**	291	0.03**	-74	-0.01**
	Expansion (17744)	150307	8.47	45661	2.57**	1360	0.08*	-3	0.00**
	Sprawl growth (9102)	18333	2.01**	15124	1.66	627	0.07**	221	0.02**
	Persistence (15276)	238499	15.61**	24962	1.63*	1529	0.10**	101	0.01**
	Proportional growth (13523)	127716	9.44*	30110	2.23**	1335	0.10**	409	0.03**
	Growth densification (10464)	86492	8.27	11550	1.10**	982	0.09**	532	0.05**

The *persistence* trajectory, by definition, has the smallest changes in population and in BPC. These regions are not abundant and mostly found in Europe, the United States, Russia and Central Asia. Nonetheless, these regions have the largest mean built-up land area in the starting year. Because we included a 1% margin around the no change, some changes in built-up land took place in these regions. In fact, this trajectory included an increase of 9.7, 8.1, and 25.0 thousand km² in the periods 1975-1990, 1990-2000, and 2000-2015, respectively (Table 5.1).

The *proportional growth* trajectory, characterized by an increase in population and a stable BPC, is mostly found in Africa, Latin America, India, and Southeast Asia, but also scattered in developed regions like Europe, the United States, and Oceania. The contribution of this trajectory to built-up land change was rather constant over time, ranging from an extra 27.8 thousand km² between 1990 and 2000 to 30.1 thousand km² in the other two periods, for all regions in this trajectory together (Table 5.1). On average, these regions see the largest increase in built-up land area per region in the periods 1975-1990 and 1990-2000, and the second largest in 2000-2015.

The *growth densification* trajectory can be found in rural areas in the Global South, as well as a few large cities in developed regions. It is characterized by an increase in population in combination with a decrease in BPC. The well-developed coastal region of China shown in Figure 5.3d provides an example of the latter. This trajectory is the only trajectory that is characterized by a decrease in BPC, but this decrease does not lead to any decrease in built-up land. Instead, it is more than compensated by the generally large increase in population in these regions. As a result, regions following this trajectory add a total of 9.9, 7.1, and 11.6 thousand km² of built-up land in the periods 1975-1990, 1990-2000, and 2000-2015, respectively (Table 5.1). The number of regions following this trajectory decreased from 17414 in the period 1975-1990 to 6935 in the period 1990-2000 and 10464 in the period 2000-2015.

5.3.2 Built-up land change in different types of regions

We analyse to what extent built-up land change trajectories differ between different types of regions. Not unexpectedly, *large city centres* (> 2M inhabitants) and *small and medium city centres* (between 0.3M and 2M inhabitants) have a much larger mean area of built-up land, and a much larger mean increase in built-up land per region over time than their peripheries (Table 5.2). While the mean area of built-up land and the mean increase in built-up land is by far the lowest for *other regions*, the sum of all new built-up land in these regions is about as large as that of all *large city centres*, *small and medium city centres*, and their *peripheries* combined (see Table 5.2). In the year 1975, BPC

in *small and medium city centres* is higher than the BPC in *large city centres*, which is again larger than the BPC in *other regions* (Table 5.3). Yet, over time, *large city centres* intensify, and in the final year of our analysis their BPC is smallest, on average. Conversely, land-use intensity in *other regions* decreases over time as expressed in an increase in BPC. The BPC in *small and medium city centres* remains more or less the same, and remains the highest of all five region types (Table 5.3). The BPC in peripheries is lower than that of the *large city centres* and *small and medium city centres* in all three time periods.

Consistent with the average changes in population and BPC, regions indicated as *large city centres* and *small and medium city centres* more often follow the *growth densification* and *proportional growth trajectories*, while *other regions* follow the *population decline* and *expansion* trajectories more often between 2000 and 2015 (Figure 5.4). Urban peripheries fall in between these two extremes, both for *large city peripheries* and *small and medium city peripheries*. This patterns holds for all three periods analysed (see supplementary Figure D3 for 1975-1990 and 1990-2000).

To investigate whether the observed densification in *large city centres* could be related to land scarcity, we analysed the proportion of built-up land in the five region classes at the four epochs. *Large city centres* have on average the largest ratio of built-up land to available land, as compared to other types of regions. Yet, built-up land in *large city centres* accounts for about 50% on average of all available land in 2015, where available land denotes total land area in the specific region but excluding water body. In contrast, built-up land in *other regions* account for a very small fraction on average to its total available land in 2015 (see supplementary Figure D4 for detailed results of this analysis).

Table 5.2: Built-up land and population for regions in different urbanization classes as well as their changes over time. For each region type, 'Sum' refers to the total built-up land area or population in the starting year, or their total changes. Similarly, 'Mean' represents the average built-up land area or population per subdivision in the starting year, or their average changes per subdivision for corresponding region types. As indicated, mean values for nearly all major world regions are significantly different from the mean of the entire regions combined (two-tailed t-test, $^{**}p<0.01$, $*p< 0.05$).

Period	Region type (# of regions)	Built-up land in the starting year [km²]		New built-up land [km²]		Population in the starting year [million]		Population change [million]	
		Sum	Mean	Sum	Mean	Sum	Mean	Sum	Mean
1975-1990	Large city centres (105)	17134	163**	4608	44**	192	1.83**	59	0.56**
	Large city peripheries (1307)	25719	20**	13031	10**	378	0.29**	126	0.10**
	Small and medium city centres (911)	52693	58**	16265	18**	454	0.50**	142	0.16**
	Small and medium city peripheries (8128)	77226	10**	41254	5**	1081	0.13**	327	0.04**
	Other regions (64651)	202292	3**	67830	1**	1976	0.03**	559	0.01**
1990-2000	Large city centres (105)	21742	207**	3760	36**	250	2.38**	51	0.48**
	Large city peripheries (1307)	38750	30**	11765	9**	504	0.39**	85	0.06**
	Small and medium city centres (911)	68958	76**	14104	15**	595	0.65**	100	0.11**
	Small and medium city peripheries (8128)	118481	15**	35164	4**	1408	0.17**	203	0.03**
	Other regions (64651)	270122	4**	66803	1**	2534	0.04**	392	0.01**
2000-2015	Large city centres (105)	25502	243**	2384	23**	301	2.87**	66	0.63**
	Large city peripheries (1307)	50515	39**	14412	11**	589	0.45**	117	0.09**
	Small and medium city centres (911)	83063	91**	11451	13**	696	0.76**	132	0.15**
	Small and medium city peripheries (8128)	153644	19**	41204	5**	1611	0.20**	260	0.03**
	Other regions (64651)	336924	5**	63408	1**	2927	0.05**	610	0.01**

Table 5.3: BPC in the starting years and relative annual changes in built-up land, population, and BPC for five region types. For nearly all periods changes in built-up land, population and BPC for each of the five region classes were significantly different from these for all settlements combined (two-tailed t-test, $^{**}p<0.01$, $^{*}p<0.05$).

Period	Region type (# of regions)	BPC in the starting year [m²/person]	Δ Built-up land	Δ Population	Δ BPC
1975-1990	Large city centres (105)	89	1.60%**	1.79%	-0.19%**
	Large city peripheries (1307)	68	2.77%**	1.94%**	0.82%
	Small and medium city centres (911)	116	1.81%**	1.83%	-0.02%**
	Small and medium city peripheries (8128)	71	2.89%**	1.78%**	1.10%**
	Other regions (64651)	102	1.95%*	1.67%	0.27%
1990-2000	Large city centres (105)	87*	1.61%**	1.87%	-0.25%**
	Large city peripheries (1307)	77*	2.69%	1.57%**	1.10%**
	Small and medium city centres (911)	116	1.88%**	1.57%	0.30%**
	Small and medium city peripheries (8128)	84	2.63%**	1.36%	1.26%**
	Other regions (64651)	107	2.23%*	1.45%	0.77%**
2000-2015	Large city centres (105)	85	0.60%**	1.33%	-0.72%**
	Large city peripheries (1307)	86**	1.69%	1.22%**	0.47%**
	Small and medium city centres (911)	119	0.86%**	1.17%	-0.30%**
	Small and medium city peripheries (8128)	95**	1.60%	1.00%*	0.59%**
	Other regions (64651)	115	1.16%	1.27%	-0.11%**

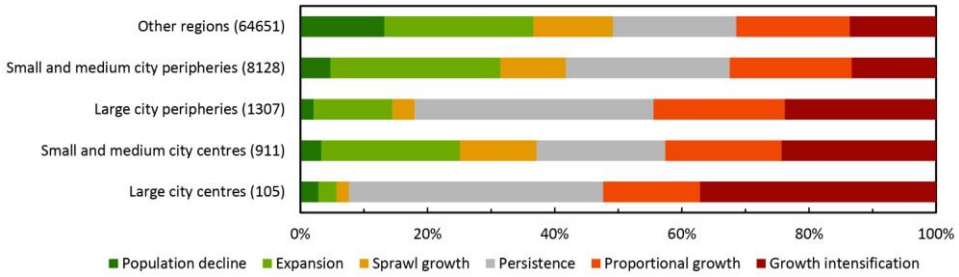


Figure 5.4: Distribution of built-up land change trajectories over the different types of regions for the period 2000-2015. Colour coding is consistent with Figure 5.3.

5.3.3 Gross and net changes for major world regions.

Globally, changes both in population and in BPC play an important role in the increase in built-up land, especially in the period 1990-2000. We find that built-up land changes due to population changes are relatively homogenous within major world regions, as almost all regions in the world increase in population (Figure 5.5k). Conversely, built-up land changes due to BPC changes are more heterogeneous, indicating that BPC increases in some regions while it decreases in others (Figure 5.5k). At the level of major world regions, some differences become visible (Figure 5.5a-j).

Canada and USA, China, and Europe experienced the largest increase in built-up land. In Europe, this is mainly driven by increases in BPC, while in China and Canada and USA this increase is driven more or less equally by population change and BPC change. However, BPC in Canada and USA was already much higher at the start of each of the periods analysed (see supplementary Tables D3 and D4). In Europe, there is only a small net increase of built-up land related to population change during 2000-2015, but gross changes are relatively large, illustrating the heterogeneity in population dynamics here. Specifically, the net increase of 0.46×10^4 km^2 built-up land due to population change in this period is the result of 1.09×10^4 km^2 gross increase and 0.63×10^4 km^2 gross decrease related to contrasting population dynamics across different regions within Europe (also see Figure 5.3). In the Global South, including Latin America, Middle-East and Northern Africa, and Sub-Saharan Africa, built-up land increase is mostly driven by population growth, whereas urban densification, i.e., a net decrease in BPC, reduced the increase in built-up land. For example, we find a net decrease of 0.48×10^4 km^2 built-up land in Sub-Saharan Africa during 2000-2015, which is due to the intensified use of built-up land (Figure 5.5j). Moreover, these regions show much less heterogeneity, especially in their built-up land changes due to

population dynamics, indicating that nearly all regions increase in population or at least stay constant.

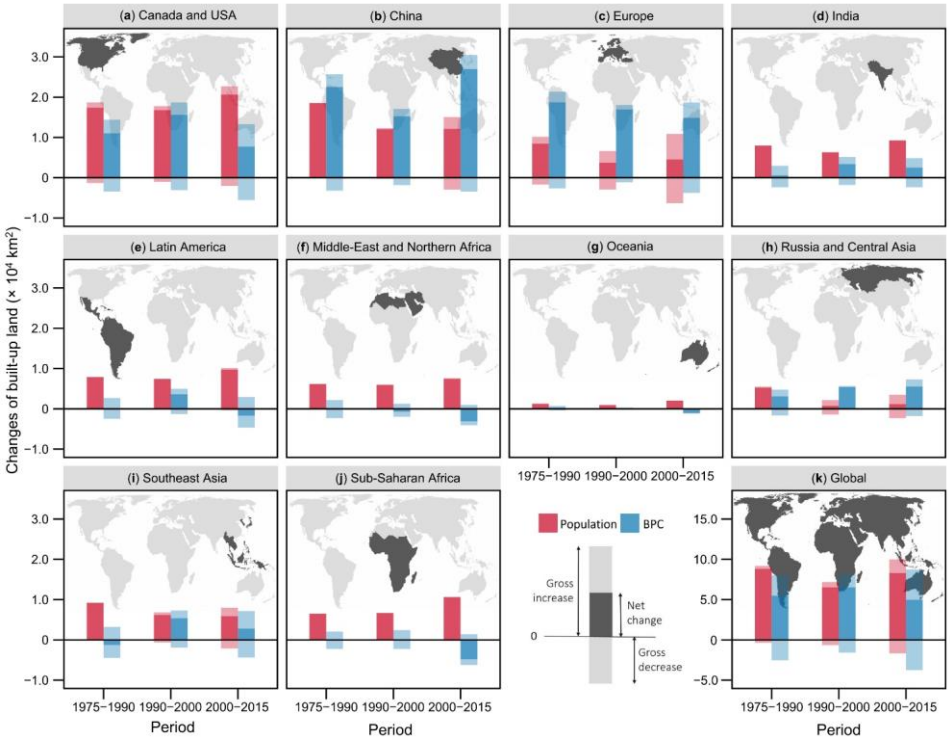


Figure 5.5: Heterogeneity in built-up land changes due to population change and BPC change in major world regions. Red and blue bars represent changes in built-up land due to changes in population and changes in BPC, respectively. The charts show gross changes in both directions, with the light-coloured bars indicating the equivalent change in opposite direction, highlighting the relationship between gross changes and net changes. Net changes that are equal or almost equal to the gross changes are interpreted as homogenous change trajectories in that major world region, while net changes that are much larger than their corresponding net change are interpreted as heterogeneous change trajectories in that major world region. See supplementary Table D5-D7 for results from the other three approaches to allocating built-up land change.

At a national level, changes in built-up land due to population dynamics are rather homogeneous for almost all countries, while changes in BPC per country are much more heterogeneous (see supplementary Figure D8). This indicates that in most countries, population either increases or decreases in almost all regions within one period. For example, population in Latvia grows in nearly all regions during 1975-

1990, while the two consecutive periods show a population decrease in all regions. Urban development in the most recent period in the Global North is a notable exception to this trend, mostly as a result of rural depopulation and related rural-urban migration. Changes in built-up land due to changes in BPC are much more heterogeneous in many countries (see supplementary Figure D8). For most European countries, the majority of all regions show an increase in BPC, indicating a decrease in land-use intensity. Yet, in recent decades, more heterogeneous developments were observed in some European countries, notably Spain, United Kingdom, and Italy, combining mostly densification of urban areas with decreasing density in rural areas (i.e., rural depopulation). Conversely, many countries mostly in Africa, but also a few in Latin America in the first and the last periods, show a homogenous pattern of densification in nearly all regions within those countries, mostly because the population increased faster than built-up land (see supplementary Figure D8).

5.4 Discussion

5.4.1 Divergent patterns of built-up land change trajectory

Globally, 38.3% (29.1%-38.3%) of all new built-up land during 1975-1990 can be attributed to BPC change, whereas this share increased to 49.6% (48.0%-49.9%) during 1990-2000, and was 37.5% (29.5-43.6%) during 2000-2015. This suggests that the decrease in land-use intensity plays a very important role in global built-up land expansion, approaching population growth in terms of their relative importance. These findings confirm earlier studies, which found that residential density has been decreasing almost everywhere for decades (Angel et al., 2010; Mahtta et al., 2019; Seto et al., 2011).

While built-up land increases faster than the population, globally, we find a large number of regions that do not follow this trend. This observation is consistent with earlier findings for specific regions in Europe (Wolff et al., 2018a) and the United States (Richter, 2020). Most notably, we find that *large city centres* of > 2M inhabitants are densifying, on average, while the opposite is true for *other regions* characterized by the absence of cities with > 0.3M inhabitants and their peripheries. These findings suggest that urbanization, and specifically the influx of people in *large cities*, could partially contribute to SDG target 11.3 by reducing land take per person. Conversely, we predominantly find a decrease in density in their peripheries. This also holds true for *other regions* characterized by depopulation, sometimes leaving empty buildings behind (Tietjen and Jørgensen, 2016), but also in some regions with population increase.

Densification in *large city centres* and *small and medium city centres* contrasts with the pattern of decreasing density in *large city peripheries, small and medium city peripheries*, and *other regions* in all three periods. Initially, *large city centres* and small and medium city centres had a BPC that was higher than other regions, but this pattern turned around over time due to an increasing concentration of people in both types of city centres. An example of this can be seen in Spain and Portugal in Figure 5.3a-c, where coastal city centres (such as Barcelona and Valencia) are characterized by urban densification, while the opposite is true for most of the adjacent inland areas. The decrease in density of *other regions* is at least partly related to a decrease in population, rather than an increase in built-up land strictly. These processes have been described elsewhere as rural land abandonment (Vannier et al., 2019; Weissteiner et al., 2011) and urban densification (Broitman and Koomen, 2015; Kyttä et al., 2013). Our study links these two processes, which could partly explain the heterogeneity in built-up land changes due to population changes within a country or major world region.

Our finding that *large city centres* densify seems to contradict earlier studies reporting a decrease in density for cities, including prior analyses of urban sprawl (Güneralp et al., 2020; Seto et al., 2011). One possible explanation could be the delineation of these cities. In this study, we characterized regions that include cities > 2M as *large city centres*, but their surrounding regions (i.e., *large city peripheries*) that are often part of their metropolitan areas are analyzed separately. For these narrowly defined *large city centers*, we found urban densification on average, which reflects the findings by Angel et al. (2021b) that in most cities investigated, there were net increases in population in the areas built before the starting year, and thus densified significantly. However, much of the sprawl related to urbanization takes place in the outskirts, rather than the city proper (Angel et al., 2021b; Salem et al., 2020). Hence, we suggest that these results do not contradict earlier findings, but that they add an important nuance: *large city centres* themselves are on average intensifying, while their surrounding areas often become less dense, for example as part of a process of peri-urbanization and urban sprawl fueled by increased accessibility and decreased costs of transport (Anas et al., 1998). This process of density changes also potentially explains the heterogeneity in built-up land changes due to BPC changes which was observed for multiple regions.

A possible explanation for urban densification is the lack of available land in *large city centers*. We found that on average built-up land in *large city centers* is approaching half the land area of these regions (see supplementary Figure D4). Given that 100% is not a realistic upper boundary, as there is always a need for green space, the lack of land might explain at least part of the observed densification. Their variation in region sizes is also reflected in the variation in the range of built-up land percentages per

region, as some of the regions with a low percentage of built-up land are in fact relatively larger regions. Therefore, the impact of land availability in *large city centers* might be stronger than reported in this study. Nonetheless, the observation that residential density on average increased in these regions is still remarkable as a step towards SDG target 11.3, as it is not inevitable. After all, a lack of land in *large city centers* could also lead to population increases in their neighboring regions. Yet, average population increases for *large city centers* were larger than these for *large city peripheries* in all three study periods. Moreover, *large city centers* also showed an annual increase in built-up land of 1.61% in the first period and 0.60% in the last period, suggesting that at least some expansion was still possible. While *small and medium city centers* also have a higher average share of built-up land than their peripheries or *other regions*, these shares remain much lower than for *large city centers*. Therefore, we expect that densification due to a lack of available land plays no important role in these categories strictly. Yet, consistent with economic theories of land-rent (Anas et al., 1998; Manganelli and Murgante, 2017), locations closer to city centers are more attractive and thus more expensive, which might trigger densification for both *large city centers* and *small and medium city centers*.

5.4.2 Limitations

As any other data-driven analysis, our study results depend on data accuracy. Currently, GHSL is the only accessible dataset providing consistent time-series of both built-up land and population density at a global scale (Balk et al., 2018; Leyk et al., 2019; Melchiorri et al., 2018), allowing for the long-term analysis presented in this study. Globally, GHSL has an overall accuracy of >70% (Blei et al., 2018; Marconcini et al., 2020), but with a relative low accuracy in some sparsely urbanized areas (Leyk et al., 2018) and some parts of Africa and South America (Gómez et al., 2019; Sliuzas et al., 2017). Because the GHSL data is produced as a time-series, built-up land is mapped consistently for different time periods (Pesaresi et al., 2016). Hence our results are potentially affected by the data accuracies, but this is unlikely to yield dramatically different outcomes (e.g., increases in built-up land where in reality there are decreases).

Population data in GHSL is only lightly modelled. More specifically, population estimates in a region are downscaled within that region based on the presence and absence of built-up land only (Leyk et al., 2019). This potentially yields large inaccuracies at a pixel level, but it will not have a large impact on the average population density in a region, as long as the region for which this average is calculated is equal to or larger than the regions within which population is downscaled.

Consistent with findings in Tellman et al. (2021), our exploratory analyses of GHSL dataset show that in rare cases, regions exhibit suspicious and dramatic population changes, which potentially affect our results. For example, we found a few regions with sudden drops in one-time period and sudden rises in the next period, which could not be confirmed using other sources. Changes in these regions are consequently categorized into *population decline* and *growth densification*, since BPC is derived from the population numbers in a particular year. For a few regions with suspected unrealistic fluctuations, we checked secondary sources to see if we could explain such changes. As a result, for a few regions we adjust population size manually. Yet, an assessment of these outliers suggests that these regions only account for a small fraction of the total change in built-up land, and therefore these irregularities will not affect main findings of this study. GHSL population estimates for target years 1975, 1990, 2000 and 2015 are modelled on the basis of GPWv4 (Doxsey-Whitfield et al., 2015), which is based on more than 12.5M input units. This number is much higher than the number of regions analysed in this study, thus making the influence of pixel-level inaccuracies on our results due to the light modelling likely very small. Moreover, downscaled population grids in GHSL are consistent with, and adjusted to, United Nations World Population Prospects (Melchiorri et al., 2019), making them suitable for the analysis of long term process of human settlement changes possible. As a result, we believe that our outcomes remain valid despite the known pixel-level inaccuracies in the GHSL. Yet, we expect that some areas with relatively large input units for population data, such as in Africa and Central Asia might have a lower accuracy (Freire et al., 2016).

Built-up land change trajectories can vary within our study regions, and consequently, our findings depend on the delineation of our study region. We selected administrative units because population data was also downscaled from administrative units, although typically at a higher level (i.e., from smaller units). As administrative units are normally nested, this choice avoids large inaccuracies because of boundaries that are not coinciding. In addition, the selected regions typically represent municipalities or counties, which allow for a meaningful classification in the context of urbanization and rural-urban migration, leading to the five types of regions reported in Section 5.3.1. Taubenböck et al. (2019) show that morphological urban areas are potentially better delineations of cities. While this holds for the delineation of larger cities, regions created by this method do not coincide with the administrative regions underlying the population data in GHSL. Moreover, urban morphological zones do not provide a regional subdivision the land that does not belong to these urban morphological zones, thus constraining their application for a study with global coverage. Regular delineations such as pixels or hexagons (e.g., Boudet et al. (2020))

also suffer from this limitation. One disadvantage of our units of analysis is their range in sizes (e.g., regions in Saudi Arabia are much larger than in Western Europe). These differences might hamper comparability between regions. Yet, larger units generally represent more remote regions, where there is very little built-up land and which do not contribute much to the overall process of urban development. Therefore, it is likely that the variation in size of the regions analysed only has minimal impacts on our results.

5.4.3 Implications for sustainable urban development

Built-up land change trajectories analysed in this study provides the basis for a better understanding of urbanization processes across the globe, and thus inform policies towards sustainable urban development. Specifically, our findings provide important nuances to the widely acknowledged narrative that residential density has been declining almost everywhere for decades (Angel et al., 2011), and suggests that progress towards SDG target 11.3 (indicator 11.3.1) requires a consideration at smaller spatial scales, including both major world regions and subnational units. On a global level, our results suggest that the process of urbanization, i.e., the increasing share of the population living in urban areas, in itself is not necessarily unsustainable from a land take point of view, as built-up land in *large city centres* as well as *small and medium city centres* is used more intensively over time. Despite other findings that such planning initiatives have not been very successful this far (Cortinovis et al., 2019), we observe many larger cities that have intensified over the last few decades, and especially since the year 2000.

Urban densification can reduce demand for future land take, but might also adversely affect other dimensions of sustainability. For instance, dense urban areas are often associated with a higher environmental quality in some parts of Greater Helsinki, but dense areas are also found harmful elsewhere, potentially leading to social inequality (Kyttä et al., 2013). In some densely populated areas like Hong Kong, Japan, and Singapore, policies aiming to promote high-density development may hamper liveability, for instance, due to a loss in urban green space (Richards et al., 2017), decrease in air quality (Grêt-Regamey et al., 2020), and an increased urban heat island effect (Seto and Christensen, 2013). Apart from *large city centres*, most land-use densification related to built-up land in the past decades is located in Latin America, Middle-East and Northern Africa, and Sub-Saharan Africa, where BPC was already much smaller than in developed regions such as Europe (see supplementary Tables D3 and D4). Urban densification in the Global South, and especially in Sub-Saharan Africa is often adversely associated with human well-being (Güneralp et al., 2017a;

Sclar et al., 2005). Many of these regions had a low built-up land area per person at the start of our analysis period, while they also score lower on other aspects of sustainability, such as urban infrastructure and energy consumption (Bakker et al., 2021; Nagendra et al., 2018). Therefore, urban densification in these regions might reduce land take, but it inevitably comes with socio-economic trade-offs. For example, in Sub-Saharan Africa, about 200 million people (United Nations, 2016) were living in slums in 2014. Although the share of population inhabited in urban slums has been declining since 1990, the development of new slums could be interpreted as urban densification (Thorn et al., 2015; Vermeiren et al., 2012). Such developments deteriorate one situation for SDG target 11.1, i.e., urban population living in slums, and therefore suggests a clear trade-off between both targets.

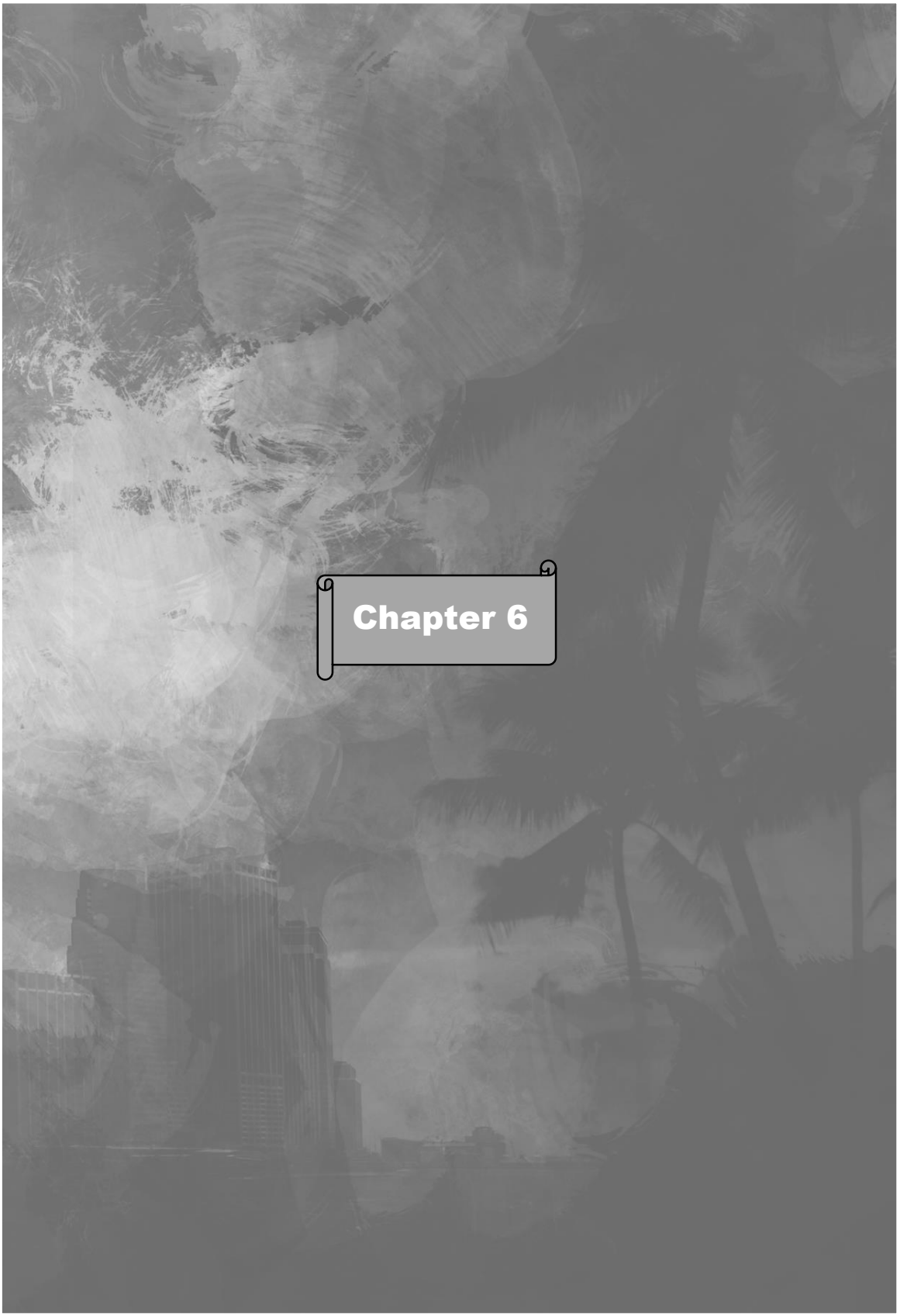
To operationalize sustainability practices, policies towards urban sustainability need to account for the spatial inequality of population, social conditions, environmental resources, etc., and implement interventions that are efficient, requiring less undesirable transformational change (Abson et al., 2017). In developed countries, where residential density is relatively low, policy-makers and urban planners could take actions to increase land-use intensity in the context of global environmental change, for instance, by optimizing land configuration (Angel et al., 2020). Yet, in densely populated regions that are socially vulnerable, local policies towards further urban densification are likely to have adverse effects, and policies to limit the development of over-populated communities and prioritize the development of public infrastructure in terms of transportation and sanitation networks would be preferable. Local context further matters in terms of available infrastructure. To avoid the hysteresis of infrastructure, newly developed built-up land could prioritize areas with available infrastructure (Espindola et al., 2017), or plan for orderly expansion of built-up land by organizing the territory on the periphery of existing cities in advance of their development. In addition, to alleviate the global competition for land, urban planning should consider for the wider contextual conditions and urban densification should be favoured at those locations that can support such development.

Data and code availability

Tabular data and python scripts for the analyses are freely available (https://cscproject.github.io). Original spatial data can be openly accessed using links provided in the materials and methods section of this article.

Acknowledgements

Mengmeng Li would like to thank China Scholarship Council for the financial support (Grant No 201706510011), and thank Chenhui Chang for the assistance with data visualization in R program. Jasper van Vliet was supported by the Netherlands Organization for Scientific Research NWO in the form of a VIDI grant (Grant No VI.Vidi.198.008). This research contributes to the Global Land Programme (www.glp.earth).


Chapter 6

6 | Synthesis

The objective of this thesis is to advance large-scale analysis of human settlements and their dynamics through the lens of land system with a specific focus on the role of land-use intensity. The preceding **Chapters 2-5** answered the following specific research questions (RQs):

RQ1: How can we represent the diversity of settlement systems?

RQ2: What can we learn from Earth Observation to better characterize urban land-use intensity?

RQ3: How do urban land-use intensity and settlement systems change over time?

This final chapter gives an overview of the main findings and their scientific relevance, followed by a discussion on future research directions. Section 6.1 presents the main findings of this thesis, which are placed in the context of recent developments in the field. Sections 6.2 and 6.3 point out implications for future research and sustainable development, respectively. In the final section 6.3, concluding remarks are presented.

6.1 Overview of findings

6.1.1 Diverse patterns and trajectories of human settlements

Chapter 2 illustrates how human settlements distribute across space and change over time. To do so, three properties derived from built-up land (the density of built-up land, the density of built-up clusters, and the size of the largest cluster) are used to characterize settlement systems in China. Results show that settlement systems are widely distributed, ranging from *large cities* to *deep rural*, and their change trajectories are typically gradual and incremental. Moreover, this chapter suggests that the total increase of built-up land in village landscapes far outweighs that of dense urban regions. This chapter concludes that to advance large-scale analysis of human settlements, we should characterize human settlements more comprehensively.

The conception of settlement systems in Chapter 2 builds on land systems that represent land-use patterns with regard to prevailing land cover, which contribute to global environmental assessments in the era of Anthropocene (Václavík et al., 2013;

van Asselen and Verburg, 2012; Verburg et al., 2015). Chapter 2 only identified settlement systems for China. Nevertheless, this approach can be easily applied elsewhere to address specific research questions. For example, de Lange (2022) recently used this approach to compare spatial difference of settlement systems between Europe and East Africa. Results exhibit a substantial difference in the spatial distribution of settlement systems and settlement patterns between Europe and Eastern Africa (de Lange, 2022). Given that the development of human settlements is often driven by distinctive social-environmental processes of urbanization across the world, with different implications for addressing sustainability issues (Müller et al., 2014; Zhou et al., 2021), the differences within a country (China) and across regions (China, Europe, Eastern Africa), show the necessity to study human settlements as systems along a gradient, as well the geographic context within which these patterns emerge and change.

Settlement systems, as coupled human-environment systems, are not confined to the three properties used in this thesis, but can also include other potentially relevant dimensions. For example, to better represent human interactions with the environment, Ellis and Ramankutty (2008) mapped global anthropogenic biomes that range from dense settlements (e.g., urban) to sparse inhabited systems (e.g., populated forests) using datasets on population density, land use, and land cover. To explore the trajectories of human settlements after peak population in China, Wang et al. (2021b) characterize human settlements by their cluster number and population density. In both studies, population density could be interpreted as the proxy for urban land-use intensity, which is, however, not considered in Chapter 2. The consecutive chapters characterize urban land-use intensity at a large spatial scale, and present its dynamics over time.

Moreover, spatial resolution matters for the characterization of settlement systems. The anthropogenic biomes (i.e., *anthromes*) are recently identified at ~9 km resolution at the equator (Ellis et al., 2021). These *anthromes* are dominated by the combination of multiple land uses and land covers that constitute the heterogeneous mosaic landscapes, particularly, a large proportion of human settlements are enclosed by agricultural land (Ellis et al., 2021). The settlement systems in China presented by Wang et al. (2021b) are characterized at a 2-km resolution, consistent with the resolution adopted in Chapter 2 of this dissertation. In comparison, Dou et al. (2021) characterized land systems in Europe at a 1-km resolution, given that European landscapes are relatively small in spatial scale. Compared with agricultural and natural landscapes, human settlements are often characterized by much more detailed scale variations and heterogeneity, and hence justifying a more detailed land system

classification. To this point, LCZs are characterized at a much detailed scale, typically, at 100 m resolution. Yet, LCZ classes are more climate-driven and are less capable to represent land-use pattern in a broader context. Uncertainties in the input datasets that characterize settlement systems are supposed to have only marginal effects on outcome with the relative coarse resolution in terms of accuracy (Ornetsmüller et al., 2018). Yet, spatial resolution has a substantial effect for rural regions with sparse settlements, as human settlements therein often exist with a smaller coverage relative to agricultural and natural areas, and a smaller size to characterize land systems could lead to a growing proportion of completely agricultural land and completely natural land given the limitation of current datasets on land use and land cover (Verburg et al., 2011), which contradicts the basis of land systems as mosaic landscapes that represent human and nature interactions (Malek and Verburg, 2017; van Asselen and Verburg, 2012).

6.1.2 Mapping urban land-use intensity

One important dimension of settlement land systems is urban land-use intensity. Here, urban land-use intensity describes to what extent urban land is used by humans. Intensive use of urban land can reduce extra land development elsewhere with mixed consequences to nature and human society. Therefore, urban land-use intensity is inherently associated with total land take, and is one of the central topics on the interactions between nature and human activities. Prior to mapping urban land-use intensity, we need to clarify the connotation of urban land-use intensity. Similarly to Erb et al. (2013), urban land-use intensity can be interpreted on the basis of inputs, outputs, and changes in system properties. The input dimension can be represented by resource consumption or infrastructure investments such as electricity consumption, road network, buildings, and other urban facilities (Lariviere and Lafrance, 1999; Xia et al., 2020). The output dimension refers to, but is not limited to, economic and demographic outcomes, depending on specific research contexts (Yan et al., 2017; Zhong et al., 2018). The dimension of changes in system properties could be, for example, the intensity of artificial light at night that reflects the degree to which natural environment is disturbed due to human activities (Levin and Zhang, 2017). These three aspects for measuring urban land-use intensity are often intertwined, for instance, high economic outcomes are often at the cost of high electricity consumption, which consequently leading to brighter artificial light at night.

In Chapters 3-5, urban land-use intensity is represented by 3D building structure and population density. The latter was operationalized more or less equivalent to the conceptualization of *urban density* by Dovey and Pafka (2013). According to Dovey

and Pafka (2013), *urban density* consists of three different but intertwined factors: *buildings-population-open space*. Consistently, Angel et al. (2021a) propose a more systematic framework for the decomposition of *urban density* (more precisely, population density herein) into constituent factors that, when multiplied, reconstitute *urban density*. As illustrated in Figure 6.1, *urban density* is decomposed into two, three, four, and seven factors, respectively. Urban land per capita as investigated in Chapter 5 is the inverse of population density (i.e., *urban density* in this context), namely built-up land per capita, while 3D building structure derived from Earth Observation and auxiliary data in Chapters 3 and 4 is indispensable for the majority of factors that determine the *urban density* (Figure. 6.1).

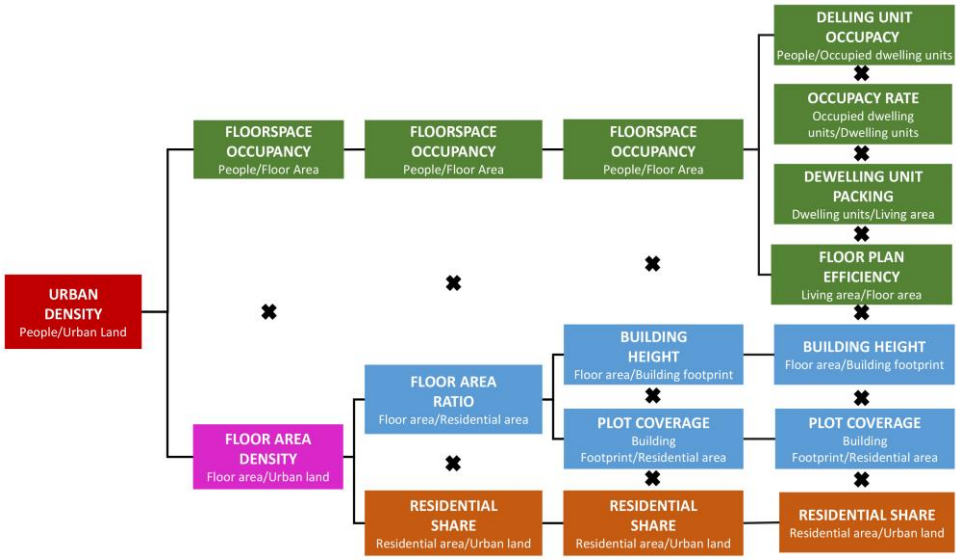


Figure 6.1: The decomposition of urban density. Modified from Angel et al. (2021a).

To portray urban land-use intensity, Chapters 3 and 4 explored the potential of mapping 3D building structure based on Earth Observation. Specifically, an advanced approach is proposed in Chapter 3 to map 3D building structure across China, Europe, and the US. This method is further optimized in Chapter 4, which allowed mapping 3D building structure at a global scale. Chapter 4 also analysed the spatial heterogeneity in 3D building structure across space. The 3D building structure data produced in both chapters provide a nuanced characterization of urban heterogeneity that is useful for the exploration of solutions to addressing urban sustainability issues, such as urban thermal environment and transportation-related carbon emission. In addition, the correlation between building height and building footprint at a pixel level

is only 0.57 worldwide, indicating the relevance of mapping these properties separately.

In light of the decomposition of urban land-use intensity, illustrated in Figure 6.1, global-scale mapping of urban land-use intensity could gain benefits from relevant projects such as WorldPop and OpenStreetMap (OSM). Worldpop project was initiated in 2013, aiming to provide publicly accessible demographic datasets for the Global South to support socioeconomic development therein. The crowdsourced OSM layers include highway, railway, water bodies, building, boundary and others, which are essential elements for measuring urban land-use intensity. Yet, heights are absent for the vast majority of buildings in OSM (Biljecki, 2020; Over et al., 2010), particularly in regions far away from urban centres, hampering the potential applications at a global scale. Moreover, global building footprints have been well mapped individually based on state-of-the-art technique. For instance, Bing Maps is releasing a global dataset of building footprints (Microsoft, 2022). Until now, nearly 800 million building footprints have been from detected from Bing Maps imagery between 2014 and 2021.

In recent years, there are increasing efforts in mapping heterogeneity within built-up areas using Earth Observation data (Demuzere et al., 2022; Esch et al., 2022; Frantz et al., 2021; Zhu et al., 2022). Yet, large-scale mapping of urban land-use intensity remains one of the most challenging open questions.

6.1.3 Global dynamics in urban land-use intensity

Chapter 5 suggests that, at a global scale, the decrease of urban land-use intensity relates to 38.3%, 49.6%, and 37.5% of the built-up land expansion in the three periods during 1975-2015, but with large local variations. These variations in urban land-use intensity dynamics also correspond to the distinctive trajectories of settlement systems presented in Chapter 2. Nevertheless, Chapter 2 illustrates an overall change towards increasingly denser urban patterns. In fact, the decrease in urban land-use intensity could explain why so many areas are characterized by the conversions of *deep rural* to *isolated villages*, and *isolated villages* to *sparse villages*. Chapter 5 suggests that centers of large cities intensify in all the three periods (i.e., 1975-1990, 1990-2000, and 2000-2015), while their rural counterparts show an opposite direction. In the Global South, densification often happens in regions where human settlements are already used intensively, suggesting potential trade-offs with other living standards.

Given that analyses of urban land-use intensity represented by 3D building structure in Chapters 3 and 4 are confined to one point of time, Chapter 5 therefore uses built-

up land per person as an operationalization for urban land-use intensity in order to investigate its temporal dynamics at a global scale. The multiple pathways in urban land-use intensity observed in Chapter 5 imply the different trajectories of settlement systems. For example, the *growth intensification* trajectory and *population decline* trajectory observed in Chapter 5 are often associated with densification and abandonment, respectively. These changes in land-use activities have profound environmental and socioeconomic impacts. In particular, compared to horizontal urban expansion, urban densification often has multiple environmental merits such as reduced carbon emission due to the decreases in energy consumption for transportation, but it is not always without adverse impacts (Næss et al., 2019). For instance, high density has long been equated to congestion, overcrowding, and high crime rates (Khan and Carville, 2017). The distinctive impacts of settlement change trajectories suggest that environmental assessments should integrate these nuanced settlement changes that goes beyond the mere presence and/or absence of built-up land.

Chapter 5 finds densification processes both in developing and developed regions, which are often associated with changes in 3D building structure. For example, the densification processes represented by the decreasing built-up land per capita usually correspond to changes in building footprints and/or heights. Specifically, the decreased built-up land per capita in deprived regions of Nairobi City could be attributed to the dense building footprints (Abascal et al., 2022). In regions such as Hong Kong and Singapore, however, the decreased built-up land per capita could be attributed to both dense building footprints and growing building heights (Lau et al., 2005; Xue et al., 2017). For the mountainous regions in Hong Kong, there is very little space for new buildings due to the topographical constraint, leading to higher buildings alternatively to host the growing population. Considering the rugged terrain as well as water bodies, rugged areas and water bodies are therefore excluded in Chapter 2 to calculate built-up land density that constitutes the properties used for the classification of settlement systems.

6.2 Implications for future research

6.2.1 Characterization of human settlements

Characterization of human settlements could benefit from and contribute to the developments in other fields to address global environmental and socioeconomic challenges. In the field of industrial ecology, for example, material stocks are considered as one of the major drivers of environmental crisis (Haberl et al., 2021). An assessment suggest that concrete, aggregates, and bricks, which are used for

buildings and infrastructure that characterize human settlements, constitute the vast majority of human-made mass (Elhacham et al., 2020). Spatially explicit characterization of human settlements by their building material have profound implications for sustainability, e.g., energy use (Samuel et al., 2013), human health (Latha et al., 2015), and risk assessment (Tierolf et al., 2021). As of today, however, there is no publicly available dataset on building material at a global or even local scale, hampering the more nuanced characterization of human settlements and the assessment of their environmental and socioeconomic impacts.

Technically, large-scale characterization of human settlements highly relies on the development of Earth Observation. Over the past few decades, Earth Observation technology has substantially help monitor the extent of human settlements, from local to global scales (Liu et al., 2018; Mundia and Aniya, 2005; Schneider et al., 2009; Van de Voorde et al., 2011). Yet, large-scale mapping of heterogeneity within human settlements is inherently different (Schug, 2021), and the usage of Earth Observation data alone could sometimes be improper (Rosier et al., 2022). In our case, mapping urban land-use intensity is dependent on the specific contexts (3D building structure and land take per capita). Moreover, Gross Domestic Product (GDP) can be used to quantify urban land-use intensity in the view of output dimension in measuring land-use intensity (Erb et al., 2013). GDP could be less relevant to 3D building structure, one of the other proxies for urban land-use intensity, which is mapped globally in Chapter 4 with Earth Observation and socioeconomic data, among others. In the era of big data, Earth Observation researchers are increasingly required to map human settlements using an integrated approach due to the multidimensionality of urban land use. Therefore, Earth Observation researchers could do more to collaborate with urban planners and practitioners to gain deeper insights into what forms urban heterogeneity (Wentz et al., 2018; Zhu et al., 2019). As a community, we need to think about how to derive the information (e.g., algorithms development), and also about what it will be used for (Zhu et al., 2019).

6.2.2 Drivers of human settlement dynamics

In the past decades, analysis of the drivers of urban expansion was at the core of urban land change science (Colsaet et al., 2018; Li et al., 2018a). This contributed to theory development, promoted the causal analysis of change, and gained insights into how human settlement development could be steered into societally more desirable pathways (Bürgi et al., 2022). However, this evidence is confined to the analysis of built-up land cover, without the nuanced representation of human settlements and their trajectories. To this point, this dissertation could be an addition in this field by

including settlement systems as well as urban land-use intensity, and a large question pertains to the drivers of changes in settlement systems and urban land-use intensity.

Human settlement dynamics and their drivers are in most cases investigated at a local or regional scale, and our understanding of forms and drivers of human settlement dynamics is therefore limited, as the evidence is often fragmented (Broitman and Koomen, 2020; Xiong et al., 2021; Yue et al., 2013). Systematic reviews, which synthesize results of case studies, have been widely conducted in land use science (van Vliet et al., 2016b). For example, van Vliet et al. (2015) conducted a systematic review on the manifestations and underlying drivers in terms of agricultural land change, which provided us with insight into mechanism in agricultural land change. Human settlements, as the most intense landscape altered by human beings, have been well discussed in case studies (Alig et al., 2004; Fang et al., 2016; Schneider et al., 2005). Yet, there is still no clear knowledge gains from the evidence-based synthesis. Therefore, there is an urgent need for a systematic review on drivers of changes in urban land-use intensity and settlement systems based on existing case studies.

Land use scientists can utilize hypothesis and theory from other fields and disciplines to better explain the conditions under which urban expansion and urban densification occur. For example, frontier development has been widely used in land science community, particularly in the field of agricultural expansion and deforestation (Meyfroidt et al., 2018). This concept was recently applied on the progress in biodiversity conservation, as the term "*conservation frontier*" which allows assessing the patterns, actors, and drivers of conservation effort (Buchadas et al., 2022). As found in Chapter 5, urban land-use intensity in the urban peripheries decreases over time, which I speculate that this decrease in urban land-use intensity can at least be partly attributed the abundance of land resource available for urban development, and local dwellers prefer to develop new land elsewhere for residency (expansion) rather than reduce built-up land per person (densification), given the low cost of developing new land therein.

6.2.3 Simulating dynamics of human settlements

Land use models are crucial tools for the integration of empirical knowledge with the practice of decision making (van Asselen and Verburg, 2013; Verburg et al., 2002). In-depth understanding the casual mechanism of human settlement dynamics enables land use models to better analyze land use change processes as well as to simulate trajectories of human settlements. However, previous studies have focused on improving the allocation mechanisms and calibration procedures (Li and Yeh, 2005;

van Vliet et al., 2016a; Wang et al., 2022), but progress in the nuanced representation of human settlements is rather limited. Technically, these models typically represent human settlements as one single land cover class, typically represented by built-up area. Therefore, the simulation of human settlements is limited to the unilateral conversion of non-built-up to built-up land. Chapter 2 characterized human settlements into several classes, ranging from *sparse villages* to *large cities*. With such a representation of diverse human settlements, historical trajectories are further explored. In contrast, these trajectories can also be integrated with land change models to simulate the gradual and incremental development of human settlements, rather than a sudden presence of built-up area.

Recently, land change studies have acknowledged the significance of nuanced changes in land-use intensity, and therefore the assessments of environmental and socioeconomic outcomes of different land change trajectories increasingly accumulated (Stürck et al., 2015; Tierolf et al., 2021; Wang et al., 2019b). For example, van Asselen and Verburg (2013) developed a CLUMondo model that is able to simulate agricultural intensification versus agricultural expansion, and urban versus rural settlement expansion based on land availability in the neighborhood of the location. Wang et al (2019) applied the CLUMondo model to present the nuanced interactions between human settlements and agricultural systems under multiple population scenarios and agricultural regulations. Recently, Chen (2022) developed a land change model that is capable to simulate the process of urban densification, thereby advancing conventional land change models that simply predict where the next urban pixels are likely to appear. Therefore, settlement systems and urban land-use intensity identified by this thesis provide the conceptual basis and technical supports for nuanced representation and trajectory analysis of human settlements in these land change models.

6.3 Implications for sustainable development

Human society is confronted with many sustainability challenges including food insecurity, and flood risks. To address these issues, policy tools are widely implemented by governments and local authorities. Particularly, cropland protection policies are adopted worldwide to preserve cropland on the periphery of human settlements by restricting built-up land expansion (Liu et al., 2014b; Perrin et al., 2018). Still, Zhou et al. (2021) found that only ~25% of cropland loss in China is due to the development of cities, while cropland is increasingly lost to rural settlements and other lands used for industry, mining, and transportation. The most prevailing land-cover-based products with a general classification of *built-up land* or *impervious surface*

cannot capture such variations in human uses, thus hindering the implementation of targeted policies in practices. Chapter 2 also found that most built-up land as well as newly added built-up land is not located in cities, suggesting that current policies focusing on the restriction of built-up land expansion in cities may fall short without consideration of land development in rural settlements, mining, and transportation, among others.

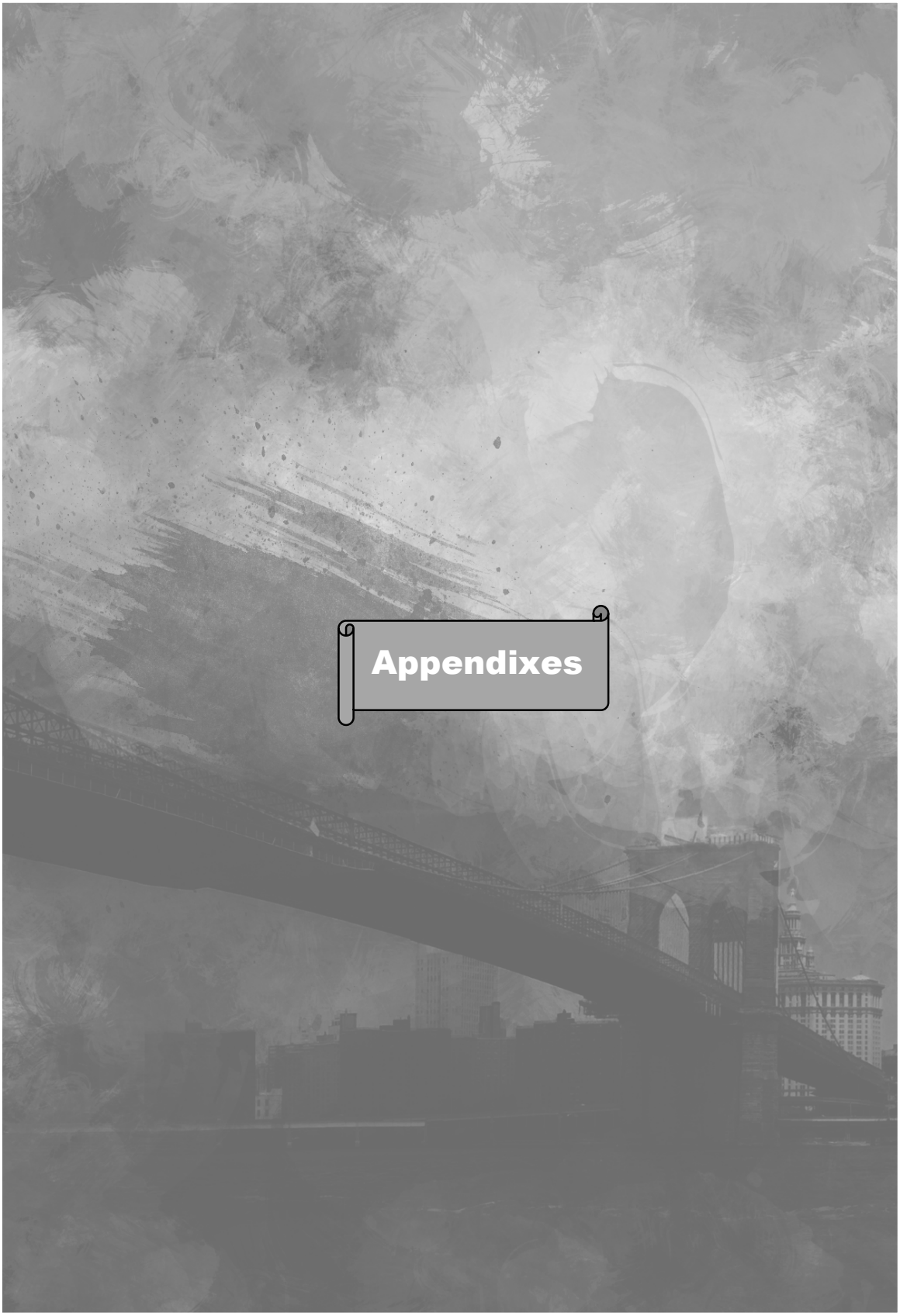
In addition, Chapters 3-5 found that urban land-use intensity, expressed as building structure or land take per person, varies substantially worldwide. Therefore, policies towards sustainable development in human settlements are supposed to account for specific contexts with regard to environmental and socioeconomic heterogeneity. In countries like the US, where residential density is relatively low, local authorities should constrain urban sprawl in a practical manner. Yet, in densely populated regions of the Global South like India and African cities, compact development is likely to have undesirable outcomes.

6.4 Concluding remark

This thesis aims to advance large-scale analysis of human settlements and their dynamics through the lens of land system with a specific focus on the role of land-use intensity. The four preceding research Chapters 2-5 provide an exploration to answering the three research questions proposed according to the objective of this dissertation.

Mega-cities have attracted wide attention from urban land-use scientists. However, a large proportion of built-up as well as new added built-up land is found in more rural regions (Chapter 2), which is further supported by the disproportional conversions from deep rural to towns, and from towns to peri-urban and to large cities. Other than the difference in horizontal pattern, 3D building structure exhibits a large variation globally, and the global maps of 3D building structure produced here provide an image of the heterogeneity within built-up areas (Chapters 3 and 4), thereby complementing global datasets on urban heterogeneity (for example, built-up land extent, and gridded population). Finally, temporal changes of urban land-use intensity are put in a global context and found that urban land expands faster than population growth (Chapter 5). Results show that while built-up land increases faster than the population at a global scale, specifically, most densification is found in large cities as well as in many regions in the Global South, where urban land is already in intensive use.

In short, this thesis reveals a large variation in urban land-use intensity, which allows for more nuanced analyses of human settlement dynamics. For integrated assessment models, the representation of urban land-use change can incorporate such variation to inform decision-making.


Appendixes

Appendix A

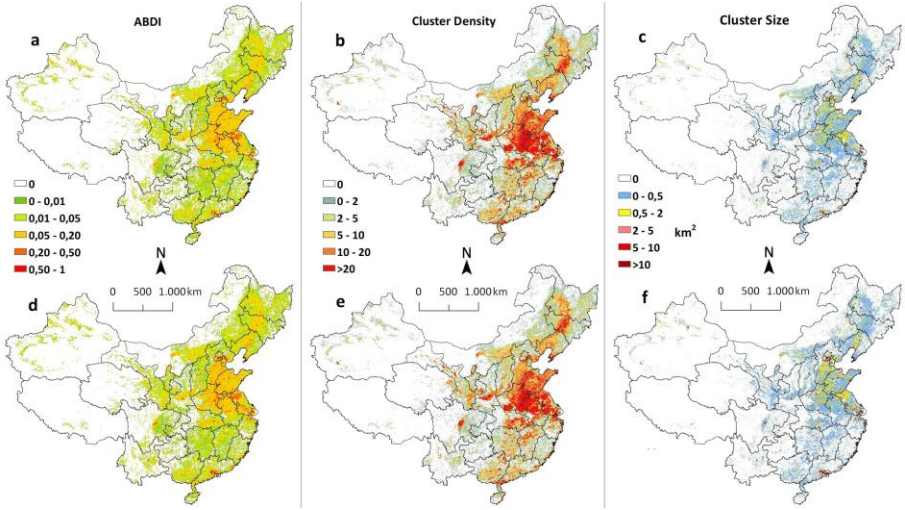


Figure A1: ABDI, cluster density and cluster size for 1990 and 2000 (a-c, and d-f respectively).

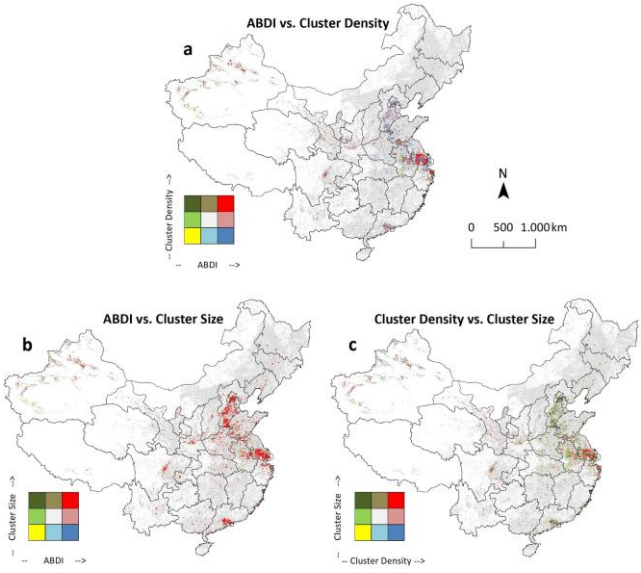


Figure A2: Relations between changes in ABDI, cluster density, and cluster size between 1990 and 2000.

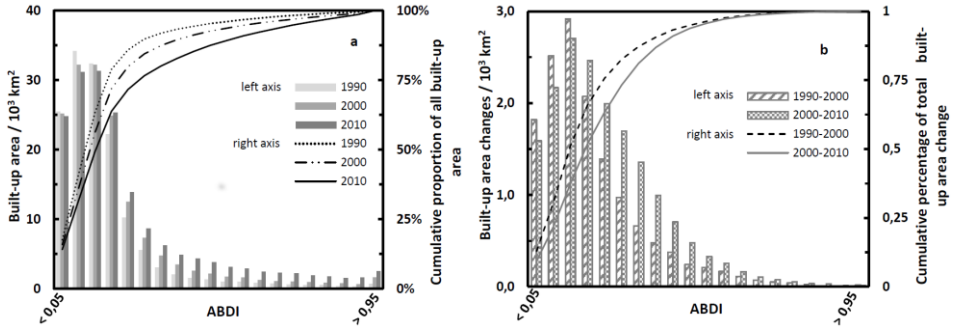


Figure A3: Distribution of built-up area in 1990, 2000, and 2010 (a) and its change during 1990-2000 and 2000-2010 (b).

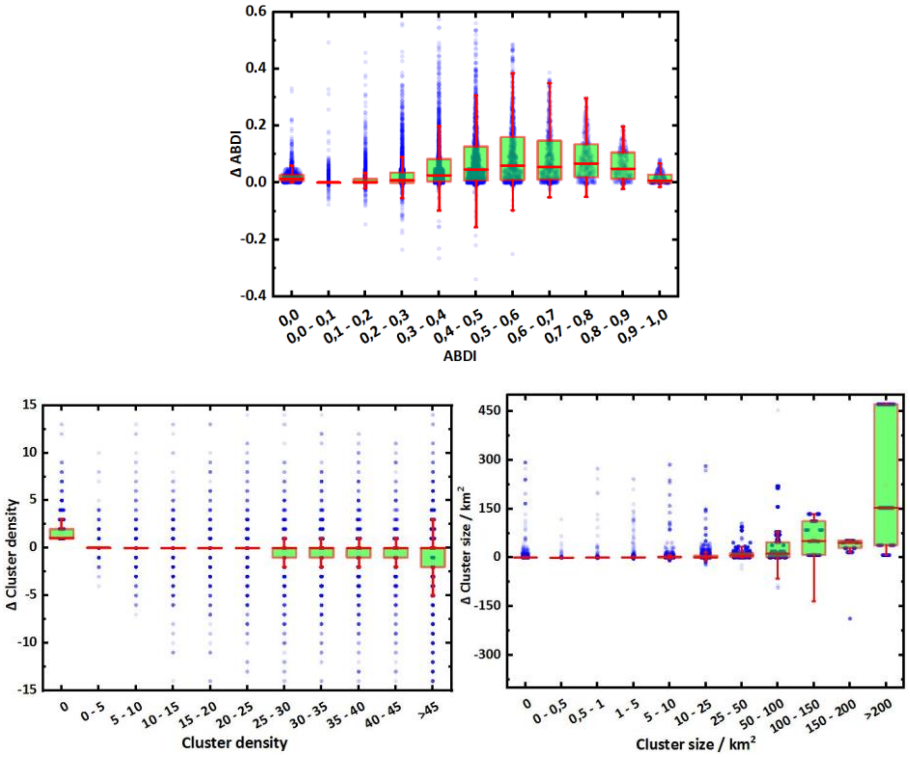


Figure A4: Changes in ABDI, cluster density and cluster size as a function of their initial value during 1990-2000. The horizontal bars in each boxplot correspond to the 25th, 50th, and 75th percentiles. The whiskers extend to 1.5 times the interquartile ranges (IQR).

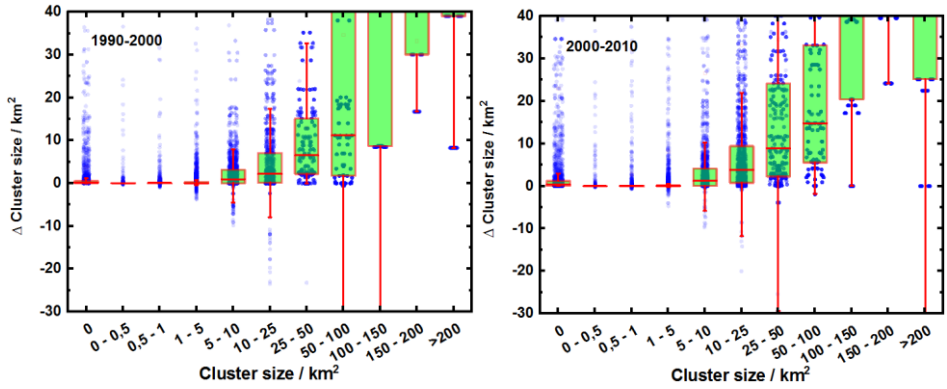


Figure A5: Boxplots of cluster size changes in relation to initial value. Y axis is capped at -30 and 50, to allow comparison and to increase visibility.

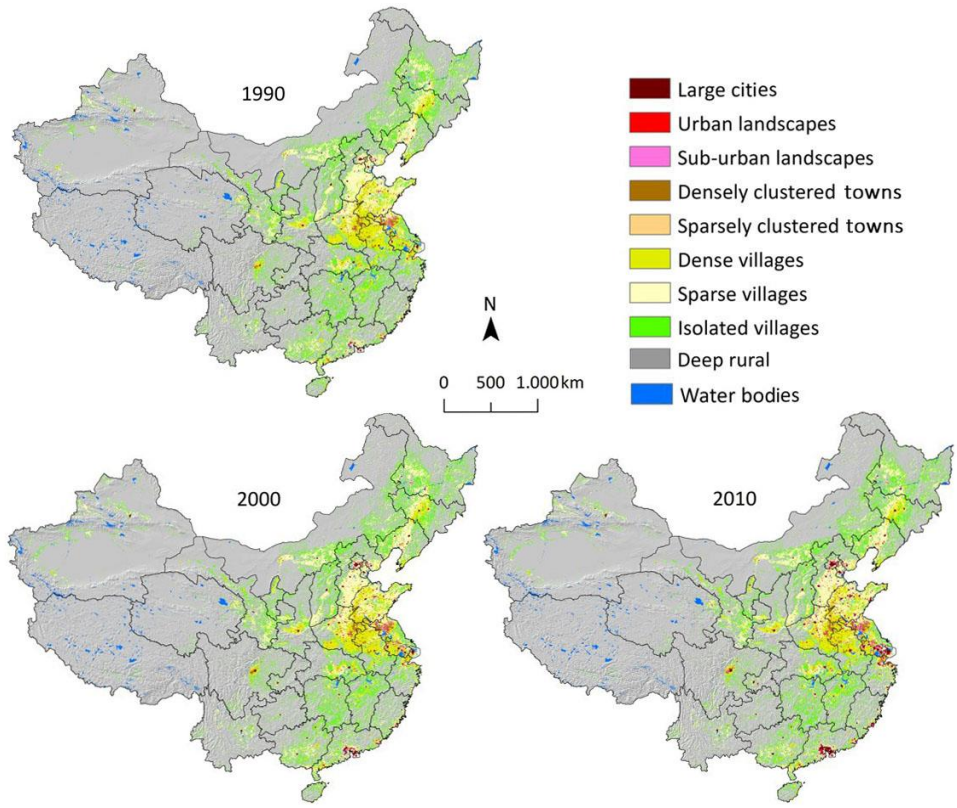


Figure A6: Settlement systems for 1990, 2000 and 2010.

Appendix B

In order to reduce the bias of reference data for the Random Forest (RF) model, we manually generated reference data for 6163 sample points across all three continents. For this we used a visual interpretation of aerial images, in combination with a visual interpretation of Google street view images for quantifying building height. First, we vectored building footprint in 100 randomly selected images manually using ArcGIS 10.4.1, in order to assess the reliability of a simpler mapping based on a grid of 20×20 meters. Results, as shown in Figure B1 show a very good match justifying the use of the simpler method for classifying built-up area footprint. Subsequently, we used the values in Table B1 for the interpretation of building height based on street view imagery of buildings included. Figures B2 and B3 provide an example sample location included in the reference data.

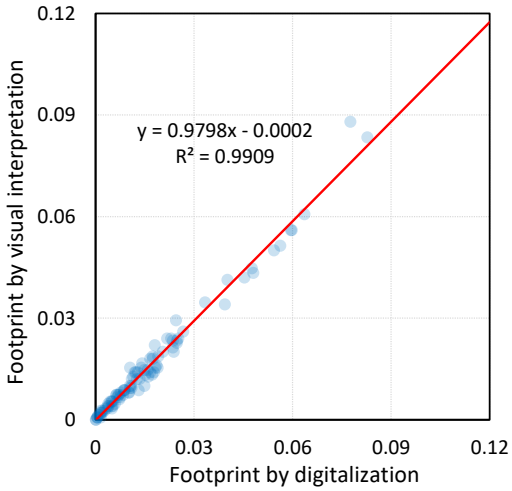


Figure B1: Reliability analysis of building footprint by visual interpretation compared to manual digitalization.

Table B1: Guidelines for building height assignment based on google street view.

Description	Height (m)
all 1-floor	3
dominantly 1-floor, with a few 2-floor buildings	4
dominantly 2-floor, with a few 1-floor buildings	5
all 2-floor	6
dominantly 2-floor, with a few 3-floor buildings	7
dominantly 3-floor, with a few 2-floor buildings	8
all 3-floor	9
dominantly 3-floor, with a few 4-floor buildings	10
dominantly 4-floor, with a few 3-floor buildings	11
…	…

Figure B2: An example of the valid VHR satellite imageries in France (47°47'55.8"N, 0°00'25.4"W). Each small square is 20×20m, and in this imagery we can observe 3 building clusters (encircled A, B and C), while the numbers in white indicate the estimation of square numbers occupied by building footprint for corresponding cluster.

Figure B3: Google street view in sites A, B and C, as signed in Figure B2. In this case, buildings are dominantly one-floor high. Besides, we can see triangular roofs on top of these buildings, thereby we assume there is additional space for living and a value of 4 m is assigned as the averaged building height as suggested in Table B1.

Figure B4 showed that for the RF models trained to map building volume and height differ between individual case region applications and all case regions combined, in terms of the relative importance of VV and VH polarized SAR data. Therefore, we further tested our analysis by running all RF models with either VV polarized SAR data or VH polarized SAR data. Results from Figure B4 show that VV and VH SAR data are highly complementary, as, both achieve highly similar R^2 values. Moreover, results of the RF models with only one of the SAR polarization modes included are almost as accurate as the RF model with both bands included. This complementarity shows why the relative importance of VV and VH in classifying building height and volume could change very much while still achieving a higher overall R^2 for all data combined, as compared to the RF models for each region specifically. Figure B5 further illustrates this by showing the variable importance corresponding to the results shown in Figure B4 based on the data for all regions combined.

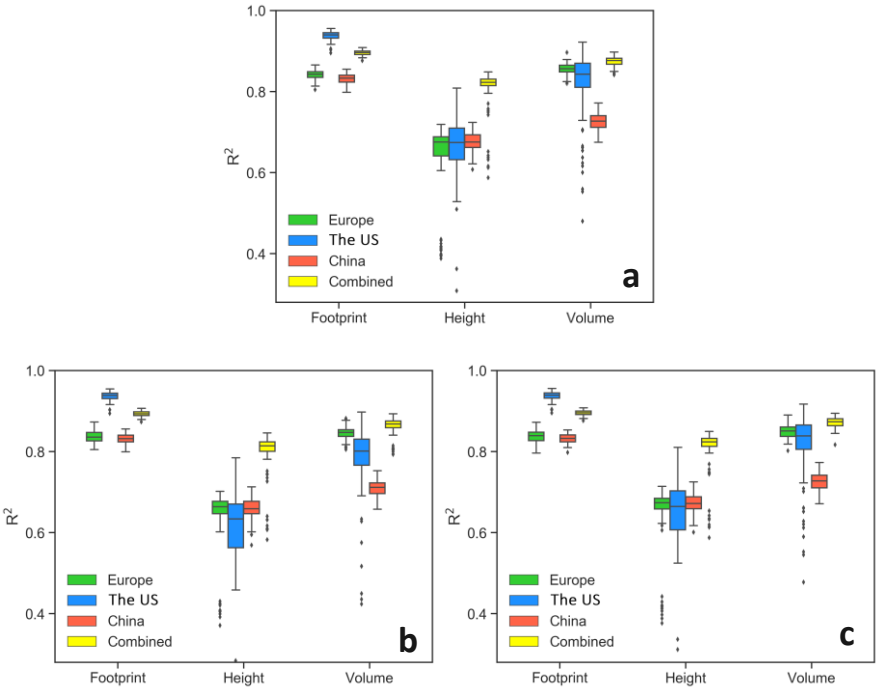


Figure B4: Model performance comparison in cases that one of the SAR polarization modes is absent. (a) all variables included; (b) no-VH; and (c) no-VV.

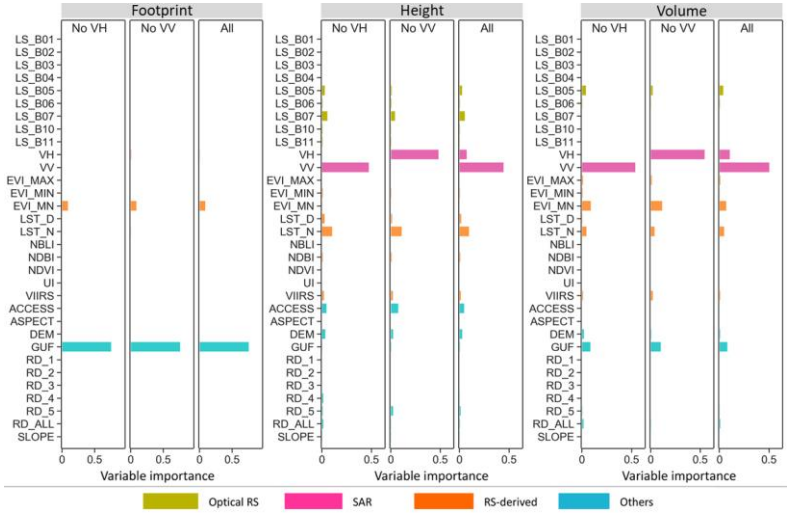


Figure B5: Normalized variable importance in cases that one of the SAR polarization modes is absent. Variable importance is identified based on best-fitted run. All reference data for the three case regions are combined in this analysis. "All" indicates that both VH and VV are included in the process model training and validation.

Table B2: Collection of 3D building dataset for the US used for model training and validation.

US City	Source
Southeast Michigan	https://maps.semcog.org/BuildingFootprints
Chicago	https://data.cityofchicago.org/Buildings/Building-Footprints-current-/hz9b-7nh8
LA	https://egis3.lacounty.gov/dataportal/2016/11/03/countywide-building-outlines-2014-update-public-domain-release/
Boulder	https://hub.arcgis.com/datasets/0d43652d038a4a0dbca68f0501151bb0_0
Fort Collins	http://hub.arcgis.com/datasets/7e577a14c83f4d83a6b58657c48027da_0/data?geometry=-105.502%2C40.489%2C-
Norman	http://hub.arcgis.com/datasets/d68b0defa057465db7167d9260c90ad9_0
Austin	https://data.austintexas.gov/browse?q=footprint&sortBy=relevance
Albuquerque	http://data-cabq.opendata.arcgis.com/datasets/e65e375b680345e0b21fa7585d83ce9c_0?uiTab=table
New York	https://data.cityofnewyork.us/Housing-Development/Building-Footprints/nqwf-w8eh
San Francisco	https://data.sfgov.org/Housing-and-Buildings/Building-Footprints-File-Geodatabase-Format-/asx6-3trm
Miami	https://mdc.maps.arcgis.com/home/item.html?id=ab4d3a61e60c441bbfc1098d701fc991
Sarasota	http://hub.arcgis.com/datasets/6c679d2949544274aee3bee8182c5611_0?geometry=-84.608%2C29.432%2C-
Boston	http://boston.maps.arcgis.com/home/item.html?id=c423eda7a64b49c98a9ebdf5a6b7e135
Roanoke	http://hub.arcgis.com/datasets/198c95ddd5f749ca9fc851dd64ba6ff0_32?geometry=-81.388%2C37.714%2C-
Santa_Clara	http://hub.arcgis.com/datasets/ee83a3518a7249fda22866117463de3f_0?page=10
Reedsburg	http://hub.arcgis.com/datasets/dbe64a71897e4982934dbd7637d576d5_0?geometry=-91.204%2C43.336%2C-
Macomb	http://hub.arcgis.com/datasets/b5bbaf4fef6e4b59b8214ccaa17b8331_0
Washtenaw	http://hub.arcgis.com/datasets/beefad4ecf334b43b883123a72bf86b7_0
Henderson	http://hub.arcgis.com/datasets/23e5f3506f034c3d99b84e54fce51584_11
Lincoln	http://hub.arcgis.com/datasets/1b6a5a2ef1b34c28950c4e720e8d7a3d_0
Monroe	http://hub.arcgis.com/datasets/61e3bf9f7da143b3b87e17dabe5b0c52_0
St._Clair	http://hub.arcgis.com/datasets/d00d7cffefd6466086dcfc7202f185f4_0/data?orderBy=MEDIAN_HGT&page=109
Sioux Falls	https://hub.arcgis.com/datasets/065e40f79b784848b403130234d95a1e_5?geometry=-97.326%2C43.445%2C-
Bernalillo	https://hub.arcgis.com/datasets/e65e375b680345e0b21fa7585d83ce9c_0
Arlington	https://hub.arcgis.com/datasets/bac045c94c144838a7e65fbcf7aa939c_0?page=1532
DC	http://opendata.dc.gov/datasets?q=3d&sort_by=relevance
Cambridge	https://www.cambridgema.gov/GIS/gisdatadictionary/Basemap/BASEMAP_Buildings.aspx

Appendix C

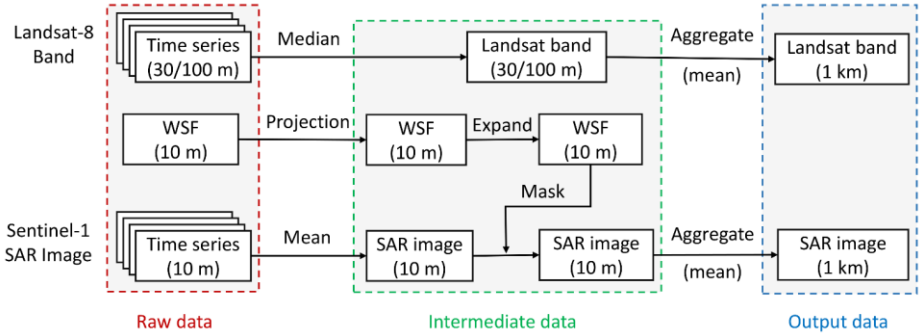


Figure C1: Processing chain of Landsat and SAR data. World Settlement Footprint (WSF-2015) data is used as the source to constrain SAR backscatter signals to settlement ambience.

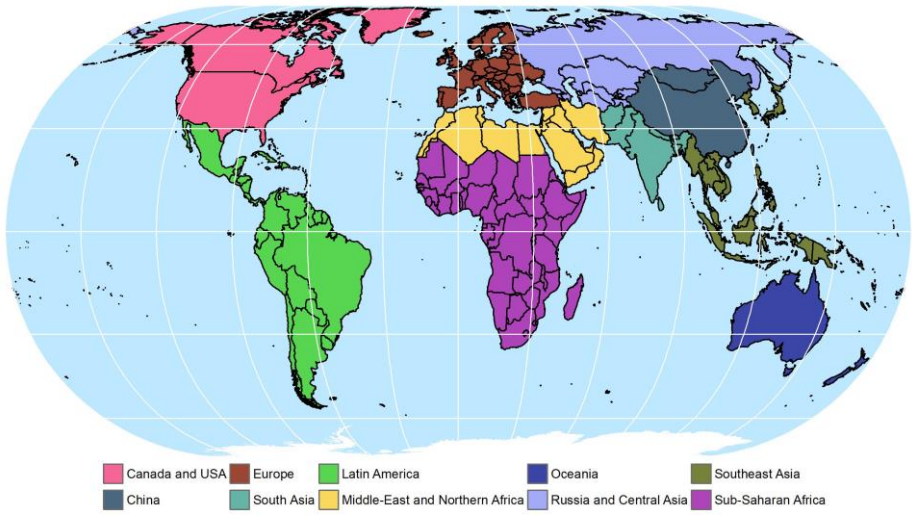


Figure C2: Ten world regions used for presenting aggregated results in this study. The world regions are World Bank regions, except for (*i*) East Asia and Pacific and (*ii*) Europe and Central Asia, which were further subdivided to reflect the differences in both urban expansion and socioeconomic development. Specifically, the East Asia and Pacific region was further subdivided into China, Southeast Asia, and Oceania, while Europe and Central Asia was divided in Europe, and Russia and Central Asia. Country boundaries are added for further spatial reference.

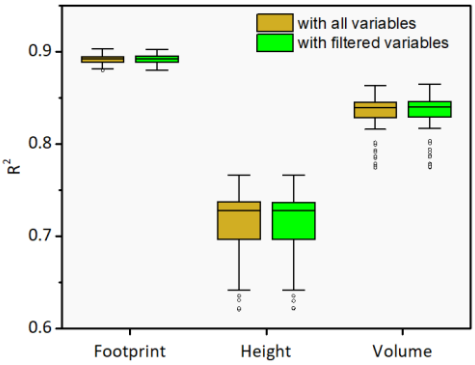


Figure C3: Model performance before and after removing less important variables. This graph compares classification accuracies before and after this procedure, based on the 100 individual models in each model ensemble. Each of the 100 R^2 values is calculated based on the comparison of observations to predictions of the 10% validation subset. We removed explanatory variables with less than 0.5% variable importance in order to avoid overfitting. The paired boxplots suggest a negligible effect of excluding less important variables on model performance.

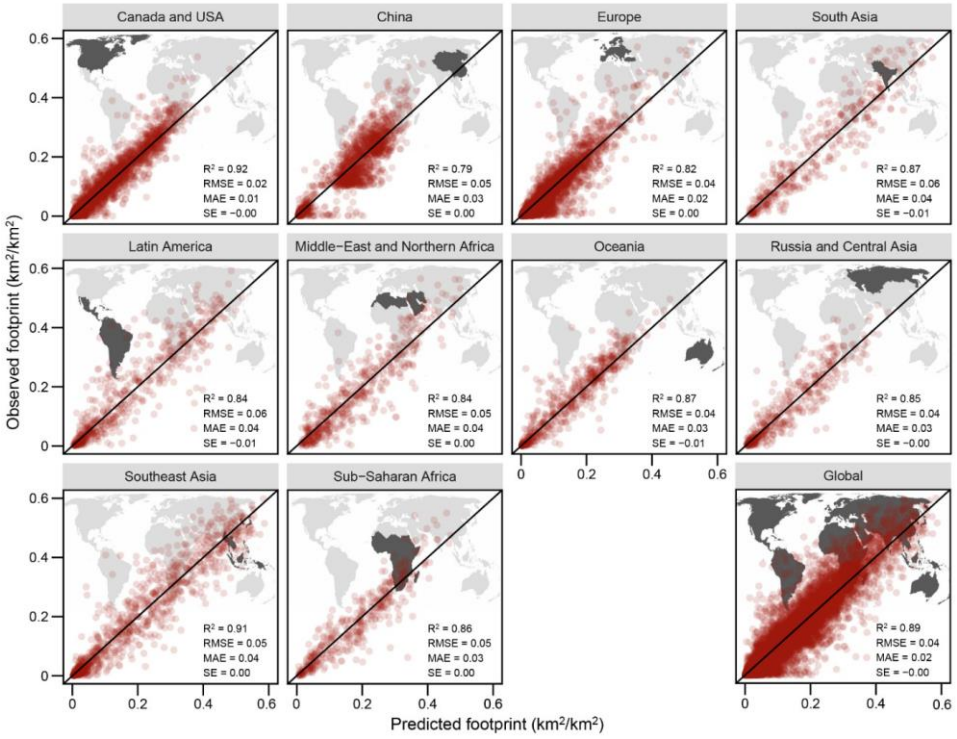


Figure C4: Scatterplots the predicted values v.s. observed values for building footprint in the independent 20% test subset. Predicted values represent the mean of 100 predicted values for each location in the test subset.

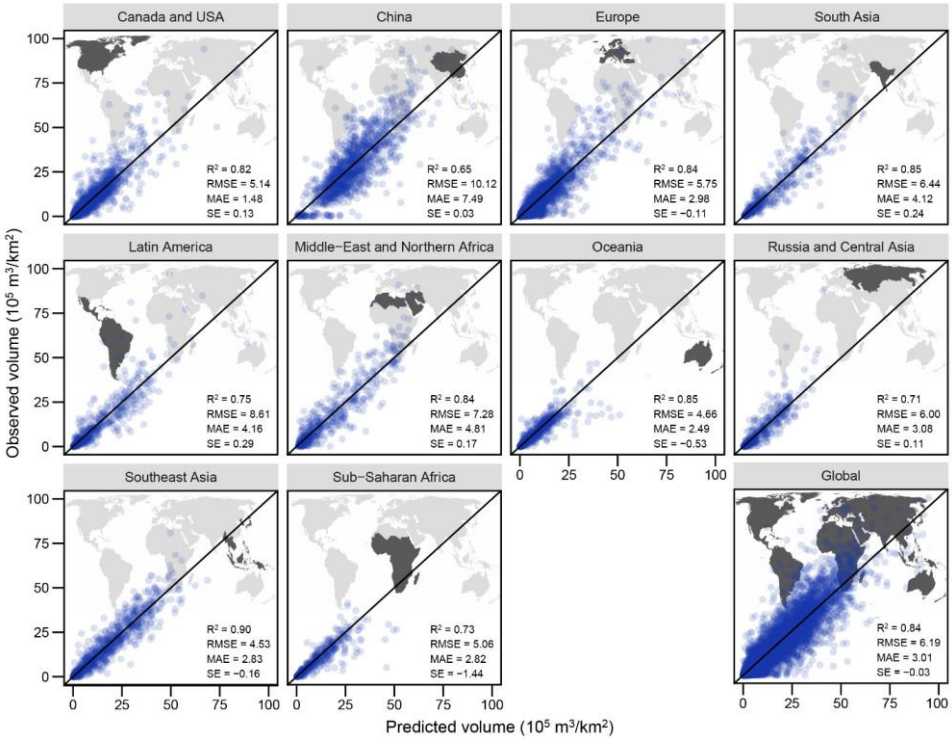


Figure C5: Scatterplots the predicted values v.s. observed values for building volume in the independent 20% test subset. Predicted values represent the mean of 100 predicted values for each location in the test subset.

Table C1: Data sourced used in this study.

Dataset		Resolution	Period	Source	Abbreviation
Optical RS	Landsat Band 1	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B01
	Landsat Band 2	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B02
	Landsat Band 3	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B03
	Landsat Band 4	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B04
	Landsat Band 5	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B05
	Landsat Band 6	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B06
	Landsat Band 7	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B07
	Landsat Band 10	100 m	2015	https://landsat.gsfc.nasa.gov/	LS-B10
	Landsat Band 11	30 m	2015	https://landsat.gsfc.nasa.gov/	LS-B11
SAR	Sentinel-1 VH	10 m	~2015	https://sentinel.esa.int/	SAR-VH
	Sentinel-1 VH	10 m	~2015	https://sentinel.esa.int/	SAR-VV
	Sentinel-1 MAX	10 m	~2015	https://sentinel.esa.int/	SAR-MAX
RS-derived	EVI maximum	1 km	2015	https://modis.gsfc.nasa.gov/	EVI-MAX
	EVI mean	1 km	2015	https://modis.gsfc.nasa.gov/	EVI-MN
	EVI minimum	1 km	2015	https://modis.gsfc.nasa.gov/	EVI-MIN
	LST day	1 km	2015	https://lpdaac.usgs.gov/	LST-D
	LST night	1 km	2015	https://lpdaac.usgs.gov/	LST-N
	NBLI	30 m	2015	Landsat	NBLI
	NDBI	30 m	2015	Landsat	NDBI
	NDVI	30 m	2015	Landsat	NDVI
	UI	30 m	2015	Landsat	UI
	Nighttime light	500 m	2015	https://earthdata.nasa.gov/	NTL
Others	Accessibility	1 km	2015	Weiss et al. (2018)	ACCESS
	Aspect	250 m	2010	https://www.usgs.gov/	ASPECT
	DEM	250 m	2010	https://www.usgs.gov/	DEM
	WSF	10 m	2015	Marconcini et al. (2020)	WSF
	Highways	Vector	2015	Meijer et al. (2018)	RD-1
	Primary roads	Vector	2015	Meijer et al. (2018)	RD-2
	Secondary roads	Vector	2015	Meijer et al. (2018)	RD-3
	Tertiary roads	Vector	2015	Meijer et al. (2018)	RD-4
	Local roads	Vector	2015	Meijer et al. (2018)	RD-5
	All (+unclassified)	Vector	2015	Meijer et al. (2018)	RD-ALL
	Slope	250 m	2010	https://www.usgs.gov/	SLOPE
	GDP per capita	Vector	~2015	Gennaioli et al. (2013)	GDP
	Gini coefficient	Vector	~2015	Solt (2020)	GINI

Table C2: Number of 1-km^2 grids in the training samples as well as in the predicted data points for each world regions. Sub-total numbers refer to the combination of data collected from multiple sources, and interpreted results refer to manually generated reference data for each world region (see Figure 4.2, and Figure C2). The percentages in bottom row indicate the ratios of training samples to total grids to be predicted.

World region	# of predicted data points	# of training samples					
		Footprint		Height		Volume	
		Sub-total	Interpreted	Sub-total	Interpreted	Sub-total	Interpreted
Canada and USA	2336135	32035	1685	26236	1685	26152	1685
China	1690674	8409	2595	6185	385	6185	385
Europe	2311616	16531	2803	16531	2800	16531	2800
Southern Asia	786108	2161	70	2161	70	2161	70
Latin America	427087	3271	1169	3271	1169	3271	1169
Middle-East and Northern Africa	275288	2701	33	2701	33	2701	33
Oceania	266232	3234	549	3234	549	3234	549
Russia and Central Asia	588014	2248	458	2248	458	2248	458
Southeast Asia	913497	6390	2278	6390	2278	6390	2278
Sub-Saharan Africa	817985	2206	243	2206	243	2206	243
Total	10412636	79186	11883	71163	9670	71079	9670
		0.76%		0.68%		0.68%	

Table C3: Variable importance values for explanatory variables before and after selecting variables that have an importance value larger than 0.5%. Table contents show the mean values of variable importance for the 100 models in a model ensemble, followed by their standard deviations. Please refer to Table C1 for variable abbreviations.

Dataset		Original importance value (%)			Filtered importance value (%)		
		Footprint	Height	Volume	Footprint	Height	Volume
Optical RS	LS-B01	0.22 ± 0.01	0.63 ± 0.02	0.30 ± 0.01	-	0.66 ± 0.02	-
	LS-B02	0.18 ± 0.00	0.47 ± 0.03	0.25 ± 0.01	-	0.54 ± 0.03	-
	LS-B03	0.20 ± 0.01	0.47 ± 0.04	0.27 ± 0.02	-	0.72 ± 0.06	-
	LS-B04	0.14 ± 0.01	0.37 ± 0.04	0.16 ± 0.01	-	-	-
	LS-B05	0.97 ± 0.08	12.63 ± 0.16	6.29 ± 0.64	1.14 ± 0.08	12.67 ± 0.16	6.49 ± 0.63
	LS-B06	0.65 ± 0.03	1.59 ± 0.12	0.46 ± 0.03	0.89 ± 0.04	1.63 ± 0.12	0.71 ± 0.04
	LS-B07	0.31 ± 0.01	1.37 ± 0.08	0.38 ± 0.03	-	1.43 ± 0.08	-
	LS-B10	0.38 ± 0.02	1.29 ± 0.09	0.46 ± 0.02	0.70 ± 0.02	1.31 ± 0.09	0.52 ± 0.03
	LS-B11	0.35 ± 0.02	0.97 ± 0.05	0.75 ± 0.06	-	0.99 ± 0.05	0.82 ± 0.06
SAR	SAR-VH	0.97 ± 0.03	5.27 ± 0.39	8.37 ± 0.73	1.08 ± 0.03	5.32 ± 0.39	8.51 ±0.73
	SAR-VV	1.98 ± 0.11	2.75 ± 0.24	5.51 ± 0.51	2.05 ± 0.11	2.77 ± 0.23	5.57 ± 0.51
	SAR-MAX	2.41 ± 0.11	2.88 ± 0.21	6.62 ± 0.54	2.48 ± 0.11	2.90 ± 0.21	6.69 ± 0.53
RS-derived	EVI-MAX	0.54 ± 0.03	0.80 ± 0.04	1.83 ± 0.31	0.63 ± 0.03	0.83 ± 0.04	1.90 ± 0.32
	EVI-MN	0.77 ± 0.04	1.08 ± 0.09	6.60 ± 0.68	0.85 ±0.04	1.11 ± 0.10	6.68 ± 0.70
	EVI-MIN	0.46 ± 0.02	1.30 ± 0.07	1.04 ± 0.26	0.56 ± 0.02	1.34 ± 0.08	1.13 ± 0.25
	LST-D	2.28 ± 0.14	5.05 ± 0.15	0.72 ± 0.04	2.40 ±0.14	5.09 ± 0.14	0.83 ± 0.04
	LST-N	2.26 ± 0.21	3.66 ± 0.21	1.97 ± 0.07	2.39 ± 0.21	3.71 ± 0.21	2.10 ± 0.07

	NBLI	0.14 ± 0.01	0.39 ± 0.04	0.16 ± 0.01	-	-	-
	NDBI	2.80 ± 0.82	0.77 ± 0.03	0.51 ± 0.03	2.91 ± 0.82	0.79 ± 0.03	0.57 ± 0.03
	NDVI	0.53 ± 0.03	0.81 ± 0.05	0.46 ± 0.02	0.70 ± 0.03	0.91 ± 0.05	0.64 ± 0.03
	UI	7.88 ± 0.86	0.86 ± 0.03	1.01 ± 0.07	8.02 ± 0.86	0.88 ± 0.03	1.09 ± 0.07
	NTL	1.00 ± 0.04	29.65 ± 0.18	1.98 ± 0.11	1.15 ± 0.05	29.73 ± 0.18	2.11 ± 0.11
Others	ACCESS	0.72 ± 0.06	2.95 ± 0.09	1.63 ± 0.34	0.73 ± 0.06	2.96 ± 0.09	1.63 ± 0.34
	ASPECT	0.37 ± 0.01	0.79 ± 0.03	0.46 ± 0.02	-	0.81 ± 0.03	0.53 ± 0.02
	DEM	0.64 ± 0.02	3.12 ± 0.16	1.42 ± 0.05	0.79 ± 0.02	3.17 ± 0.16	1.56 ± 0.05
	WSF	61.70 ± 0.15	0.83 ± 0.03	43.30 ± 0.66	61.97 ± 0.15	0.86 ± 0.03	43.45 ± 0.66
	RD-1	0.10 ± 0.01	0.20 ± 0.02	0.07 ± 0.01	-	-	-
	RD-2	0.15 ± 0.00	0.44 ± 0.02	0.25 ± 0.01	-	-	-
	RD-3	0.19 ± 0.01	0.79 ± 0.07	0.42 ± 0.02	-	0.80 ± 0.07	-
	RD-4	0.20 ± 0.01	2.69 ± 0.47	0.33 ± 0.01	-	2.71 ± 0.47	-
	RD-5	0.61 ± 0.02	1.43 ± 0.09	1.25 ± 0.14	0.69 ± 0.02	1.46 ± 0.09	1.34 ± 0.14
	RD-ALL	1.29 ± 0.03	1.33 ± 0.09	1.72 ± 0.18	1.43 ± 0.03	1.39 ± 0.09	1.87 ± 0.19
	SLOPE	0.36 ± 0.01	1.05 ± 0.03	0.53 ± 0.03	-	1.09 ± 0.03	0.62 ± 0.03
	GDP	0.71 ± 0.04	5.36 ± 0.75	1.08 ± 0.06	0.79 ± 0.04	5.40 ± 0.75	1.16 ± 0.06
	GINI	5.55 ± 0.22	3.98 ± 0.18	1.41 ± 0.12	5.65 ± 0.22	4.00 ± 0.18	1.48 ± 0.12

Table C4: Comparison of WSF-3D dataset with the global results produced this study. Note that accuracy metrics are inherently not comparable due to the difference in cell size, according to which we convert metric values of building volume reported in Esch et al. (2022), allowing for better comparison.

Measure	WSF-3D			This study		
	Footprint (%)	Height (m)	Volume ($\times 10^5$ m^3/km^2)	Footprint (km^2/km^2)	Height (m)	Volume ($\times 10^5$ m^3/km^2)
RMSE	14.09	6.04	15.39	0.04	2.56	6.19
MAE	10.24	3.56	8.49	0.02	1.39	3.01
SE	3.06	-2.30	-3.47	0.00	-0.05	-0.03

Appendix D

To examine the impact of different ways of attributing new built-up land to changes in population and changes in BPC, we tested several options. In the main paper, we decompose the total change in built-up land into changes in population and changes in BPC, so that the multiplication of both equals the total change in built-up land. Here we explain three other methods of attributing changes in built-up land in order to assess the sensitivity for these attributions.

The first alternative is similar to the approach for attributing crop production change to either cropland expansion or yield increase, used for example in Eitelberg et al. (2016) and Wang et al. (2019b). This approach could be relevant for SDG target 11.3, which states that the rate of increase in built-up land should not exceed the rate of increase in population, as it effectively takes the BPC in the starting year ($\mathrm{BPC_0}$) of each analysis period as the baseline. Specifically, the area of built-up land change in a given region due to changes in population ($\mathrm{A_{POP}}$) and due to changes in BPC ($\mathrm{A_{BPC}}$), respectively, can be calculated as:

$$\mathrm{A_{POP}} = \mathrm{BPC_0} \times (\mathrm{POP_1} - \mathrm{POP_0}) \qquad \text{(D1)}$$

$$\mathrm{A_{BPC}} = \mathrm{POP_1} \times (\mathrm{BPC_1} - \mathrm{BPC_0}) \qquad \text{(D2)}$$

Conversely, Eq. (D3) and (D4) use $\mathrm{BPC_1}$ as a reference level, thereby effectively assuming that increased/decreased population individually consumes the same amount of built-up land in end year of this period:

$$\mathrm{A_{POP}} = \mathrm{BPC_1} \times (\mathrm{POP_1} - \mathrm{POP_0}) \qquad \text{(D3)}$$

$$\mathrm{A_{BPC}} = \mathrm{POP_0} \times (\mathrm{BPC_1} - \mathrm{BPC_0}) \qquad \text{(D4)}$$

Eq. (D5) and (D6) use a reference level dependent on the direction of population change: increased population individually consumes the same amount of built-up land in end year of this period. In contrast, decreased population individually consumes the same amount of built-up land in starting year of this period.

$$\mathrm{A_{POP}} = f(\mathrm{BPC}) \times (\mathrm{POP_1} - \mathrm{POP_0}) \qquad \text{(D5)}$$

$$A_{BPC} = f(POP) \times (BPC_1 - BPC_0) \qquad (D6)$$

Here, if $POP_1 - POP_0 \geq 0$, then $f(BPC) = BPC_1$, and $f(POP) = POP_0$; Otherwise, $f(BPC) = BPC_0$, and $f(POP) = POP_1$.

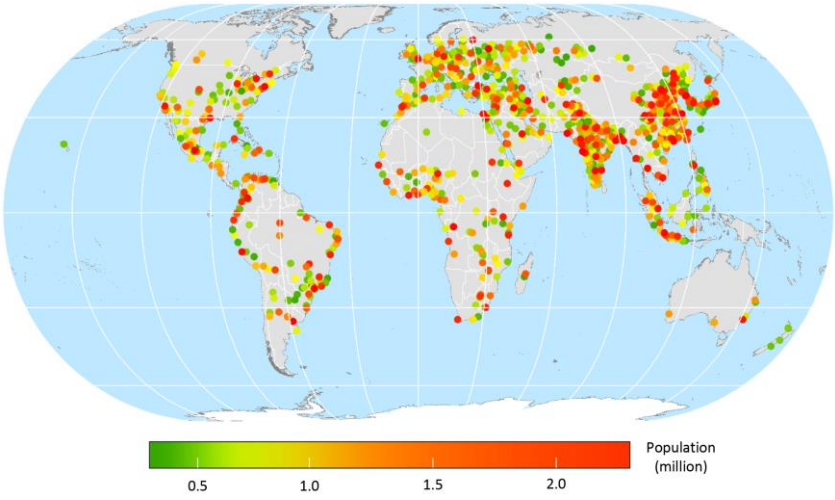


Figure D1: Distribution of cities with a population size large than 0.3 million.

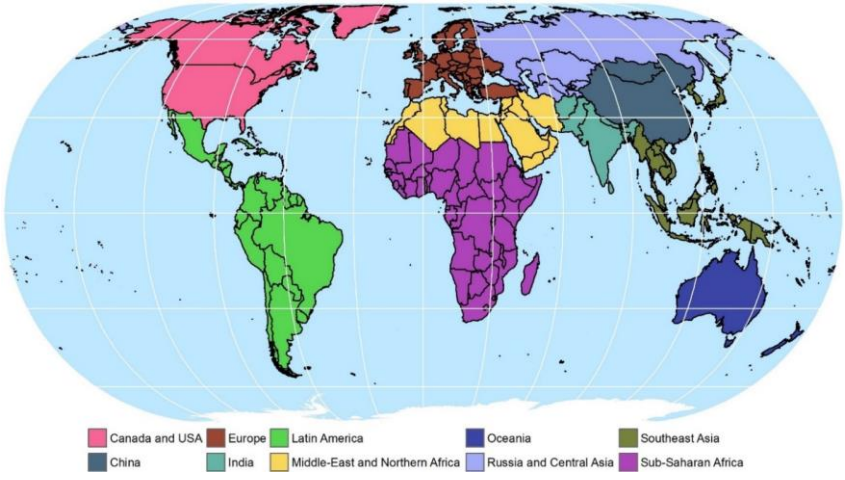


Figure D2: Ten world regions used in this article. Country boundaries are added for further spatial reference.

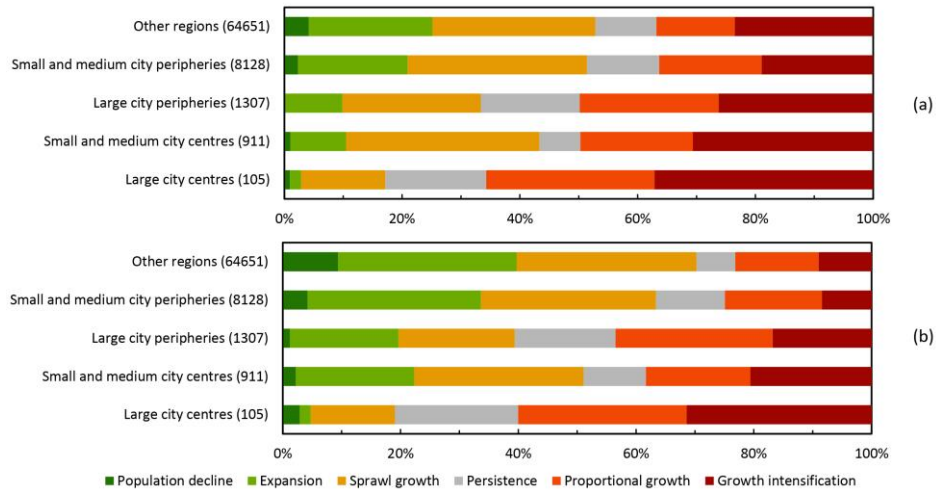


Figure D3: Distribution of built-up land change trajectories over the different types of regions for the periods a) 1975-1990, and b)1990-2000.

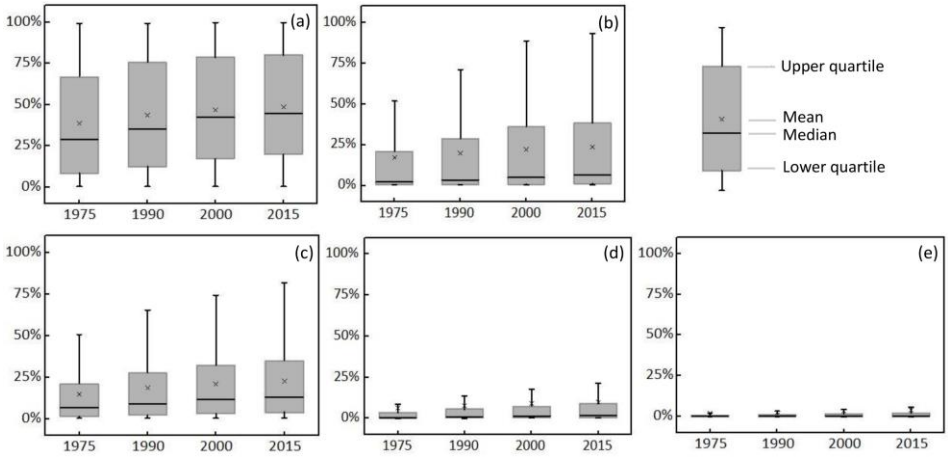


Figure D4: Percentage of built-up land to total available land in each region class. (a) large city centres; (b) large city peripheries; (c) small and medium city centres; (d) small and medium city peripheries; (e) other regions. Each boxplot is based on all regions that are classified as the specific region class, and average are taken over all regions within this class without weighing for their area of amount of built-up land. Calculations are based on the ~38 m GHSL dataset, in which water layers are excluded already.

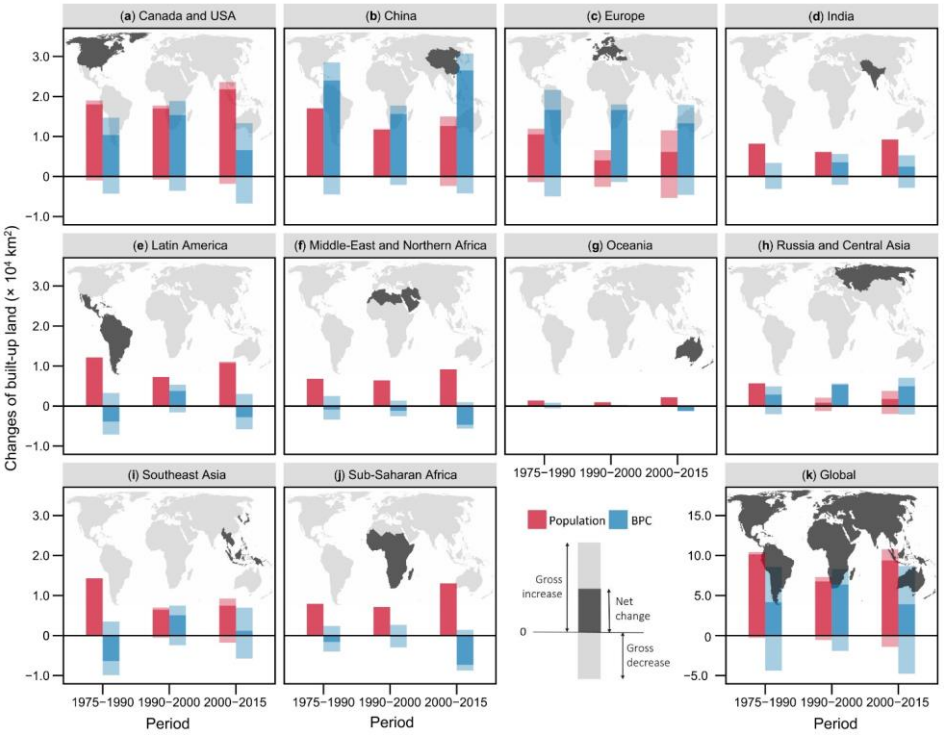


Figure D5: Heterogeneity in built-up land area changes. A_{POP} and A_{BPC} are calculated following Eq. (D1) and (D2).

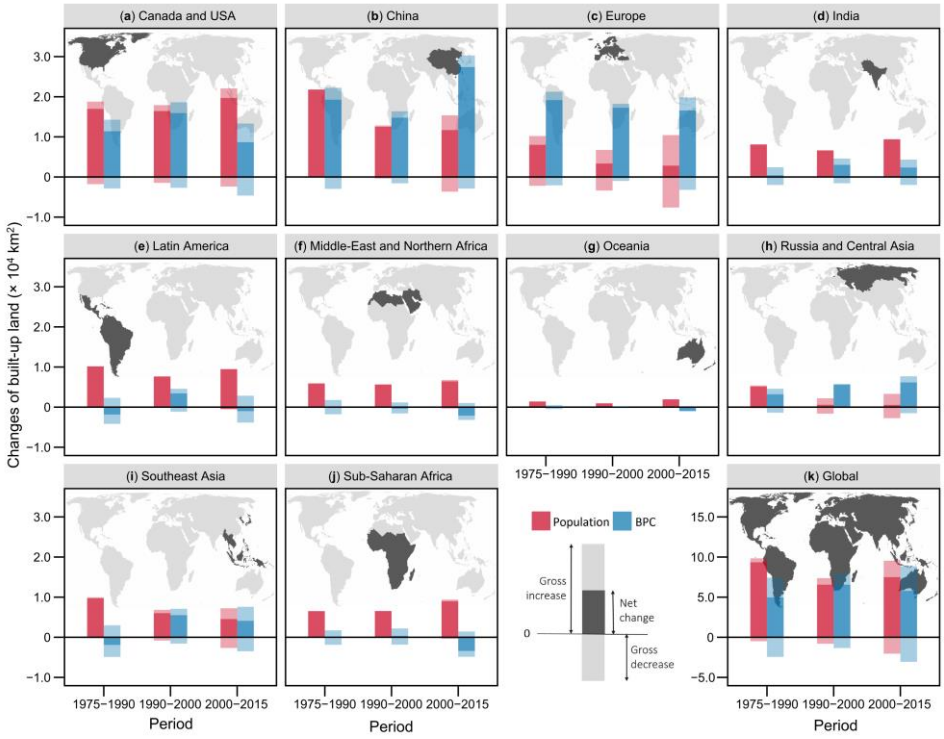


Figure D6: Heterogeneity in built-up land area changes. A_{POP} and A_{BPC} are calculated following Eq. (D3) and (D4).

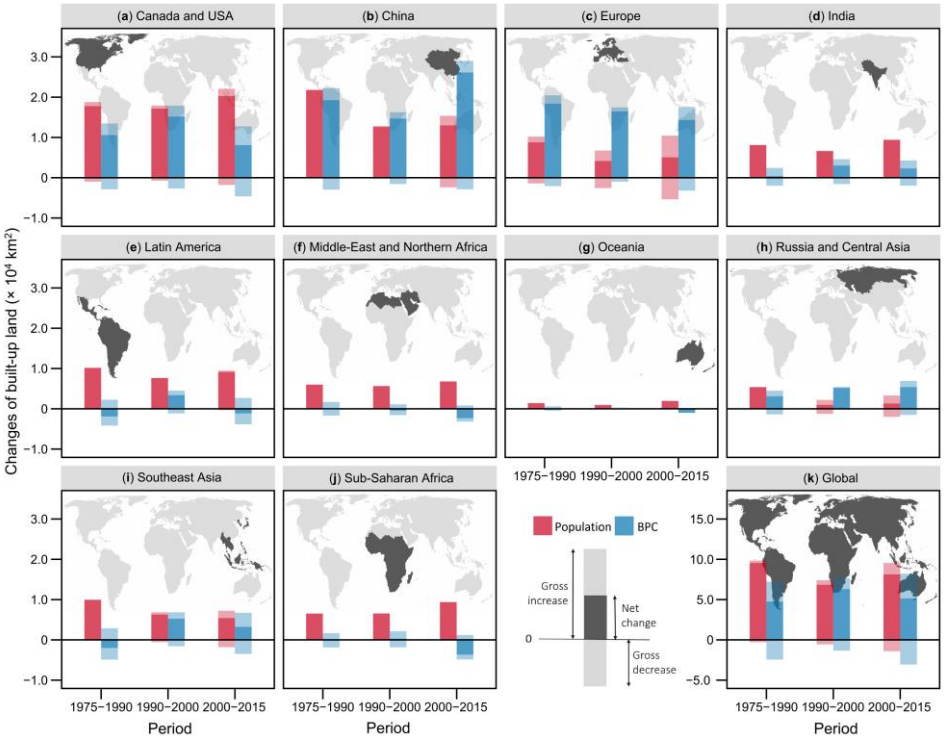


Figure D7: Heterogeneity in built-up land area changes. A_{POP} and A_{BPC} are calculated following Eq. (D5) and (D6).

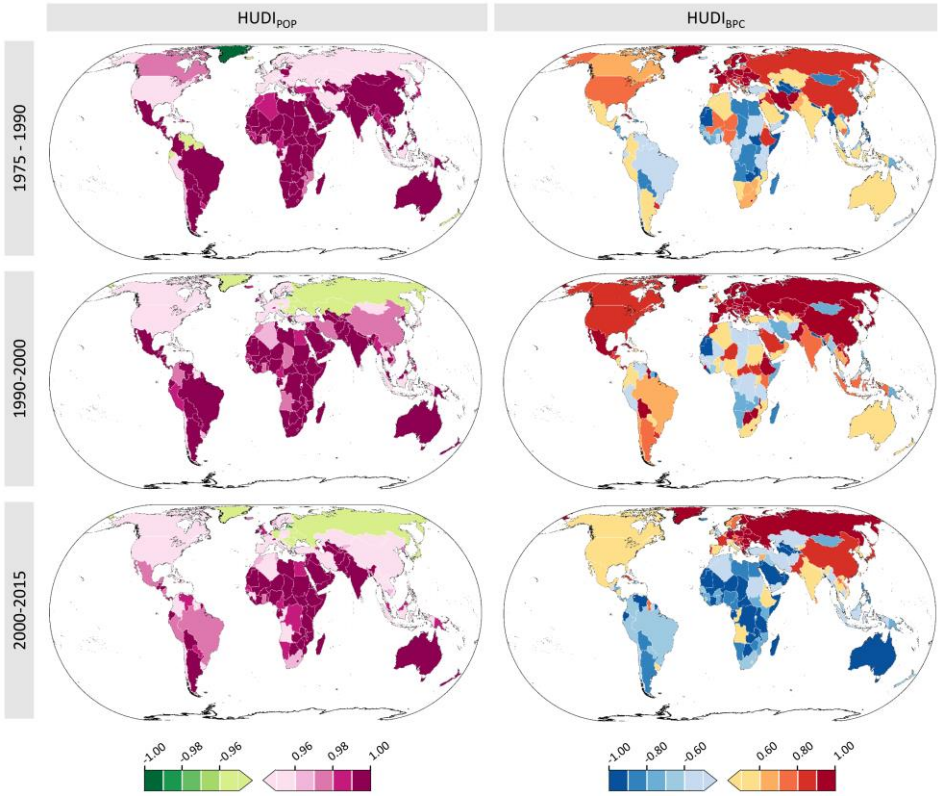


Figure D8: Homogeneity in built-up land change due to changes in population and due to changes in BPC within countries. Values closer to 1 or closer to -1 indicate that regions within a country all develop homogeneously, in terms of population dynamics or BPC changes, while values closer to 0 indicate more heterogeneous developments within a country.

Table D1: Analysis units for each country. For a full list of this table, please see the online version at https://doi.org/10.1016/j.landurbplan.2021.104308

ID	Country/Region code	Country/Region name	Most detailed unit available	Mean population per Level-2 unit	Analysis unit
1	ABW	Aruba	Level-0	-	-
2	AFG	Afghanistan	Level-2	-	Level-2
3	AGO	Angola	Level-3	146573	Level-2
4	AIA	Anguilla	Level-0	-	-
5	ALA	Åland	Level-1	-	Level-1
6	ALB	Albania	Level-3	78002	Level-2
7	AND	Andorra	Level-1	-	Level-1
8	ARE	United Arab Emirates	Level-3	42726	Level-2
9	ARG	Argentina	Level-2	-	Level-2
10	ARM	Armenia	Level-1	-	Level-1
11	ASM	American Samoa	Level-3	3281	Level-2
12	ATA	Antarctica	Level-0	-	-
13	ATF	French Southern Territories	Level-1	-	Level-1
14	ATG	Antigua and Barbuda	Level-1	-	Level-1
15	AUS	Australia	Level-2	-	Level-2
16	AUT	Austria	Level-3	78228	Level-2
17	AZE	Azerbaijan	Level-2	-	Level-2
18	BDI	Burundi	Level-4	79857	Level-2
19	BEL	Belgium	Level-4	1025858	Level-3
20	BEN	Benin	Level-2	-	Level-2
27	BIH	Bosnia and Herzegovina	Level-3	210094	Level-3
28	BLM	Saint-Barthélemy	Level-0	-	-
29	BLR	Belarus	Level-2	-	Level-2
30	BLZ	Belize	Level-1	-	Level-1
31	BMU	Bermuda	Level-1	-	Level-1
32	BOL	Bolivia	Level-3	89685	Level-2
33	BRA	Brazil	Level-3	29478	Level-2
…	…	…	…	…	…
256	ZWE	Zimbabwe	Level-2	-	Level-2

Table D2: Comparison of different approaches to allocating built-up land changes (in thousand km²) to changes in population and to changes in built-up land area per capita (BPC).

Period	Eq. (D1) and (D2)		Eq. (D3) and (D4)		Eq. (D5) and (D6)		Eq. (5.1) and (5.2)		Mean		S.D.	
	A_{POP}	A_{BPC}	A_{POP}	A_{BPC}	A_{POP}	A_{BPC}	A_{POP}	A_{BPC}	A_{POP}	A_{BPC}	A_{POP}	A_{BPC}
1975-1990	101.46 (70.9%)	41.70 (29.1%)	93.43 (65.3%)	49.74 (34.7%)	95.46 (66.7%)	47.71 (33.3%)	88.22 (61.7%)	54.77 (38.3%)	94.64 (66.1%)	48.48 (33.9%)	5.48 (3.8%)	5.41 (3.8%)
1990-2000	67.70 (51.5%)	63.87 (48.5%)	65.96 (50.1%)	65.61 (49.9%)	68.43 (52.0%)	63.14 (48.0%)	66.14 (50.4%)	65.21 (49.6%)	67.12 (51.0%)	64.46 (49.0%)	1.15 (0.9%)	1.16 (0.9%)
2000-2015	93.71 (70.5%)	39.13 (29.5%)	74.98 (46.4%)	57.86 (43.6%)	81.33 (61.2%)	51.51 (38.8%)	83.05 (62.5%)	49.58 (37.5%)	83.27 (62.7%)	49.58 (37.3%)	7.78 (5.9%)	7.78 (5.9%)

Table D3: Built-up land and population for major world regions as well as their changes over time. For each world region, 'Sum' refers to the total built-up land area or population in the starting year, or their total changes. Similarly, 'Mean' represents the average built-up land area or population per subdivision in the starting year, or their average changes per subdivision for corresponding world region. As indicated, mean values for nearly all world regions are significantly different from the mean of the entire regions combined (two-tailed t-test, $^{**}p<0.01$, $^{*}p<0.05$).

Period	World region	Built-up land in the starting year [km²]		New built-up land [km²]		Population in the starting year [million]		Population change [million]	
		Sum	Mean	Sum	Mean	Sum	Mean	Sum	Mean
1975-1990	Canada and USA	86965	25**	28386	8**	242	0.07**	38	0.01**
	China	47651	17**	41050	15**	939	0.34**	252	0.09**
	Europe	95541	5	27187	1**	580	0.03**	44	0.00**
	India	18411	5	8534	2*	801	0.20**	333	0.08**
	Latin America	27067	2**	8030	1*	330	0.03**	114	0.01**
	Middle-East and Northern Africa	12334	3**	5815	2*	168	0.05**	79	0.02**
	Oceania	6252	8**	1553	2	17	0.02**	4	0.00**
	Russia and Central Asia	27055	9**	8330	3**	184	0.06*	30	0.01**
	Southeast Asia	40095	3**	7792	1**	469	0.04**	148	0.01**
	Sub-Saharan Africa	13694	1**	6311	1**	349	0.03**	172	0.01**
1990-2000	Canada and USA	115351	33**	32380	9**	280	0.08*	35	0.01
	China	88701	32**	27299	10**	1191	0.43**	117	0.04**
	Europe	122728	6**	20631	1**	624	0.03**	16	0.00**
	India	26945	7	9662	2**	1134	0.29**	253	0.06**
	Latin America	35097	3**	10962	1**	444	0.03**	80	0.01**

	Middle-East and Northern Africa	18149	5**	5119	1**	247	0.07	71	0.02**
	Oceania	7805	10**	1034	1	21	0.03**	3	0.00**
	Russia and Central Asia	35386	12**	6158	2*	214	0.07	3	0.00**
	Southeast Asia	47887	4**	11534	1**	617	0.05**	95	0.01**
	Sub-Saharan Africa	20005	2**	6816	1**	521	0.04**	159	0.01**
2000-2015	Canada and USA	147731	42**	28386	8**	315	0.09	41	0.01**
	China	116000	42**	39120	14**	1308	0.47**	99	0.04**
	Europe	143359	7**	19431	1**	640	0.03**	29	0.00**
	India	36607	9	11672	3**	1386	0.35**	354	0.09**
	Latin America	46060	4**	7985	1**	524	0.04**	107	0.01**
	Middle-East and Northern Africa	23268	6**	4269	1**	318	0.09	103	0.03**
	Oceania	8840	11*	1066	1	24	0.03**	6	0.01**
	Russia and Central Asia	41544	14**	6721	2**	217	0.07*	10	0.00**
	Southeast Asia	59420	4**	8649	1**	712	0.05**	118	0.01**
	Sub-Saharan Africa	26821	2**	5559	0**	680	0.06**	319	0.03**

Table D4: BPC and relative changes in built-up land, population, and BPC for ten major world regions. For large parts of the world regions, BPC in the starting year, annual built-up land change, annual population change, and annual BPC change are significantly different from these for all world regions combined (two-tailed t-test, **p<0.01, *p<0.05).

Period	World region	BPC in the starting year [m²/person]	Δ Built-up land	Δ Population	Δ BPC
1975-1990	Canada and USA	359**	1.90%**	0.97%**	0.92%**
	China	51**	4.23%**	1.60%**	2.59%**
	Europe	165**	1.68%**	0.49%**	1.19%**
	India	23**	2.57%**	2.34%**	0.22%**
	Latin America	82**	1.75%**	1.99%	-0.24%**
	Middle-East and Northern Africa	73*	2.61%**	2.59%**	0.01%**
	Oceania	361**	1.49%**	1.29%**	0.19%
	Russia and Central Asia	147	1.81%**	1.00%**	0.80%*
	Southeast Asia	85**	1.19%**	1.84%**	-0.64%**
	Sub-Saharan Africa	39**	2.56%**	2.70%**	-0.14%
1990-2000	Canada and USA	412**	2.51%**	1.20%	1.29%**
	China	74**	2.72%	0.94%**	1.76%**
	Europe	197**	1.57%**	0.25%**	1.31%**
	India	24**	3.11%**	2.03%**	1.06%**
	Latin America	79**	2.76%**	1.66%	1.07%**
	Middle-East and Northern Africa	73**	2.52%**	2.55%**	-0.03%**

	Oceania	371**	1.25%**	1.16%	0.09%**
	Russia and Central Asia	166**	1.62%**	0.15%**	1.46%**
	Southeast Asia	78**	2.18%**	1.44%**	0.73%**
	Sub-Saharan Africa	38**	2.98%**	2.71%**	0.26%**
2000-2015	Canada and USA	469**	1.18%**	0.83%**	0.35%**
	China	89**	1.96%**	0.49%**	1.46%**
	Europe	224**	0.85%**	0.29%**	0.56%**
	India	26**	1.86%**	1.53%**	0.33%**
	Latin America	88**	1.07%	1.24%*	-0.17%
	Middle-East and Northern Africa	73**	1.13%*	1.89%**	-0.75%**
	Oceania	375**	0.76%	1.43%**	-0.66%**
	Russia and Central Asia	192**	1.00%**	0.30%**	0.70%**
	Southeast Asia	83**	0.91%	1.03%**	-0.11%**
	Sub-Saharan Africa	39**	1.26%**	2.60%**	-1.30%**

References

References

Abascal, A., Rodríguez-Carreño, I., Vanhuysse, S., Georganos, S., Sliuzas, R., Wolff, E., Kuffer, M., 2022, Identifying degrees of deprivation from space using deep learning and morphological spatial analysis of deprived urban areas, *Computers, Environment and Urban Systems* **95**. https://doi.org/10.1016/j.compenvurbsys.2022.101820

Abson, D. J., Fischer, J., Leventon, J., Newig, J., Schomerus, T., Vilsmaier, U., von Wehrden, H., Abernethy, P., Ives, C. D., Jager, N. W., Lang, D. J., 2017, Leverage points for sustainability transformation, *Ambio* **46**(1)**:**30-39. https://doi.org/10.1007/s13280-016-0800-y

Agergaard, J., Tacoli, C., Steel, G., Ørtenblad, S. B., 2019, Revisiting Rural–Urban Transformations and Small Town Development in Sub-Saharan Africa, *The European Journal of Development Research* **31**(1)**:**2-11. https://doi.org/10.1057/s41287-018-0182-z

Alberti, M., Booth, D., Hill, K., Coburn, B., Avolio, C., Coe, S., Spirandelli, D., 2007, The impact of urban patterns on aquatic ecosystems: An empirical analysis in Puget lowland sub-basins, *Landscape and Urban Planning* **80**(4)**:**345-361. https://doi.org/https://doi.org/10.1016/j.landurbplan.2006.08.001

Alexander Wandl, D. I., Nadin, V., Zonneveld, W., Rooij, R., 2014, Beyond urban–rural classifications: Characterising and mapping territories-in-between across Europe, *Landscape and Urban Planning* **130:**50-63. https://doi.org/10.1016/j.landurbplan.2014.06.010

Alig, R. J., Kline, J. D., Lichtenstein, M., 2004, Urbanization on the US landscape: looking ahead in the 21st century, *Landscape and Urban Planning* **69**(2-3)**:**219-234. https://doi.org/10.1016/j.landurbplan.2003.07.004

Alonso, W., 1964, Location and land use, Harvard University Press.

Anas, A., Arnott, R., Small, K. A., 1998, Urban spatial structure, *Journal of Economic Literature* **36**(3)**:**1426-1464.

Andrade-Núñez, M. J., Aide, T. M., 2018, Built-up expansion between 2001 and 2011 in South America continues well beyond the cities, *Environmental Research Letters* **13**(8). https://doi.org/10.1088/1748-9326/aad2e3

Angel, S., Arango Franco, S., Liu, Y., Blei, A. M., 2020, The shape compactness of urban footprints, *Progress in Planning* **139:**1-26. https://doi.org/10.1016/j.progress.2018.12.001

Angel, S., Lamson-Hall, P., Blanco, Z. G., 2021a, Anatomy of density: measurable factors that constitute urban density, *Buildings and Cities* **2**(1)**:**264-282. https://doi.org/10.5334/bc.91

Angel, S., Lamson-Hall, P., Blei, A., Shingade, S., Kumar, S., 2021b, Densify and Expand: A Global Analysis of Recent Urban Growth, *Sustainability* **13**(7). https://doi.org/10.3390/su13073835

Angel, S., Parent, J., Civco, D. L., Blei, A., 2010, The Persistent Decline in Urban Densities: Global and Historical Evidence of 'Sprawl', Lincoln Institute of Land Policy Cambridge.

Angel, S., Parent, J., Civco, D. L., Blei, A., Potere, D., 2011, The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050, *Progress in Planning* **75**(2)**:**53-107. https://doi.org/10.1016/j.progress.2011.04.001

Badmos, O., Rienow, A., Callo-Concha, D., Greve, K., Jürgens, C., 2018, Urban Development in West Africa—Monitoring and Intensity Analysis of Slum Growth in Lagos: Linking Pattern and Process, *Remote Sensing* **10**(7)**:**1044. https://doi.org/10.3390/rs10071044

Bagan, H., Yamagata, Y., 2012, Landsat analysis of urban growth: How Tokyo became the world's largest megacity during the last 40 years, *Remote sensing of Environment* **127:**210-222. https://doi.org/10.1016/j.rse.2012.09.011

Bagheri, H., Schmitt, M., d'Angelo, P., Zhu, X. X., 2018, A framework for SAR-optical stereogrammetry over urban areas, *ISPRS Journal of Photogrammetry and Remote Sensing* **146:**389-408. https://doi.org/10.1016/j.isprsjprs.2018.10.003

Baiocchi, G., Creutzig, F., Minx, J., Pichler, P.-P., 2015, A spatial typology of human settlements and their CO2 emissions in England, *Global Environmental Change* **34:**13-21. https://doi.org/10.1016/j.gloenvcha.2015.06.001

Bakker, V., Verburg, P. H., van Vliet, J., 2021, Trade-offs between prosperity and urban land per capita in major world cities, *Geography and Sustainability* **2**(2)**:**134-138. https://doi.org/10.1016/j.geosus.2021.05.004

Balk, D., Leyk, S., Jones, B., Montgomery, M. R., Clark, A., 2018, Understanding urbanization: A study of census and satellite-derived urban classes in the United States, 1990-2010, *PLoS One* **13**(12)**:**e0208487. https://doi.org/10.1371/journal.pone.0208487

Barrington-Leigh, C., Millard-Ball, A., 2015, A century of sprawl in the United States, *Proceedings of the National Academy of Sciences* **112**(27)**:**8244-9. https://doi.org/10.1073/pnas.1504033112

Bartholomé, E., Belward, A. S., 2007, GLC2000: a new approach to global land cover mapping from Earth observation data, *International Journal of Remote Sensing* **26**(9)**:**1959-1977. https://doi.org/10.1080/01431160412331291297

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., Portugali, Y., 2012, Smart cities of the future, *The European Physical Journal Special Topics* **214**(1)**:**481-518. https://doi.org/10.1140/epjst/e2012-01703-3

Berger, C., Rosentreter, J., Voltersen, M., Baumgart, C., Schmullius, C., Hese, S., 2017, Spatio-temporal analysis of the relationship between 2D/3D urban site characteristics and land surface temperature, *Remote Sensing of Environment* **193:**225-243. https://doi.org/10.1016/j.rse.2017.02.020

Bertaud, A., 2001, Metropolis: A measure of the spatial organization of 7 large cities, *Unpublished working paper***:**1-22.

Biljecki, F., 2020, Exploration of open data in southeast Asia to generate 3d building models, *ISPRS Annals of Photogrammetry, Remote Sensing Spatial Information Sciences* **6:**37-44. https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-37-2020

Biljecki, F., Chow, Y. S., 2022, Global Building Morphology Indicators, *Computers, Environment and Urban Systems* **95**. https://doi.org/10.1016/j.compenvurbsys.2022.101809

Biljecki, F., Ledoux, H., Stoter, J., 2016, An improved LOD specification for 3D building models, *Computers, Environment and Urban Systems* **59:**25-37. https://doi.org/10.1016/j.compenvurbsys.2016.04.005

Blei, A. M., Angel, S., Civco, D. L., Liu, Y., Zhang, X., 2018, Accuracy assessment and map comparisons for monitoring urban expansion: the Atlas of urban expansion and the global human settlement layer, *Lincoln Institute of Land Policy: Cambridge, MA, USA*.

Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Mayaux, P., Boettcher, M., Brockmann, C., Kirches, G., 2013, Consistent global land cover maps for climate modelling communities: current achievements of the ESA's land cover CCI, in: *Proceedings of the ESA living planet symposium, Edimburgh*, pp. 9-13.

Boudet, F., MacDonald, G. K., Robinson, B. E., Samberg, L. H., 2020, Rural-urban connectivity and agricultural land management across the Global South, *Global Environmental Change* **60:**101982. https://doi.org/10.1016/j.gloenvcha.2019.101982

Breiman, L., 2001, Random forests, *Machine learning* **45**(1)**:**5-32.

Bren d'Amour, C., Reitsma, F., Baiocchi, G., Barthel, S., Guneralp, B., Erb, K. H., Haberl, H., Creutzig, F., Seto, K. C., 2017, Future urban land expansion and implications for global croplands, *Proceedings of the National Academy of Sciences* **114**(34)**:**8939-8944. https://doi.org/10.1073/pnas.1606036114

Broitman, D., Koomen, E., 2015, Residential density change: Densification and urban expansion, *Computers, Environment and Urban Systems* **54:**32-46. https://doi.org/10.1016/j.compenvurbsys.2015.05.006

Broitman, D., Koomen, E., 2020, The attraction of urban cores: Densification in Dutch city centres, *Urban Studies* **57**(9)**:**1920-1939. https://doi.org/10.1177/0042098019864019

Buchadas, A., Qin, S., Meyfroidt, P., Kuemmerle, T., 2022, Conservation frontiers: understanding the geographic expansion of conservation, *Journal of Land Use Science***:**1-14. https://doi.org/10.1080/1747423x.2021.2018516

Bürgi, M., Celio, E., Diogo, V., Hersperger, A. M., Kizos, T., Lieskovsky, J., Pazur, R., Plieninger, T., Prishchepov, A. V., Verburg, P. H., 2022, Advancing the study of driving forces of landscape change, *Journal of Land Use Science***:**1-16. https://doi.org/10.1080/1747423x.2022.2029599

Cao, S., Cai, Y., Du, M., Weng, Q., Lu, L., 2022, Seasonal and diurnal surface urban heat islands in China: an investigation of driving factors with three-dimensional urban morphological parameters, *GIScience & Remote Sensing* **59**(1)**:**1121-1142. https://doi.org/10.1080/15481603.2022.2100100

Cao, Y., Huang, X., 2021, A deep learning method for building height estimation using high-resolution multi-view imagery over urban areas: A case study of 42 Chinese cities, *Remote Sensing of Environment* **264**. https://doi.org/10.1016/j.rse.2021.112590

Cao, Y., Zou, J., Fang, X., Wang, J., Cao, Y., Li, G., 2020, Effect of land tenure fragmentation on the decision-making and scale of agricultural land transfer in China, *Land Use Policy* **99**. https://doi.org/10.1016/j.landusepol.2020.104996

Cao, Z., Zheng, X., Liu, Y., Li, Y., Chen, Y., 2018, Exploring the changing patterns of China's migration and its determinants using census data of 2000 and 2010, *Habitat International* **82:**72-82. https://doi.org/10.1016/j.habitatint.2018.09.006

Carlson, T., Sanchez-Azofeifa, G. A., 1999, Satellite remote sensing of land use changes in and around San Jose, Costa Rica, *Remote Sensing of Environment* **70**(3)**:**247-256. https://doi.org/10.1016/S0034-4257(99)00018-8

CGTN, 2017, Beijing city gov't begins move to Tongzhou.

Chai, B., Seto, K. C., 2019, Conceptualizing and characterizing micro-urbanization: A new perspective applied to Africa, *Landscape and Urban Planning* **190:**103595. https://doi.org/10.1016/j.landurbplan.2019.103595

Chapman, S., Watson, J. E. M., Salazar, A., Thatcher, M., McAlpine, C. A., 2017, The impact of urbanization and climate change on urban temperatures: a systematic review, *Landscape Ecology* **32**(10)**:**1921-1935. https://doi.org/10.1007/s10980-017-0561-4

Chen, T.-H. K., Qiu, C., Schmitt, M., Zhu, X. X., Sabel, C. E., Prishchepov, A. V., 2020a, Mapping horizontal and vertical urban densification in Denmark with Landsat time-series

from 1985 to 2018: A semantic segmentation solution, *Remote Sensing of Environment* **251:**112096. https://doi.org/10.1016/j.rse.2020.112096

Chen, Y., 2022, An extended patch-based cellular automaton to simulate horizontal and vertical urban growth under the shared socioeconomic pathways, *Computers, Environment and Urban Systems* **91**. https://doi.org/10.1016/j.compenvurbsys.2021.101727

Chen, Y., Yue, W., La Rosa, D., 2020b, Which communities have better accessibility to green space? An investigation into environmental inequality using big data, *Landscape and Urban Planning* **204**. https://doi.org/10.1016/j.landurbplan.2020.103919

Chi, G., Marcouiller, D. W., 2013, In-migration to remote rural regions: The relative impacts of natural amenities and land developability, *Landscape and Urban Planning* **117:**22-31. https://doi.org/10.1016/j.landurbplan.2013.04.012

Christaller, W., 1933, Die zentralen Orte in Suddeutschland: Eine okonomisch-geographische Untersuchung uber die Gesetzmassigkeit der Verbreitung und Entwicklung der Siedlungen mit stadtischen Funktionen, *Jena*.

Colsaet, A., Laurans, Y., Levrel, H., 2018, What drives land take and urban land expansion? A systematic review, *Land Use Policy* **79:**339-349. https://doi.org/10.1016/j.landusepol.2018.08.017

Connors, J. P., Galletti, C. S., Chow, W. T. L., 2012, Landscape configuration and urban heat island effects: assessing the relationship between landscape characteristics and land surface temperature in Phoenix, Arizona, *Landscape Ecology* **28**(2)**:**271-283. https://doi.org/10.1007/s10980-012-9833-1

Cortinovis, C., Haase, D., Zanon, B., Geneletti, D., 2019, Is urban spatial development on the right track? Comparing strategies and trends in the European Union, *Landscape and Urban Planning* **181:**22-37. https://doi.org/10.1016/j.landurbplan.2018.09.007

Creutzig, F., Baiocchi, G., Bierkandt, R., Pichler, P. P., Seto, K. C., 2015, Global typology of urban energy use and potentials for an urbanization mitigation wedge, *Proceedings of the National Academy of Sciences* **112**(20)**:**6283-8. https://doi.org/10.1073/pnas.1315545112

Cullum, C., Brierley, G., Perry, G. L. W., Witkowski, E. T. F., 2016, Landscape archetypes for ecological classification and mapping, *Progress in Physical Geography: Earth and Environment* **41**(1)**:**95-123. https://doi.org/10.1177/0309133316671103

Cumming, G. S., Buerkert, A., Hoffmann, E. M., Schlecht, E., von Cramon-Taubadel, S., Tscharntke, T., 2014, Implications of agricultural transitions and urbanization for ecosystem services, *Nature* **515**(7525)**:**50-7. https://doi.org/10.1038/nature13945

Dadashpoor, H., Ahani, S., 2019, A conceptual typology of the spatial territories of the peripheral areas of metropolises, *Habitat International* **90**. https://doi.org/10.1016/j.habitatint.2019.102015

Dai, X., Guo, Z., Zhang, L., Wu, J., 2010, Spatio-temporal pattern of urban land cover evolvement with urban renewal and expansion in Shanghai based on mixed-pixel classification for remote sensing imagery, *International Journal of Remote Sensing* **31**(23)**:**6095-6114. https://doi.org/10.1080/01431160903376407

de Lange, M., 2022, Human settlement systems: Mapping settlement systems and their population distributions in Europe and Eastern Africa, Vrije Universiteit Amsterdam.

Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D., Raska, R., 2018, Deepglobe 2018: A challenge to parse the earth through satellite images, in: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, pp. 172-17209.

Demuzere, M., Bechtel, B., Middel, A., Mills, G., 2019, Mapping Europe into local climate zones, *PLoS One* **14**(4)**:**e0214474. https://doi.org/10.1371/journal.pone.0214474

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., Bechtel, B., 2022, A global map of local climate zones to support earth system modelling and urban-scale environmental science, *Earth System Science Data* **2022**(14)**:**3835–3873. https://doi.org/10.5194/essd-14-3835-2022

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Li, F.-f., 2009, Imagenet: A large-scale hierarchical image database, in: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 248-255.

Dong, P., Ramesh, S., Nepali, A., 2010, Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data, *International Journal of Remote Sensing* **31**(21)**:**5571-5586. https://doi.org/10.1080/01431161.2010.496804

Dong, T., Jiao, L., Xu, G., Yang, L., Liu, J., 2019, Towards sustainability? Analyzing changing urban form patterns in the United States, Europe, and China, *Science of the Total Environment* **671:**632-643. https://doi.org/10.1016/j.scitotenv.2019.03.269

Dou, Y., Cosentino, F., Malek, Z., Maiorano, L., Thuiller, W., Verburg, P. H., 2021, A new European land systems representation accounting for landscape characteristics, *Landscape Ecology* **36**(8)**:**2215-2234. https://doi.org/10.1007/s10980-021-01227-5

Dovey, K., Pafka, E., 2013, The urban density assemblage: Modelling multiple measures, *URBAN DESIGN International* **19**(1)**:**66-76. https://doi.org/10.1057/udi.2013.13

Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., Baptista, S. R., 2015, Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4, *Papers in Applied Geography* **1**(3)**:**226-234. https://doi.org/10.1080/23754931.2015.1014272

Du, S., He, C., Huang, Q., Shi, P., 2018, How did the urban land in floodplains distribute and expand in China from 1992–2015?, *Environmental Research Letters* **13**(3)**:**034018. https://doi.org/10.1088/1748-9326/aaac07

Easterlin, R. A., Angelescu, L., Zweig, J. S., 2011, The Impact of Modern Economic Growth on Urban–Rural Differences in Subjective Well-Being, *World Development* **39**(12)**:**2187-2198. https://doi.org/10.1016/j.worlddev.2011.04.015

Ehrlich, D., Kemper, T., Pesaresi, M., Corbane, C., 2018, Built-up area and population density: Two Essential Societal Variables to address climate hazard impact, *Environmental Science & Policy* **90:**73-82. https://doi.org/10.1016/j.envsci.2018.10.001

Ehrlich, D., Melchiorri, M., Capitani, C., 2021, Population Trends and Urbanisation in Mountain Ranges of the World, *Land* **10**(3). https://doi.org/10.3390/land10030255

Eitelberg, D. A., van Vliet, J., Doelman, J. C., Stehfest, E., Verburg, P. H., 2016, Demand for biodiversity protection and carbon storage as drivers of global land change scenarios, *Global Environmental Change* **40:**101-111. https://doi.org/10.1016/j.gloenvcha.2016.06.014

Elhacham, E., Ben-Uri, L., Grozovski, J., Bar-On, Y. M., Milo, R., 2020, Global human-made mass exceeds all living biomass, *Nature* **588**(7838)**:**442-444. https://doi.org/10.1038/s41586-020-3010-5

Ellis, E. C., Gauthier, N., Klein Goldewijk, K., Bliege Bird, R., Boivin, N., Diaz, S., Fuller, D. Q., Gill, J. L., Kaplan, J. O., Kingston, N., Locke, H., McMichael, C. N. H., Ranco, D., Rick, T. C., Shaw, M. R., Stephens, L., Svenning, J. C., Watson, J. E. M., 2021, People have shaped most of terrestrial nature for at least 12,000 years, *Proceedings of the National Academy of Sciences* **118**(17). https://doi.org/10.1073/pnas.2023483118

Ellis, E. C., Kaplan, J. O., Fuller, D. Q., Vavrus, S., Klein Goldewijk, K., Verburg, P. H., 2013, Used planet: a global history, *Proceedings of the National Academy of Sciences* **110**(20)**:**7978-85. https://doi.org/10.1073/pnas.1217241110

Ellis, E. C., Ramankutty, N., 2008, Putting people in the map: anthropogenic biomes of the world, *Frontiers in Ecology and the Environment* **6**(8)**:**439-447. https://doi.org/10.1890/070062

Elvidge, C. D., Tuttle, B. T., Sutton, P. C., Baugh, K. E., Howard, A. T., Milesi, C., Bhaduri, B., Nemani, R., 2007, Global Distribution and Density of Constructed Impervious Surfaces, *Sensors* **7**(9)**:**1962-1979.

Engelfriet, L., Koomen, E., 2017, The impact of urban form on commuting in large Chinese cities, *Transportation* **45**(5)**:**1269-1295. https://doi.org/10.1007/s11116-017-9762-6

Erb, K. H., Haberl, H., Jepsen, M. R., Kuemmerle, T., Lindner, M., Muller, D., Verburg, P. H., Reenberg, A., 2013, A conceptual framework for analysing and measuring land-use intensity, *Current Opinion in Environmental Sustainability* **5**(5)**:**464-470. https://doi.org/10.1016/j.cosust.2013.07.010

ESA, 2019, The Sentinel-1 Toolbox, in: *https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-1*.

Esch, T., Brzoska, E., Dech, S., Leutner, B., Palacios-Lopez, D., Metz-Marconcini, A., Marconcini, M., Roth, A., Zeidler, J., 2022, World Settlement Footprint 3D - A first three-dimensional survey of the global building stock, *Remote Sensing of Environment* **270**. https://doi.org/10.1016/j.rse.2021.112877

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., Strano, E., 2017, Breaking new ground in mapping human settlements from space – The Global Urban Footprint, *ISPRS Journal of Photogrammetry and Remote Sensing* **134:**30-42. https://doi.org/10.1016/j.isprsjprs.2017.10.012

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenbock, H., Muller, A., Dech, S., 2013, Urban Footprint Processor—Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission, *IEEE Geoscience and Remote Sensing Letters* **10**(6)**:**1617-1621. https://doi.org/10.1109/lgrs.2013.2272953

Espindola, G. M. d., Carneiro, E. L. N. d. C., Façanha, A. C., 2017, Four decades of urban sprawl and population growth in Teresina, Brazil, *Applied Geography* **79:**73-83. https://doi.org/10.1016/j.apgeog.2016.12.018

Ewing, R., Cervero, R., 2010, Travel and the built environment: a meta-analysis, *Journal of the American Planning Association* **76**(3)**:**265-294. https://doi.org/10.1080/01944361003766766

Fan, P., Yue, W., Zhang, J., Huang, H., Messina, J., Verburg, P. H., Qi, J., Moore, N., Ge, J., 2020, The spatial restructuring and determinants of industrial landscape in a mega city under rapid urbanization, *Habitat International* **95**. https://doi.org/10.1016/j.habitatint.2019.102099

Fang, C., Li, G., Wang, S., 2016, Changing and Differentiated Urban Landscape in China: Spatiotemporal Patterns and Driving Forces, *Environ Sci Technol* **50**(5)**:**2217-27. https://doi.org/10.1021/acs.est.5b05198

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F., Zanchetta, L., 2019a, GHSL Data Package 2019, Luxembourg.

Florczyk, A. J., Corbane, C., Schiavina, M., Pesaresi, M., Maffenini, L., Melchiorri, M., Zanchetta, L., 2019b, GHS Urban Centre Database 2015, multitemporal and

multidimensional attributes, R2019A, *Luxembourg: European Commission, Joint Research Centre (JRC)*.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., 2005, Global consequences of land use, *Science* **309**(5734)**:**570-574. https://doi.org/10.1126/science.1111772

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockstrom, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D. P., 2011, Solutions for a cultivated planet, *Nature* **478**(7369)**:**337-42. https://doi.org/10.1038/nature10452

Forman, R. T., Wu, J., 2016, Where to put the next billion people, *Nature* **537**(7622)**:**608. https://doi.org/10.1038/537608a

Frantz, D., Schug, F., Okujeni, A., Navacchi, C., Wagner, W., van der Linden, S., Hostert, P., 2021, National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series, *Remote Sensing of Environment* **252**. https://doi.org/10.1016/j.rse.2020.112128

Freire, S., Doxsey-Whitfield, E., MacManus, K., Mills, J., Pesaresi, M., 2016, Development of new open and free multi-temporal global population grids at 250 m resolution, *AGILE 2016 Helsinki*.

Frolking, S., Milliman, T., Seto, K. C., Friedl, M. A., 2013, A global fingerprint of macro-scale changes in urban structure from 1999 to 2009, *Environmental Research Letters* **8**(2)**:**024004. https://doi.org/10.1088/1748-9326/8/2/024004

Fujita, M., Krugman, P., 1995, When is the economy monocentric?: von Thünen and Chamberlin unified, *Regional Science and Urban Economics* **25**(4)**:**505-528. https://doi.org/10.1016/0166-0462(95)02098-F

Gago, E. J., Roldan, J., Pacheco-Torres, R., Ordóñez, J., 2013, The city and urban heat islands: A review of strategies to mitigate adverse effects, *Renewable and Sustainable Energy Reviews* **25:**749-758. https://doi.org/10.1016/j.rser.2013.05.057

Gao, J., O'Neill, B. C., 2020, Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways, *Nature Communications* **11**(1)**:**2302. https://doi.org/10.1038/s41467-020-15788-7

Gardi, C., Panagos, P., Van Liedekerke, M., Bosco, C., De Brogniez, D., 2015, Land take and food security: assessment of land take on the agricultural production in Europe, *Journal of Environmental Planning and Management* **58**(5)**:**898-912. https://doi.org/10.1080/09640568.2014.899490

Geiß, C., Leichtle, T., Wurm, M., Pelizari, P. A., Standfuß, I., Zhu, X. X., So, E., Siedentop, S., Esch, T., Taubenböck, H., 2019, Large-Area Characterization of Urban Morphology—Mapping of Built-Up Height and Density Using TanDEM-X and Sentinel-2 Data, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12**(8)**:**2912-2927. https://doi.org/10.1109/JSTARS.2019.2917755

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2013, Human Capital and Regional Development, *The Quarterly Journal of Economics* **128**(1)**:**105-164. https://doi.org/10.1093/qje/qjs050

Georg, I., Blaschke, T., Taubenböck, H., 2018, Are we in Boswash yet? A multi-source geodata approach to spatially delimit urban corridors, *ISPRS International Journal of Geo-Information* **7**(1)**:**15. https://doi.org/10.3390/ijgi7010015

Gómez, J. A., Patiño, J. E., Duque, J. C., Passos, S., 2019, Spatiotemporal Modeling of Urban Growth Using Machine Learning, *Remote Sensing* **12**(1)**:**109. https://doi.org/10.3390/rs12010109

Gong, J., Hu, Z., Chen, W., Liu, Y., Wang, J., 2018, Urban expansion dynamics and modes in metropolitan Guangzhou, China, *Land Use Policy* **72:**100-109. https://doi.org/10.1016/j.landusepol.2017.12.025

Gong, P., Chen, B., Li, X., Liu, H., Wang, J., Bai, Y., Chen, J., Chen, X., Fang, L., Feng, S., Feng, Y., Gong, Y., Gu, H., Huang, H., Huang, X., Jiao, H., Kang, Y., Lei, G., Li, A., Li, X., Li, X., Li, Y., Li, Z., Li, Z., Liu, C., Liu, C., Liu, M., Liu, S., Mao, W., Miao, C., Ni, H., Pan, Q., Qi, S., Ren, Z., Shan, Z., Shen, S., Shi, M., Song, Y., Su, M., Ping Suen, H., Sun, B., Sun, F., Sun, J., Sun, L., Sun, W., Tian, T., Tong, X., Tseng, Y., Tu, Y., Wang, H., Wang, L., Wang, X., Wang, Z., Wu, T., Xie, Y., Yang, J., Yang, J., Yuan, M., Yue, W., Zeng, H., Zhang, K., Zhang, N., Zhang, T., Zhang, Y., Zhao, F., Zheng, Y., Zhou, Q., Clinton, N., Zhu, Z., Xu, B., 2020a, Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018, *Science Bulletin* **65**(3)**:**182-187. https://doi.org/10.1016/j.scib.2019.12.007

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., Zhou, Y., 2020b, Annual maps of global artificial impervious area (GAIA) between 1985 and 2018, *Remote Sensing of Environment* **236:**111510. https://doi.org/10.1016/j.rse.2019.111510

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., 2013, Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data, *International Journal of Remote Sensing* **34**(7)**:**2607-2654. https://doi.org/10.1080/01431161.2012.748992

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017, Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sensing of Environment* **202:**18-27. https://doi.org/10.1016/j.rse.2017.06.031

Grace Wong, K. M., 2004, Vertical cities as a solution for land scarcity: the tallest public housing development in Singapore, *URBAN DESIGN International* **9**(1)**:**17-30. https://doi.org/10.1057/palgrave.udi.9000108

Grêt-Regamey, A., Galleguillos-Torres, M., Dissegna, A., Weibel, B., 2020, How urban densification influences ecosystem services—a comparison between a temperate and a tropical city, *Environmental Research Letters* **15**(7)**:**075001. https://doi.org/10.1088/1748-9326/ab7acf

Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., Seto, K. C., 2017a, Urbanization in Africa: challenges and opportunities for conservation, *Environmental Research Letters* **13**(1)**:**015002. https://doi.org/10.1088/1748-9326/aa94fe

Güneralp, B., Reba, M., Hales, B. U., Wentz, E. A., Seto, K. C., 2020, Trends in urban land expansion, density, and land transitions from 1970 to 2010: a global synthesis, *Environmental Research Letters* **15**(4)**:**044015. https://doi.org/10.1088/1748-9326/ab6669

Güneralp, B., Zhou, Y., Ürge-Vorsatz, D., Gupta, M., Yu, S., Patel, P. L., Fragkias, M., Li, X., Seto, K. C., 2017b, Global scenarios of urban density and its impacts on building energy use through 2050, *Proceedings of the National Academy of Sciences* **114**(34)**:**8945-8950. https://doi.org/10.1073/pnas.1606035114

Guo, A., Yue, W., Yang, J., He, T., Zhang, M., Li, M., 2022, Divergent impact of urban 2D/3D morphology on thermal environment along urban gradients, *Urban Climate* **45**. https://doi.org/10.1016/j.uclim.2022.101278

Haberl, H., Wiedenhofer, D., Schug, F., Frantz, D., Virag, D., Plutzar, C., Gruhler, K., Lederer, J., Schiller, G., Fishman, T., Lanau, M., Gattringer, A., Kemper, T., Liu, G., Tanikawa, H., van der Linden, S., Hostert, P., 2021, High-Resolution Maps of Material Stocks in Buildings and Infrastructures in Austria and Germany, *Environmental Science and Technology* **55**(5)**:**3368-3379. https://doi.org/10.1021/acs.est.0c05642

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., 2013, High-resolution global maps of 21st-century forest cover change, *Science* **342**(6160)**:**850-853. https://doi.org/10.1126/science.1244693

Harun, S., Ogneva-Himmelberger, Y., 2013, Distribution of industrial farms in the United States and socioeconomic, health, and environmental characteristics of counties, *Geography Journal* **2013**. https://doi.org/10.1155/2013/385893

He, C., Liu, Z., Tian, J., Ma, Q., 2014, Urban expansion dynamics and natural habitat loss in China: a multiscale landscape perspective, *Global Change Biology* **20**(9)**:**2886-902. https://doi.org/10.1111/gcb.12553

He, M., Zhu, Q., Du, Z., Hu, H., Ding, Y., Chen, M., 2016, A 3D Shape Descriptor Based on Contour Clusters for Damaged Roof Detection Using Airborne LiDAR Point Clouds, *Remote Sensing* **8**(3). https://doi.org/10.3390/rs8030189

Hemerijckx, L.-M., Janusz, K., Van Emelen, S., Tumwesigye, S., Davis, J., Lwasa, S., Van Rompaey, A., 2022, Food accessibility of different socioeconomic groups in sub-Saharan African cities: a mixed-method analysis in Kampala, Uganda, *Food Security* **14**(3)**:**677-694. https://doi.org/10.1007/s12571-021-01248-7

Henderson, J. V., Quigley, J., Lim, E., 2009, Urbanization in China: Policy issues and options, *Unpublished manuscript, Brown University*.

Hersperger, A. M., Oliveira, E., Pagliarin, S., Palka, G., Verburg, P., Bolliger, J., Grădinaru, S., 2018, Urban land-use change: The role of strategic spatial planning, *Global Environmental Change* **51:**32-42. https://doi.org/10.1016/j.gloenvcha.2018.05.001

Heuvelink, G. B., Burrough, P. A., Stein, A., 1989, Propagation of errors in spatial modelling with GIS, *International Journal of Geographical Information System* **3**(4)**:**303-322. https://doi.org/10.1080/02693798908941518

Howley, P., 2009, Attitudes towards compact city living: Towards a greater understanding of residential behaviour, *Land use policy* **26**(3)**:**792-798. https://doi.org/10.1016/j.landusepol.2008.10.004

Hu, J., Wang, Y., Taubenböck, H., Zhu, X. X., 2021, Land consumption in cities: A comparative study across the globe, *Cities* **113:**103163. https://doi.org/10.1016/j.cities.2021.103163

Huang, H., Chen, P., Xu, X., Liu, C., Wang, J., Liu, C., Clinton, N., Gong, P., 2022, Estimating building height in China from ALOS AW3D30, *ISPRS Journal of Photogrammetry and Remote Sensing* **185:**146-157. https://doi.org/10.1016/j.isprsjprs.2022.01.022

Huang, J., Lu, X. X., Sellers, J. M., 2007, A global comparative analysis of urban form: Applying spatial metrics and remote sensing, *Landscape and Urban Planning* **82**(4)**:**184-197. https://doi.org/10.1016/j.landurbplan.2007.02.010

Huang, X., Schneider, A., Friedl, M. A., 2016, Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights, *Remote Sensing of Environment* **175:**92-108. https://doi.org/10.1016/j.rse.2015.12.042

Hudeček, T., Hnilička, P., Dlouhý, M., Leňo Cutáková, L., Leňo, M., 2019, Urban structures, population density and municipal expenditures: An empirical study in the Czech Republic, *Urban Studies* **56**(16)**:**3450-3465. https://doi.org/10.1177/0042098018813268

Ianoş, I., Jones, R., 2019, Local aspects of change in the rural-urban fringe of a metropolitan area: A study of Bucharest, Romania, *Habitat International* **91**. https://doi.org/10.1016/j.habitatint.2019.102026

Kamusoko, C., Gamba, J., 2015, Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model, *ISPRS International Journal of Geo-Information* **4**(2)**:**447-470. https://doi.org/10.3390/ijgi4020447

Kedron, P., Zhao, Y., Frazier, A. E., 2019, Three dimensional (3D) spatial metrics for objects, *Landscape Ecology* **34**(9)**:**2123-2132. https://doi.org/10.1007/s10980-019-00861-4

Kennedy, C. A., Stewart, I., Facchini, A., Cersosimo, I., Mele, R., Chen, B., Uda, M., Kansal, A., Chiu, A., Kim, K. G., Dubeux, C., Lebre La Rovere, E., Cunha, B., Pincetl, S., Keirstead, J., Barles, S., Pusaka, S., Gunawan, J., Adegbile, M., Nazariha, M., Hoque, S., Marcotullio, P. J., Gonzalez Otharan, F., Genena, T., Ibrahim, N., Farooqui, R., Cervantes, G., Sahin, A. D., 2015, Energy and material flows of megacities, *Proceedings of the National Academy of Sciences* **112**(19)**:**5985-90. https://doi.org/10.1073/pnas.1504315112

Khan, S., Carville, A., 2017, To follow the Australian dream or to embrace urban densification, *Growing Compact: Urban Form, Density and Sustainability***:**301.

Kim, G., Miller, P. A., Nowak, D. J., 2018, Urban vacant land typology: A tool for managing urban vacant land, *Sustainable Cities and Society* **36:**144-156. https://doi.org/10.1016/j.scs.2017.09.014

Klotz, M., Kemper, T., Geiß, C., Esch, T., Taubenböck, H., 2016, How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe, *Remote Sensing of Environment* **178:**191-212. https://doi.org/10.1016/j.rse.2016.03.001

Koks, E. E., Rozenberg, J., Zorn, C., Tariverdi, M., Vousdoukas, M., Fraser, S. A., Hall, J. W., Hallegatte, S., 2019, A global multi-hazard risk analysis of road and railway infrastructure assets, *Nature Communications* **10**(1)**:**2677. https://doi.org/10.1038/s41467-019-10442-3

Kourtit, K., Nijkamp, P., Reid, N., 2014, The new urban world: Challenges and policy, *Applied Geography* **49:**1-3. https://doi.org/10.1016/j.apgeog.2014.01.007

Krehl, A., Siedentop, S., 2018, Towards a typology of urban centers and subcenters – evidence from German city regions, *Urban Geography* **40**(1)**:**58-82. https://doi.org/10.1080/02723638.2018.1500245

Kuang, W., Chi, W., Lu, D., Dou, Y., 2014, A comparative analysis of megacity expansions in China and the U.S.: Patterns, rates and driving forces, *Landscape and Urban Planning* **132:**121-135. https://doi.org/10.1016/j.landurbplan.2014.08.015

Kuang, W., Liu, J., Dong, J., Chi, W., Zhang, C., 2016, The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers, *Landscape and Urban Planning* **145:**21-33. https://doi.org/10.1016/j.landurbplan.2015.10.001

Kuang, W., Yang, T., Yan, F., 2017, Examining urban land-cover characteristics and ecological regulation during the construction of Xiong'an New District, Hebei Province, China, *Journal of Geographical Sciences* **28**(1)**:**109-123. https://doi.org/10.1007/s11442-018-1462-4

Kuehne, H., Jhuang, H., Garrote, E., Poggio, T., Serre, T., 2011, HMDB: a large video database for human motion recognition, in: *2011 International Conference on Computer Vision*, IEEE, pp. 2556-2563.

Kuemmerle, T., Erb, K., Meyfroidt, P., Muller, D., Verburg, P. H., Estel, S., Haberl, H., Hostert, P., Jepsen, M. R., Kastner, T., Levers, C., Lindner, M., Plutzar, C., Verkerk, P. J., van der Zanden, E. H., Reenberg, A., 2013, Challenges and opportunities in mapping land use intensity globally, *Current Opinion in Environmental Sustainbility* **5**(5)**:**484-493. https://doi.org/10.1016/j.cosust.2013.06.002

Kummu, M., Taka, M., Guillaume, J. H. A., 2018, Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015, *Scientific Data* **5:**180004. https://doi.org/10.1038/sdata.2018.4

Kusno, A., 2019, Middling urbanism: the megacity and the kampung, *Urban Geography***:**1-17. https://doi.org/10.1080/02723638.2019.1688535

Kyttä, M., Broberg, A., Tzoulas, T., Snabb, K., 2013, Towards contextually sensitive urban densification: Location-based softGIS knowledge revealing perceived residential environmental quality, *Landscape and Urban Planning* **113:**30-46. https://doi.org/10.1016/j.landurbplan.2013.01.008

Labetski, A., Vitalis, S., Biljecki, F., Arroyo Ohori, K., Stoter, J., 2022, 3D building metrics for urban morphology, *International Journal of Geographical Information Science***:**1-32. https://doi.org/10.1080/13658816.2022.2103818

Lambin, E. F., Meyfroidt, P., 2011, Global land use change, economic globalization, and the looming land scarcity, *Proceedings of the National Academy of Sciences* **108**(9)**:**3465-3472. https://doi.org/10.1073/pnas.1100480108

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., Folke, C., 2001, The causes of land-use and land-cover change: moving beyond the myths, *Global Environmental Change* **11**(4)**:**261-269. https://doi.org/10.1016/S0959-3780(01)00007-3

Lao, J., Wang, C., Zhu, X., Xi, X., Nie, S., Wang, J., Cheng, F., Zhou, G., 2021, Retrieving building height in urban areas using ICESat-2 photon-counting LiDAR data, *International Journal of Applied Earth Observation and Geoinformation* **104**. https://doi.org/10.1016/j.jag.2021.102596

Lariviere, I., Lafrance, G. J. E. e., 1999, Modelling the electricity consumption of cities: effect of urban density, *Energy Economics* **21**(1)**:**53-66. https://doi.org/10.1016/S0140-9883(98)00007-3

Laroche, P. C. S. J., Schulp, C. J. E., Kastner, T., Verburg, P. H., 2020, Telecoupled environmental impacts of current and alternative Western diets, *Global Environmental Change* **62:**102066. https://doi.org/10.1016/j.gloenvcha.2020.102066

Latha, P. K., Darshana, Y., Venugopal, V., 2015, Role of building material in thermal comfort in tropical climates – A review, *Journal of Building Engineering* **3:**104-113. https://doi.org/10.1016/j.jobe.2015.06.003

Lattes, A. E., Rodríguez, J., Villa, M., 2017, Population dynamics and urbanization in Latin America: Concepts and data limitations, in: *New Forms of Urbanization*, Routledge, pp. 89-111.

Lau, S. S. Y., Giridharan, R., Ganesan, S., 2005, Multiple and intensive land use: case studies in Hong Kong, *Habitat International* **29**(3)**:**527-546. https://doi.org/10.1016/j.habitatint.2004.04.007

LeCun, Y., Bengio, Y., Hinton, G., 2015, Deep learning, *Nature* **521**(7553)**:**436-44. https://doi.org/10.1038/nature14539

Lefsky, M. A., 2010, A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System, *Geophysical Research Letters* **37**(15)**:**n/a-n/a. https://doi.org/10.1029/2010gl043622

Leichtle, T., Lakes, T., Zhu, X. X., Taubenböck, H., 2019, Has Dongying developed to a ghost city? - Evidence from multi-temporal population estimation based on VHR remote sensing and census counts, *Computers, Environment and Urban Systems* **78:**101372. https://doi.org/10.1016/j.compenvurbsys.2019.101372

Levers, C., Müller, D., Erb, K., Haberl, H., Jepsen, M. R., Metzger, M. J., Meyfroidt, P., Plieninger, T., Plutzar, C., Stürck, J., Verburg, P. H., Verkerk, P. J., Kuemmerle, T., 2015, Archetypical patterns and trajectories of land systems in Europe, *Regional Environmental Change* **18**(3)**:**715-732. https://doi.org/10.1007/s10113-015-0907-x

Levin, N., Zhang, Q., 2017, A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas, *Remote Sensing of Environment* **190:**366-382. https://doi.org/10.1016/j.rse.2017.01.006

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J., Pesaresi, M., 2019, The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use, *Earth System Science Data* **11**(3)**:**1385-1409. https://doi.org/10.5194/essd-11-1385-2019

Leyk, S., Uhl, J. H., Balk, D., Jones, B., 2018, Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States, *Remote Sensing of Environment* **204:**898-917. https://doi.org/10.1016/j.rse.2017.08.035

Li, G., Sun, S., Fang, C., 2018a, The varying driving forces of urban expansion in China: Insights from a spatial-temporal analysis, *Landscape and Urban Planning* **174:**63-77. https://doi.org/10.1016/j.landurbplan.2018.03.004

Li, M., Koks, E., Taubenböck, H., van Vliet, J., 2020, Continental-scale mapping and analysis of 3D building structure, *Remote Sensing of Environment* **245:**111859. https://doi.org/10.1016/j.rse.2020.111859

Li, M., van Vliet, J., Ke, X., Verburg, P. H., 2019, Mapping settlement systems in China and their change trajectories between 1990 and 2010, *Habitat International* **94:**102069. https://doi.org/10.1016/j.habitatint.2019.102069

Li, M., Verburg, P. H., van Vliet, J., 2022, Global trends and local variations in land take per person, *Landscape and Urban Planning* **218**. https://doi.org/10.1016/j.landurbplan.2021.104308

Li, X., Yeh, A. G. O., 2005, Integration of genetic algorithms and GIS for optimal location search, *International Journal of Geographical Information Science* **19**(5)**:**581-601. https://doi.org/10.1080/13658810500032388

Li, Y., Wu, W., Liu, Y., 2018b, Land consolidation for rural sustainability in China: Practical reflections and policy implications, *Land Use Policy* **74:**137-141. https://doi.org/10.1016/j.landusepol.2017.07.003

Liasis, G., Stavrou, S., 2016, Satellite images analysis for shadow detection and building height estimation, *ISPRS Journal of Photogrammetry and Remote Sensing* **119:**437-450. https://doi.org/10.1016/j.isprsjprs.2016.07.006

Lin, L., Homma, R., Iki, K., 2018, Preferences for a lake landscape: Effects of building height and lake width, *Environmental Impact Assessment Review* **70:**22-33. https://doi.org/10.1016/j.eiar.2018.03.001

Lipson, M. J., Nazarian, N., Hart, M. A., Nice, K. A., Conroy, B., 2022, A Transformation in City-Descriptive Input Data for Urban Climate Models, *Frontiers in Environmental Science* **10**. https://doi.org/10.3389/fenvs.2022.866398

Liu, J., Kuang, W., Zhang, Z., Xu, X., Qin, Y., Ning, J., Zhou, W., Zhang, S., Li, R., Yan, C., Wu, S., Shi, X., Jiang, N., Yu, D., Pan, X., Chi, W., 2014a, Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s, *Journal of Geographical Sciences* **24**(2)**:**195-210. https://doi.org/10.1007/s11442-014-1082-6

Liu, J., Liu, M., Zhuang, D., Zhang, Z., Deng, X., 2003, Study on spatial pattern of land-use change in China during 1995–2000, *Science in China Series D: Earth Sciences* **46**(4)**:**373-384. https://doi.org/10.1360/03yd9033

Liu, T., Yang, X., 2015, Monitoring land changes in an urban area using satellite imagery, GIS and landscape metrics, *Applied Geography* **56:**42-54. https://doi.org/10.1016/j.apgeog.2014.10.002

Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., Wang, S., 2018, High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform, *Remote Sensing of Environment* **209:**227-239. https://doi.org/10.1016/j.rse.2018.02.055

Liu, Y., Fang, F., Li, Y., 2014b, Key issues of land use in China and implications for policy making, *Land Use Policy* **40:**6-12. https://doi.org/10.1016/j.landusepol.2013.03.013

Liu, Y., Li, Y., 2017, Revitalize the world's countryside, *Nature* **548**(7667)**:**275. https://doi.org/10.1038/548275a

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., Shi, L., 2015, Social Sensing: A New Approach to Understanding Our Socioeconomic Environments, *Annals of the Association of American Geographers* **105**(3)**:**512-530. https://doi.org/10.1080/00045608.2015.1018773

Liu, Y., Liu, Y., Chen, Y., Long, H., 2010, The process and driving forces of rural hollowing in China under rapid urbanization, *Journal of Geographical Sciences* **20**(6)**:**876-888. https://doi.org/10.1007/s11442-010-0817-2

Liu, Z., He, C., Zhou, Y., Wu, J., 2014c, How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion, *Landscape Ecology* **29**(5)**:**763-771. https://doi.org/10.1007/s10980-014-0034-y

Llaguno-Munitxa, M., Bou-Zeid, E., 2020, The environmental neighborhoods of cities and their spatial extent, *Environmental Research Letters* **15**(7). https://doi.org/10.1088/1748-9326/ab8d7e

Long, H., Li, Y., Liu, Y., Woods, M., Zou, J., 2012, Accelerated restructuring in rural China fueled by 'increasing vs. decreasing balance' land-use policy for dealing with hollowed villages, *Land Use Policy* **29**(1)**:**11-22. https://doi.org/10.1016/j.landusepol.2011.04.003

Long, H., Liu, Y., Hou, X., Li, T., Li, Y., 2014, Effects of land use transitions due to rapid urbanization on ecosystem services: Implications for urban planning in the new developing area of China, *Habitat International* **44:**536-544. https://doi.org/10.1016/j.habitatint.2014.10.011

Long, H., Zou, J., Liu, Y., 2009, Differentiation of rural development driven by industrialization and urbanization in eastern coastal China, *Habitat International* **33**(4)**:**454-462. https://doi.org/10.1016/j.habitatint.2009.03.003

Loures, L., Vaz, E., 2018, Exploring expert perception towards brownfield redevelopment benefits according to their typology, *Habitat International* **72:**66-76. https://doi.org/10.1016/j.habitatint.2016.11.003

Lowry, J. H., Lowry, M. B., 2014, Comparing spatial metrics that quantify urban form, *Computers Environment and Urban Systems* **44:**59-67. https://doi.org/10.1016/j.compenvurbsys.2013.11.005

Mahiny, A. S., Clarke, K. C., 2012, Guiding SLEUTH land-use/land-cover change modeling using multicriteria evaluation: towards dynamic sustainable land-use planning, *Environment and Planning B: Planning and Design* **39**(5)**:**925-944. https://doi.org/10.1068/b37092

Mahtta, R., Mahendra, A., Seto, K. C., 2019, Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+, *Environmental Research Letters* **14**(12)**:**124077. https://doi.org/10.1088/1748-9326/ab59bf

Malek, Ž., Verburg, P., 2017, Mediterranean land systems: Representing diversity and intensity of complex land systems in a dynamic region, *Landscape and Urban Planning* **165:**102-116. https://doi.org/10.1016/j.landurbplan.2017.05.012

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R., Berger, M., 2012, Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land, *Remote Sensing of Environment* **120:**91-101. https://doi.org/10.1016/j.rse.2011.09.026

Manganelli, B., Murgante, B., 2017, The Dynamics of Urban Land Rent in Italian Regional Capital Cities, *Land* **6**(3)**:**54. https://doi.org/10.3390/land6030054

Manoli, G., Fatichi, S., Bou-Zeid, E., Katul, G. G., 2020, Seasonal hysteresis of surface urban heat islands, *Proceedings of the National Academy of Sciences* **117**(13)**:**7082-7089. https://doi.org/10.1073/pnas.1917554117

Manoli, G., Fatichi, S., Schlapfer, M., Yu, K., Crowther, T. W., Meili, N., Burlando, P., Katul, G. G., Bou-Zeid, E., 2019, Magnitude of urban heat islands largely explained by climate and population, *Nature* **573**(7772)**:**55-60. https://doi.org/10.1038/s41586-019-1512-9

Marconcini, M., Metz-Marconcini, A., Ureyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A., Paganini, M., Strano, E., 2020, Outlining where humans live, the World Settlement Footprint 2015, *Scientific Data* **7**(1)**:**242. https://doi.org/10.1038/s41597-020-00580-5

Mathews, A. J., Frazier, A. E., Nghiem, S. V., Neumann, G., Zhao, Y., 2019, Satellite scatterometer estimation of urban built-up volume: Validation with airborne lidar data, *International Journal of Applied Earth Observation and Geoinformation* **77:**100-107. https://doi.org/10.1016/j.jag.2019.01.004

McDonald, R. I., Mansur, A. V., Ascensão, F., Colbert, M. l., Crossman, K., Elmqvist, T., Gonzalez, A., Güneralp, B., Haase, D., Hamann, M., Hillel, O., Huang, K., Kahnt, B., Maddox, D., Pacheco, A., Pereira, H. M., Seto, K. C., Simkin, R., Walsh, B., Werner, A. S., Ziter, C., 2019, Research gaps in knowledge of the impact of urban growth on biodiversity, *Nature Sustainability*. https://doi.org/10.1038/s41893-019-0436-6

McFarlane, C., 2016, The geographies of urban density: Topology, politics and the city, *Progress in Human Geography* **40**(5)**:**629-648. https://doi.org/10.1177/0309132515608694

McFarlane, C., 2019, The Urbanization of the Sanitation Crisis: Placing Waste in the City, *Development and Change* **50**(5)**:**1239-1262. https://doi.org/10.1111/dech.12533

McGarigal, K., Cushman, S., 2005, The gradient concept of landscape structure [Chapter 12], *In: Wiens, John A.; Moss, Michael R., eds. Issues and Perspectives in Landscape Ecology. Cambridge University Press. p. 112-119.***:**112-119.

McGinnis, M. D., Ostrom, E., 2014, Social-ecological system framework: initial changes and continuing challenges, *Ecology and Society* **19**(2). https://doi.org/10.5751/ES-06387-190230

Meijer, J., Huijbregts, M. A., Schotten, K., Schipper, A., 2018, Global patterns of current and future road infrastructure, *Environmental Research Letters* **13**(6)**:**064006. https://doi.org/10.1088/1748-9326/aabd42

Melchiorri, M., Florczyk, A., Freire, S., Schiavina, M., Pesaresi, M., Kemper, T., 2018, Unveiling 25 Years of Planetary Urbanization with Remote Sensing: Perspectives from the Global Human Settlement Layer, *Remote Sensing* **10**(5). https://doi.org/10.3390/rs10050768

Melchiorri, M., Pesaresi, M., Florczyk, A., Corbane, C., Kemper, T., 2019, Principles and Applications of the Global Human Settlement Layer as Baseline for the Land Use Efficiency Indicator—SDG 11.3.1, *ISPRS International Journal of Geo-Information* **8**(2)**:**96. https://doi.org/10.3390/ijgi8020096

Mertes, C. M., Schneider, A., Sulla-Menashe, D., Tatem, A. J., Tan, B., 2015, Detecting change in urban areas at continental scales with MODIS data, *Remote Sensing of Environment* **158:**331-347. https://doi.org/10.1016/j.rse.2014.09.023

Messerli, P., Heinimann, A., Epprecht, M., 2009, Finding homogeneity in heterogeneity—A new approach to quantifying landscape mosaics developed for the Lao PDR, *Human Ecology* **37**(3)**:**291-304. https://doi.org/10.1007/s10745-009-9238-1

Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E. C., Erb, K. H., Filatova, T., Garrett, R. D., Grove, J. M., Heinimann, A., Kuemmerle, T., Kull, C. A., Lambin, E. F., Landon, Y., le Polain de Waroux, Y., Messerli, P., Müller, D., Nielsen, J. Ø., Peterson, G. D., Rodriguez García, V., Schlüter, M., Turner, B. L., Verburg, P. H., 2018, Middle-range theories of land system change, *Global Environmental Change* **53:**52-67. https://doi.org/10.1016/j.gloenvcha.2018.08.006

Michishita, R., Jiang, Z., Xu, B., 2012, Monitoring two decades of urbanization in the Poyang Lake area, China through spectral unmixing, *Remote Sensing of Environment* **117:**3-18. https://doi.org/10.1016/j.rse.2011.06.021

Microsoft, 2022, GlobalMLBuildingFootprints, Github.

Middel, A., Häb, K., Brazel, A. J., Martin, C. A., Guhathakurta, S., 2014, Impact of urban form and design on mid-afternoon microclimate in Phoenix Local Climate Zones, *Landscape and Urban Planning* **122:**16-28. https://doi.org/10.1016/j.landurbplan.2013.11.004

Mills, S., Weiss, S., Liang, C., 2013, VIIRS day/night band (DNB) stray light characterization and correction, *Earth Observing Systems XVIII*. https://doi.org/10.1117/12.2023107

Mohammed, N. Z., Ghazi, A., Mustafa, H. E., 2013, Positional accuracy testing of Google Earth, *International Journal of Multidisciplinary Sciences and Engineering* **4**(6)**:**6-9.

Mozumder, C., Tripathi, N. K., Losiri, C., 2016, Comparing three transition potential models: A case study of built-up transitions in North-East India, *Computers, Environment and Urban Systems* **59:**38-49. https://doi.org/10.1016/j.compenvurbsys.2016.04.009

Müller, D., Sun, Z., Vongvisouk, T., Pflugmacher, D., Xu, J., Mertz, O., 2014, Regime shifts limit the predictability of land-system change, *Global Environmental Change* **28:**75-83. https://doi.org/10.1016/j.gloenvcha.2014.06.003

Mundia, C. N., Aniya, M., 2005, Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS, *International journal of Remote sensing* **26**(13)**:**2831-2849. https://doi.org/10.1080/01431160500117865

Mushore, T. D., Odindi, J., Dube, T., Mutanga, O., 2017, Prediction of future urban surface temperatures using medium resolution satellite data in Harare metropolitan city, Zimbabwe, *Building and Environment* **122:**397-410. https://doi.org/10.1016/j.buildenv.2017.06.033

Mustafa, A., Van Rompaey, A., Cools, M., Saadi, I., Teller, J., 2018, Addressing the determinants of built-up expansion and densification processes at the regional scale, *Urban Studies*. https://doi.org/10.1177/0042098017749176

Næss, P., Saglie, I.-L., Richardson, T., 2019, Urban sustainability: is densification sufficient?, *European Planning Studies* **28**(1)**:**146-165. https://doi.org/10.1080/09654313.2019.1604633

Nagendra, H., Bai, X., Brondizio, E. S., Lwasa, S., 2018, The urban south and the predicament of global sustainability, *Nature Sustainability* **1**(7)**:**341-349. https://doi.org/10.1038/s41893-018-0101-5

Nethercote, M., 2019, Melbourne's vertical expansion and the political economies of high-rise residential development, *Urban Studies* **56**(16)**:**3394-3414. https://doi.org/10.1177/0042098018817225

Ornetsmüller, C., Heinimann, A., Verburg, P. H., 2018, Operationalizing a land systems classification for Laos, *Landscape and Urban Planning* **169:**229-240. https://doi.org/10.1016/j.landurbplan.2017.09.018

Over, M., Schilling, A., Neubauer, S., Zipf, A., 2010, Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany, *Computers, Environment and Urban Systems* **34**(6)**:**496-507. https://doi.org/10.1016/j.compenvurbsys.2010.05.001

Paprotny, D., Kreibich, H., Morales-Nápoles, O., Terefenko, P., Schröter, K., 2020, Estimating exposure of residential assets to natural hazards in Europe using open data, *Natural Hazards and Earth System Sciences* **20**(1)**:**323-343. https://doi.org/10.5194/nhess-20-323-2020

Park, J., 2014, Land rent theory revisited, *Science & Society* **78**(1)**:**88-109. https://doi.org/10.1521/siso.2014.78.1.88

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* **12**(Oct)**:**2825-2830.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A. S., 2016, High-resolution mapping of global surface water and its long-term changes, *Nature* **540:**418. https://doi.org/10.1038/nature20584

Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016, Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas, *Remote Sensing of Environment* **187:**156-168. https://doi.org/10.1016/j.rse.2016.10.010

Perrin, C., Nougarèdes, B., Sini, L., Branduini, P., Salvati, L., 2018, Governance changes in peri-urban farmland protection following decentralisation: A comparison between Montpellier (France) and Rome (Italy), *Land Use Policy* **70:**535-546. https://doi.org/10.1016/j.landusepol.2017.09.027

Pesaresi, M., Ehrlich, D., Florczyk, A. J., Freire, S., Julea, A., Kemper, T., Syrris, V., 2016, The global human settlement layer from landsat imagery, in: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, pp. 7276-7279.

Poelmans, L., Rompaey, A. V., Ntegeka, V., Willems, P., 2011, The relative impact of climate change and urban expansion on peak flows: a case study in central Belgium, *Hydrological Processes* **25**(18)**:**2846-2858. https://doi.org/10.1002/hyp.8047

Potere, D., Schneider, A., Angel, S., Civco, D. L., 2009, Mapping urban areas on a global scale: which of the eight maps now available is more accurate?, *International Journal of Remote Sensing* **30**(24)**:**6531-6558. https://doi.org/10.1080/01431160903121134

Pulighe, G., Baiocchi, V., Lupia, F., 2015, Horizontal accuracy assessment of very high resolution Google Earth images in the city of Rome, Italy, *International Journal of Digital Earth* **9**(4)**:**342-362. https://doi.org/10.1080/17538947.2015.1031716

Ramachandra, T. V., Bharath, A. H., Sowmyashree, M. V., 2015, Monitoring urbanization and its implications in a mega city from space: spatiotemporal patterns and its indicators, *Journal of Environmental Management* **148:**67-81. https://doi.org/10.1016/j.jenvman.2014.02.015

Ramankutty, N., Evan, A. T., Monfreda, C., Foley, J. A., 2008, Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, *Global Biogeochemical Cycles* **22**(1)**:**1-19. https://doi.org/10.1029/2007gb002952

Ramaswami, A., Russell, A. G., Culligan, P. J., Sharma, K. R., Kumar, E., 2016, Meta-principles for developing smart, sustainable, and healthy cities, *Science* **352**(6288)**:**940-943. https://doi.org/10.1126/science.aaf7160

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019, Deep learning and process understanding for data-driven Earth system science, *Nature* **566**(7743)**:**195-204. https://doi.org/10.1038/s41586-019-0912-1

Ren, Q., He, C., Huang, Q., Shi, P., Zhang, D., Güneralp, B., 2022, Impacts of urban expansion on natural habitats in global drylands, *Nature Sustainability*. https://doi.org/10.1038/s41893-022-00930-8

Richards, D. R., Passy, P., Oh, R. R. Y., 2017, Impacts of population density and wealth on the quantity and structure of urban green space in tropical Southeast Asia, *Landscape and Urban Planning* **157:**553-560. https://doi.org/10.1016/j.landurbplan.2016.09.005

Richter, S. M., 2020, Revisiting urban expansion in the continental United States, *Landscape and Urban Planning* **204:**103911. https://doi.org/10.1016/j.landurbplan.2020.103911

Rosier, J. F., Taubenböck, H., Verburg, P. H., van Vliet, J., 2022, Fusing Earth observation and socioeconomic data to increase the transferability of large-scale urban land use classification, *Remote Sensing of Environment* **278**. https://doi.org/10.1016/j.rse.2022.113076

Rounsevell, M. D. A., Pedroli, B., Erb, K.-H., Gramberger, M., Busck, A. G., Haberl, H., Kristensen, S., Kuemmerle, T., Lavorel, S., Lindner, M., Lotze-Campen, H., Metzger, M. J., Murray-Rust, D., Popp, A., Pérez-Soba, M., Reenberg, A., Vadineanu, A., Verburg, P. H., Wolfslehner, B., 2012, Challenges for land system science, *Land Use Policy* **29**(4)**:**899-910. https://doi.org/10.1016/j.landusepol.2012.01.007

Sabo, F., Corbane, C., Florczyk, A. J., Ferri, S., Pesaresi, M., Kemper, T., 2018, Comparison of built - up area maps produced within the global human settlement framework, *Transactions in GIS* **22**(6)**:**1406-1436. https://doi.org/10.1111/tgis.12480

Salem, M., Tsurusaki, N., Divigalpitiya, P., 2020, Land use/land cover change detection and urban sprawl in the peri-urban area of greater Cairo since the Egyptian revolution of 2011, *Journal of Land Use Science* **15**(5)**:**592-606. https://doi.org/10.1080/1747423x.2020.1765425

Samuel, D. G. L., Nagendra, S. M. S., Maiya, M. P., 2013, Passive alternatives to mechanical air conditioning of building: A review, *Building and Environment* **66:**54-64. https://doi.org/10.1016/j.buildenv.2013.04.016

Schiavina, M., Freire, S., MacManus, K., 2019, GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A, *Joint Research Centre (JRC), European Commission*.

Schindler, M., Caruso, G., 2014, Urban compactness and the trade-off between air pollution emission and exposure: Lessons from a spatially explicit theoretical model, *Computers, Environment and Urban Systems* **45:**13-23.

Schmitt, M., Hughes, L. H., Qiu, C., Zhu, X. X., 2019, SEN12MS--A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion, *arXiv preprint arXiv:1906.07789*.

Schneider, A., Chang, C., Paulsen, K., 2015, The changing spatial form of cities in Western China, *Landscape and Urban Planning* **135:**40-61. https://doi.org/10.1016/j.landurbplan.2014.11.005

Schneider, A., Friedl, M. A., Potere, D., 2009, A new map of global urban extent from MODIS satellite data, *Environmental Research Letters* **4**(4)**:**044003. https://doi.org/10.1088/1748-9326/4/4/044003

Schneider, A., Friedl, M. A., Potere, D., 2010, Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions', *Remote Sensing of Environment* **114**(8)**:**1733-1746. https://doi.org/10.1016/j.rse.2010.03.003

Schneider, A., Mertes, C., 2014, Expansion and growth in Chinese cities, 1978–2010, *Environmental Research Letters* **9**(2)**:**024008.

Schneider, A., Seto, K. C., Webster, D. R. J. E., Planning, P. B., Design, 2005, Urban growth in Chengdu, Western China: application of remote sensing to assess planning and policy outcomes, **32**(3)**:**323-345.

Schug, F., 2021, A multi-dimensional characterization of settlements with Earth Observation data, *Humboldt-Universität zu Berlin*.

Schug, F., Frantz, D., Okujeni, A., van der Linden, S., Hostert, P., 2020, Mapping urban-rural gradients of settlements and vegetation at national scale using Sentinel-2 spectral-temporal metrics and regression-based unmixing with synthetic training data, *Remote Sensing of Environment* **246:**111810. https://doi.org/10.1016/j.rse.2020.111810

Schwarz, N., 2010, Urban form revisited—Selecting indicators for characterising European cities, *Landscape and Urban Planning* **96**(1)**:**29-47. https://doi.org/10.1016/j.landurbplan.2010.01.007

Sclar, E. D., Garau, P., Carolini, G., 2005, The 21st century health challenge of slums and cities, *The Lancet* **365**(9462)**:**901-903. https://doi.org/10.1016/s0140-6736(05)71049-7

See, L., Bayas, J. C. L., Lesiv, M., Schepaschenko, D., Danylo, O., McCallum, I., Dürauer, M., Georgieva, I., Domian, D., Fraisl, D., Hager, G., Karanam, S., Moorthy, I., Sturn, T., Subash, A., Fritz, S., 2022a, Lessons learned in developing reference data sets with the contribution of citizens: the Geo-Wiki experience, *Environmental Research Letters* **17**(6). https://doi.org/10.1088/1748-9326/ac6ad7

See, L., Georgieva, I., Duerauer, M., Kemper, T., Corbane, C., Maffenini, L., Gallego, J., Pesaresi, M., Sirbu, F., Ahmed, R., Blyshchyk, K., Magori, B., Blyshchyk, V., Melnyk, O., Zadorozhniuk, R., Mandici, M. T., Su, Y. F., Rabia, A. H., Perez-Hoyos, A., Vasylyshyn, R., Pawe, C. K., Bilous, S., Kovalevskyi, S. B., Kovalevskyi, S. S., Bordoloi, K., Bilous, A., Panging, K., Bilous, V., Prestele, R., Sahariah, D., Deka, A., Nath, N., Neves, R., Myroniuk, V., Karner, M., Fritz, S., 2022b, A crowdsourced global data set for validating built-up surface layers, *Scientific Data* **9**(1)**:**13. https://doi.org/10.1038/s41597-021-01105-4

Seto, K. C., Christensen, P., 2013, Remote sensing science to inform urban climate change mitigation strategies, *Urban Climate* **3:**1-6. https://doi.org/10.1016/j.uclim.2013.03.001

Seto, K. C., Fragkias, M., 2005, Quantifying Spatiotemporal Patterns of Urban Land-use Change in Four Cities of China with Time Series Landscape Metrics, *Landscape Ecology* **20**(7)**:**871-888. https://doi.org/10.1007/s10980-005-5238-8

Seto, K. C., Fragkias, M., Güneralp, B., Reilly, M. K., 2011, A meta-analysis of global urban land expansion, *PloS One* **6**(8)**:**e23777. https://doi.org/10.1371/journal.pone.0023777

Seto, K. C., Guneralp, B., Hutyra, L. R., 2012a, Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools, *Proceedings of the National Academy of Sciences* **109**(40)**:**16083-8. https://doi.org/10.1073/pnas.1211658109

Seto, K. C., Pandey, B., 2019, Urban Land Use: Central to Building a Sustainable Future, *One Earth* **1**(2)**:**168-170. https://doi.org/10.1016/j.oneear.2019.10.002

Seto, K. C., Ramankutty, N., 2016, Hidden linkages between urbanization and food systems, *Science* **352**(6288)**:**943-945. https://doi.org/10.1126/science.aaf7439

Seto, K. C., Reenberg, A., Boone, C. G., Fragkias, M., Haase, D., Langanke, T., Marcotullio, P., Munroe, D. K., Olah, B., Simon, D., 2012b, Urban land teleconnections and sustainability, *Proceedings of the National Academy of Sciences* **109**(20)**:**7687-92. https://doi.org/10.1073/pnas.1117622109

Seto, K. C., Sánchez-Rodríguez, R., Fragkias, M., 2010, The New Geography of Contemporary Urbanization and the Environment, *Annual Review of Environment and Resources* **35**(1)**:**167-194. https://doi.org/10.1146/annurev-environ-100809-125336

Seto, K. C., Shepherd, J. M., 2009, Global urban land-use trends and climate impacts, *Current Opinion in Environmental Sustainability* **1**(1)**:**89-95. https://doi.org/10.1016/j.cosust.2009.07.012

Shaw, B. J., van Vliet, J., Verburg, P. H., 2020, The peri-urbanization of Europe: A systematic review of a multifaceted process, *Landscape and Urban Planning* **196:**103733. https://doi.org/10.1016/j.landurbplan.2019.103733

Shi, L., Taubenböck, H., Zhang, Z., Liu, F., Wurm, M., 2017, Urbanization in China from the end of 1980s until 2010 – spatial dynamics and patterns of growth using EO-data, *International Journal of Digital Earth***:**1-17. https://doi.org/10.1080/17538947.2017.1400599

Simard, M., Pinto, N., Fisher, J. B., Baccini, A., 2011, Mapping forest canopy height globally with spaceborne lidar, *Journal of Geophysical Research* **116**(G4). https://doi.org/10.1029/2011jg001708

Sliuzas, R., Kuffer, M., Kemper, T., 2017, Assessing the quality of Global Human Settlement Layer products for Kampala, Uganda, in: *2017 Joint Urban Remote Sensing Event (JURSE)*, IEEE, pp. 1-4.

Solt, F., 2020, Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database, *Social Science Quarterly* **101**(3)**:**1183-1199. https://doi.org/10.1111/ssqu.12795

Soomro, K., Zamir, A. R., Shah, M., 2012, UCF101: A dataset of 101 human actions classes from videos in the wild, *arXiv preprint arXiv:1212.0402*.

Stewart, I. D., Oke, T. R., 2012, Local Climate Zones for Urban Temperature Studies, *Bulletin of the American Meteorological Society* **93**(12)**:**1879-1900. https://doi.org/10.1175/bams-d-11-00019.1

Stokes, E. C., Seto, K. C., 2019, Characterizing urban infrastructural transitions for the Sustainable Development Goals using multi-temporal land, population, and nighttime light data, *Remote Sensing of Environment* **234**. https://doi.org/10.1016/j.rse.2019.111430

Straka, M., Sodoudi, S., 2019, Evaluating climate change adaptation strategies and scenarios of enhanced vertical and horizontal compactness at urban scale (a case study for Berlin), *Landscape and Urban Planning* **183:**68-78. https://doi.org/10.1016/j.landurbplan.2018.11.006

Stürck, J., Levers, C., van der Zanden, E. H., Schulp, C. J. E., Verkerk, P. J., Kuemmerle, T., Helming, J., Lotze-Campen, H., Tabeau, A., Popp, A., Schrammeijer, E., Verburg, P., 2015,

Simulating and delineating future land change trajectories across Europe, *Regional Environmental Change*. https://doi.org/10.1007/s10113-015-0876-0

Sumbul, G., Charfuelan, M., Demir, B., Markl, V., 2019, BigEarthNet: A Large-Scale Benchmark Archive For Remote Sensing Image Understanding, *arXiv preprint arXiv:1902.06148*.

Susaki, J., Kajimoto, M., Kishimoto, M., 2014, Urban density mapping of global megacities from polarimetric SAR images, *Remote Sensing of Environment* **155:**334-348. https://doi.org/10.1016/j.rse.2014.09.006

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., Feuston, B. P., 2003, Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling, *Journal of Chemical Information and Computer Sciences* **43**(6)**:**1947-1958. https://doi.org/10.1021/ci034160g

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D., Oimoen, M., Zhang, Z., Danielson, J., Krieger, T., Curtis, B., Haase, J., 2011, ASTER Global Digital Elevation Model Version 2–Summary of Validation Results August 31, 2011.

Tappert, S., Klöti, T., Drilling, M., 2018, Contested urban green spaces in the compact city: The (re-)negotiation of urban gardening in Swiss cities, *Landscape and Urban Planning* **170:**69-78. https://doi.org/10.1016/j.landurbplan.2017.08.016

Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y., Zhu, X. X., 2020, Seven city types representing morphologic configurations of cities across the globe, *Cities* **105**. https://doi.org/10.1016/j.cities.2020.102814

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S., 2012, Monitoring urbanization in mega cities from space, *Remote Sensing of Environment* **117:**162-176. https://doi.org/10.1016/j.rse.2011.09.015

Taubenböck, H., Kraff, N. J., Wurm, M., 2018a, The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data, *Applied Geography* **92:**150-167. https://doi.org/10.1016/j.apgeog.2018.02.002

Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., Dech, S., 2019, A new ranking of the world's largest cities—Do administrative units obscure morphological realities?, *Remote Sensing of Environment* **232:**111353. https://doi.org/10.1016/j.rse.2019.111353

Taubenböck, H., Wurm, M., Geiß, C., Dech, S., Siedentop, S., 2018b, Urbanization between compactness and dispersion: designing a spatial model for measuring 2D binary settlement landscape configurations, *International Journal of Digital Earth***:**1-20. https://doi.org/10.1080/17538947.2018.1474957

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., Slayback, D. A., 2021, Satellite imaging reveals increased proportion of population exposed to floods, *Nature* **596**(7870)**:**80-86. https://doi.org/10.1038/s41586-021-03695-w

Theobald, D. M., 2001, Land-use dynamics beyond the American urban fringe, *Geographical Review* **91**(3)**:**544-564. https://doi.org/10.1111/j.1931-0846.2001.tb00240.x

Thorn, J., Thornton, T. F., Helfgott, A., 2015, Autonomous adaptation to global environmental change in peri-urban settlements: Evidence of a growing culture of innovation and revitalisation in Mathare Valley Slums, Nairobi, *Global Environmental Change* **31:**121-131. https://doi.org/10.1016/j.gloenvcha.2014.12.009

Tian, G., Jiang, J., Yang, Z., Zhang, Y., 2011, The urban growth, size distribution and spatio-temporal dynamic pattern of the Yangtze River Delta megalopolitan region, China, *Ecological Modelling* **222**(3)**:**865-878. https://doi.org/10.1016/j.ecolmodel.2010.09.036

Tian, G., Yang, Z., Zhang, Y., 2007, The spatio-temporal dynamic pattern of rural residential land in China in the 1990s using Landsat TM images and GIS, *Environmental Management* **40**(5)**:**803-13. https://doi.org/10.1007/s00267-006-0048-6

Tierolf, L., de Moel, H., van Vliet, J., 2021, Modeling urban development and its exposure to river flood risk in Southeast Asia, *Computers, Environment and Urban Systems* **87**. https://doi.org/10.1016/j.compenvurbsys.2021.101620

Tietjen, A., Jørgensen, G., 2016, Translating a wicked problem: A strategic planning approach to rural shrinkage in Denmark, *Landscape and Urban Planning* **154:**29-43. https://doi.org/10.1016/j.landurbplan.2016.01.009

Tomás, L., Fonseca, L., Almeida, C., Leonardi, F., Pereira, M., 2015, Urban population estimation based on residential buildings volume using IKONOS-2 images and lidar data, *International Journal of Remote Sensing* **37**(sup1)**:**1-28. https://doi.org/10.1080/01431161.2015.1121301

Tóth, G., 2012, Impact of land-take on the land resource base for crop production in the European Union, *Science of the Total Environment* **435:**202-214. https://doi.org/10.1016/j.scitotenv.2012.06.103

Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E., Papale, D., 2015, Uncertainty analysis of gross primary production upscaling using Random Forests, remote sensing and eddy covariance data, *Remote Sensing of Environment* **168:**360-373. https://doi.org/10.1016/j.rse.2015.07.015

Turner, B. L., Meyer, W. B., Skole, D. L., 1994, Global land-use/land-cover change: towards an integrated study, in: *Ambio*, pp. 91-95.

United Nations, 2016, World Cities Report 2016: Urbanization and Development–Emerging Futures, *Publisher: UN-Habitat*.

United Nations, 2018, 2018 revision of world urbanization prospects, and 2017 revision of world population prospects.

Václavík, T., Lautenbach, S., Kuemmerle, T., Seppelt, R., 2013, Mapping global land system archetypes, *Global Environmental Change* **23**(6)**:**1637-1647. https://doi.org/10.1016/j.gloenvcha.2013.09.004

van Asselen, S., Verburg, P. H., 2012, A Land System representation for global assessments and land-use modeling, *Global Change Biology* **18**(10)**:**3125-3148. https://doi.org/10.1111/j.1365-2486.2012.02759.x

van Asselen, S., Verburg, P. H., 2013, Land cover change or land-use intensification: simulating land system change with a global-scale land change model, *Global Change Biology* **19**(12)**:**3648-67. https://doi.org/10.1111/gcb.12331

Van de Voorde, T., Jacquet, W., Canters, F., 2011, Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data, *Landscape and Urban Planning* **102**(3)**:**143-155. https://doi.org/10.1016/j.landurbplan.2011.03.017

Van der Horst, D., 2006, Spatial cost–benefit thinking in multi-functional forestry; towards a framework for spatial targeting of policy interventions, *Ecological economics* **59**(1)**:**171-180. https://doi.org/10.1016/j.ecolecon.2005.10.005

van der Wal, D., Herman, P. M. J., Wielemaker-van den Dool, A., 2005, Characterisation of surface roughness and sediment texture of intertidal flats using ERS SAR imagery, *Remote Sensing of Environment* **98**(1)**:**96-109. https://doi.org/10.1016/j.rse.2005.06.004

van der Zanden, E. H., Levers, C., Verburg, P. H., Kuemmerle, T., 2016, Representing composition, spatial structure and management intensity of European agricultural landscapes: A new typology, *Landscape and Urban Planning* **150:**36-49. https://doi.org/10.1016/j.landurbplan.2016.02.005

van Vliet, J., 2019, Direct and indirect loss of natural area from urban expansion, *Nature Sustainability* **2**(8)**:**755-763. https://doi.org/10.1038/s41893-019-0340-0

van Vliet, J., Birch-Thomsen, T., Gallardo, M., Hemerijckx, L.-M., Hersperger, A. M., Li, M., Tumwesigye, S., Twongyirwe, R., van Rompaey, A., 2020, Bridging the rural-urban dichotomy in land use science, *Journal of Land Use Science* **15**(5)**:**585-591. https://doi.org/10.1080/1747423x.2020.1829120

van Vliet, J., Bregt, A. K., Brown, D. G., van Delden, H., Heckbert, S., Verburg, P. H., 2016a, A review of current calibration and validation practices in land-change modeling, *Environmental Modelling & Software* **82:**174-182.

van Vliet, J., de Groot, H. L. F., Rietveld, P., Verburg, P. H., 2015, Manifestations and underlying drivers of agricultural land use change in Europe, *Landscape and Urban Planning* **133:**24-36. https://doi.org/10.1016/j.landurbplan.2014.09.001

van Vliet, J., Eitelberg, D. A., Verburg, P. H., 2017, A global analysis of land take in cropland areas and production displacement from urbanization, *Global Environmental Change* **43:**107-115. https://doi.org/10.1016/j.gloenvcha.2017.02.001

van Vliet, J., Magliocca, N. R., Buchner, B., Cook, E., Rey Benayas, J. M., Ellis, E. C., Heinimann, A., Keys, E., Lee, T. M., Liu, J., Mertz, O., Meyfroidt, P., Moritz, M., Poeplau, C., Robinson, B. E., Seppelt, R., Seto, K. C., Verburg, P. H., 2016b, Meta-studies in land use science: Current coverage and prospects, *Ambio* **45**(1)**:**15-28. https://doi.org/10.1007/s13280-015-0699-8

van Vliet, J., Verburg, P. H., Grădinaru, S. R., Hersperger, A. M., 2019, Beyond the urban-rural dichotomy: Towards a more nuanced analysis of changes in built-up land, *Computers, Environment and Urban Systems* **74:**41-49. https://doi.org/10.1016/j.compenvurbsys.2018.12.002

Vannier, C., Bierry, A., Longaretti, P.-Y., Nettier, B., Cordonnier, T., Chauvin, C., Bertrand, N., Quétier, F., Lasseur, R., Lavorel, S., 2019, Co-constructing future land-use scenarios for the Grenoble region, France, *Landscape and Urban Planning* **190:**103614. https://doi.org/10.1016/j.landurbplan.2019.103614

Verburg, P. H., Crossman, N., Ellis, E. C., Heinimann, A., Hostert, P., Mertz, O., Nagendra, H., Sikor, T., Erb, K.-H., Golubiewski, N., Grau, R., Grove, M., Konaté, S., Meyfroidt, P., Parker, D. C., Chowdhury, R. R., Shibata, H., Thomson, A., Zhen, L., 2015, Land system science and sustainable development of the earth system: A global land project perspective, *Anthropocene* **12:**29-41. https://doi.org/10.1016/j.ancene.2015.09.004

Verburg, P. H., Erb, K. H., Mertz, O., Espindola, G., 2013, Land System Science: between global challenges and local realities, *Current Opinion in Environmental Sustainability* **5**(5)**:**433-437. https://doi.org/10.1016/j.cosust.2013.08.001

Verburg, P. H., Neumann, K., Nol, L., 2011, Challenges in using land use and land cover data for global change studies, *Global Change Biology* **17**(2)**:**974-989. https://doi.org/10.1111/j.1365-2486.2010.02307.x

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S. S. A., 2002, Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model, *Environmental Management* **30**(3)**:**391-405. https://doi.org/10.1007/s00267-002-2630-x

Vermeiren, K., Van Rompaey, A., Loopmans, M., Serwajja, E., Mukwaya, P., 2012, Urban growth of Kampala, Uganda: Pattern analysis and scenario development, *Landscape and Urban Planning* **106**(2)**:**199-206. https://doi.org/10.1016/j.landurbplan.2012.03.006

Visser, H., de Nijs, T., 2006, The Map Comparison Kit, *Environmental Modelling & Software* **21**(3)**:**346-358. https://doi.org/10.1016/j.envsoft.2004.11.013

von Thünen, J., 1826, Isolated State. Translated by Carla M. Wartenberg, Pergamon, Oxford.

Wan, Z., Hook, S., Hulley, G., 2015, MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006, *NASA EOSDIS Land Processes DAAC* **10**. https://doi.org/10.5067/MODIS/MOD11A2.006

Wang, J., Hadjikakou, M., Hewittc, J. R., Bryan, A. B., 2022, Simulating large-scale urban land-use pattern and dynamics using the U-Net deep learning architecture, *Computers, Environment and Urban Systems*. https://doi.org/10.1016/j.compenvurbsys.2022.101855

Wang, J., Kuffer, M., Roy, D., Pfeffer, K., 2019a, Deprivation pockets through the lens of convolutional neural networks, *Remote Sensing of Environment* **234:**111448. https://doi.org/10.1016/j.rse.2019.111448

Wang, L., Wang, S., Zhou, Y., Liu, W., Hou, Y., Zhu, J., Wang, F., 2018, Mapping population density in China between 1990 and 2010 using remote sensing, *Remote Sensing of Environment* **210:**269-281. https://doi.org/10.1016/j.rse.2018.03.007

Wang, S., Bai, X., Zhang, X., Reis, S., Chen, D., Xu, J., Gu, B., 2021a, Urbanization can benefit agricultural production with large-scale farming in China, *Nature Food* **2**(3)**:**183-191. https://doi.org/10.1038/s43016-021-00228-6

Wang, X.-R., Hui, E. C.-M., Choguill, C., Jia, S.-H., 2015, The new urbanization policy in China: Which way forward?, *Habitat International* **47:**279-284. https://doi.org/10.1016/j.habitatint.2015.02.001

Wang, Y., Liu, Y., Li, Y., Li, T., 2016, The spatio-temporal patterns of urban–rural development transformation in China since 1990, *Habitat International* **53:**178-187. https://doi.org/10.1016/j.habitatint.2015.11.011

Wang, Y., van Vliet, J., Debonne, N., Pu, L., Verburg, P. H., 2021b, Settlement changes after peak population: Land system projections for China until 2050, *Landscape and Urban Planning* **209**. https://doi.org/10.1016/j.landurbplan.2021.104045

Wang, Y., van Vliet, J., Pu, L., Verburg, P. H., 2019b, Modeling different urban change trajectories and their trade-offs with food production in Jiangsu Province, China, *Computers, Environment and Urban Systems* **77:**101355. https://doi.org/10.1016/j.compenvurbsys.2019.101355

Watson, R. T., Noble, I. R., Bolin, B., Ravindranath, N., Verardo, D. J., Dokken, D. J., 2000, IPCC special report on land use, land-use change, and forestry.

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., Battle, K. E., Bhatt, S., Gething, P. W., 2018, A global map of travel time to cities to assess inequalities in accessibility in 2015, *Nature* **553**(7688)**:**333-336. https://doi.org/10.1038/nature25181

Weissbrod, L., Marshall, F. B., Valla, F. R., Khalaily, H., Bar-Oz, G., Auffray, J. C., Vigne, J. D., Cucchi, T., 2017, Origins of house mice in ecological niches created by settled hunter-

gatherers in the Levant 15,000 y ago, *Proceedings of the National Academy of Sciences* **114**(16)**:**4099-4104. https://doi.org/10.1073/pnas.1619137114

Weissgerber, F., Colin-Koeniguer, E., Nicolas, J.-M., Trouvé, N., 2017, 3D Monitoring of Buildings Using TerraSAR-X InSAR, DInSAR and PolSAR Capacities, *Remote Sensing* **9**(10)**:**1-24. https://doi.org/10.3390/rs9101010

Weissteiner, C. J., Boschetti, M., Böttcher, K., Carrara, P., Bordogna, G., Brivio, P. A., 2011, Spatial explicit assessment of rural land abandonment in the Mediterranean area, *Global and Planetary Change* **79**(1-2)**:**20-36. https://doi.org/10.1016/j.gloplacha.2011.07.009

Weng, Q., 2012, Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends, *Remote Sensing of Environment* **117:**34-49. https://doi.org/10.1016/j.rse.2011.02.030

Wentz, E. A., York, A. M., Alberti, M., Conrow, L., Fischer, H., Inostroza, L., Jantz, C., Pickett, S. T. A., Seto, K. C., Taubenböck, H., 2018, Six fundamental aspects for conceptualizing multidimensional urban form: A spatial mapping perspective, *Landscape and Urban Planning* **179:**55-62. https://doi.org/10.1016/j.landurbplan.2018.07.007

Wetherley, E. B., Roberts, D. A., McFadden, J. P., 2017, Mapping spectrally similar urban materials at sub-pixel scales, *Remote Sensing of Environment* **195:**170-183. https://doi.org/10.1016/j.rse.2017.04.013

Winsemius, H. C., Aerts, Jeroen C. J. H., van Beek, Ludovicus P. H., Bierkens, Marc F. P., Bouwman, A., Jongman, B., Kwadijk, Jaap C. J., Ligtvoet, W., Lucas, Paul L., van Vuuren, Detlef P., Ward, Philip J., 2015, Global drivers of future river flood risk, *Nature Climate Change* **6**(4)**:**381-385. https://doi.org/10.1038/nclimate2893

Wolff, M., Haase, D., Haase, A., 2018a, Compact or spread? A quantitative spatial model of urban areas in Europe since 1990, *PLoS One* **13**(2)**:**e0192326. https://doi.org/10.1371/journal.pone.0192326

Wolff, S., Schrammeijer, E. A., Schulp, C. J., Verburg, P. H., 2018b, Meeting global land restoration and protection targets: What would the world look like in 2050?, *Global Environmental Change* **52:**259-272. https://doi.org/10.1016/j.gloenvcha.2018.08.002

Wu, W., Zhao, S., Zhu, C., Jiang, J., 2015, A comparative study of urban expansion in Beijing, Tianjin and Shijiazhuang over the past three decades, *Landscape and Urban Planning* **134:**93-106. https://doi.org/10.1016/j.landurbplan.2014.10.010

Wu, Y., Chen, Y., Deng, X., Hui, E. C. M., 2018, Development of characteristic towns in China, *Habitat International* **77:**21-31. https://doi.org/10.1016/j.habitatint.2017.12.008

Xia, C., Yeh, A. G.-O., Zhang, A., 2020, Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities, *Landscape and Urban Planning* **193:**103669. https://doi.org/10.1016/j.landurbplan.2019.103669

Xiong, G., Cao, X., Hamm, N. A. S., Lin, T., Zhang, G., Chen, B., 2021, Unbalanced Development Characteristics and Driving Mechanisms of Regional Urban Spatial Form: A Case Study of Jiangsu Province, China, *Sustainability* **13**(6). https://doi.org/10.3390/su13063121

Xu, C., Liu, M., Zhang, C., An, S., Yu, W., Chen, J. M., 2007, The spatiotemporal dynamics of rapid urban growth in the Nanjing metropolitan region of China, *Landscape Ecology* **22**(6)**:**925-937. https://doi.org/10.1007/s10980-007-9079-5

Xu, G., Jiao, L., Yuan, M., Dong, T., Zhang, B., Du, C., 2019, How does urban population density decline over time? An exponential model for Chinese cities with international

comparisons, *Landscape and Urban Planning* **183:**59-67. https://doi.org/10.1016/j.landurbplan.2018.11.005

Xu, G., Zhou, Z., Jiao, L., Zhao, R., 2020, Compact Urban Form and Expansion Pattern Slow Down the Decline in Urban Densities: A Global Perspective, *Land Use Policy* **94:**104563. https://doi.org/10.1016/j.landusepol.2020.104563

Xu, H., Shi, T., Wang, M., Fang, C., Lin, Z., 2018a, Predicting effect of forthcoming population growth–induced impervious surface increase on regional thermal environment: Xiong'an New Area, North China, *Building and Environment* **136:**98-106. https://doi.org/10.1016/j.buildenv.2018.03.035

Xu, L., Huang, Q., Ding, D., Mei, M., Qin, H., 2018b, Modelling urban expansion guided by land ecological suitability: A case study of Changzhou City, China, *Habitat International* **75:**12-24. https://doi.org/10.1016/j.habitatint.2018.04.002

Xue, F., Gou, Z., Lau, S., 2017, The green open space development model and associated use behaviors in dense urban settings: Lessons from Hong Kong and Singapore, *URBAN DESIGN International* **22**(4)**:**287-302. https://doi.org/10.1057/s41289-017-0049-5

Yan, H., Liu, F., Liu, J., Xiao, X., Qin, Y., 2017, Status of land use intensity in China and its impacts on land carrying capacity, *Journal of Geographical Sciences* **27**(4)**:**387-402. https://doi.org/10.1007/s11442-017-1383-7

Yang, C., Zhao, S., 2022, Urban vertical profiles of three most urbanized Chinese cities and the spatial coupling with horizontal urban expansion, *Land Use Policy* **113**. https://doi.org/10.1016/j.landusepol.2021.105919

You, H., Yang, X., 2017, Urban expansion in 30 megacities of China: categorizing the driving force profiles to inform the urbanization policy, *Land Use Policy* **68:**531-551. https://doi.org/10.1016/j.landusepol.2017.06.020

Yu, H., Verburg, P. H., Liu, L., Eitelberg, D. A., 2016, Spatial Analysis of Cultural Heritage Landscapes in Rural China: Land Use Change and Its Risks for Conservation, *Environmental management* **57**(6)**:**1304-1318. https://doi.org/10.1007/s00267-016-0683-5

Yu, Q., Hu, Q., van Vliet, J., Verburg, P. H., Wu, W., 2018, GlobeLand30 shows little cropland area loss but greater fragmentation in China, *International Journal of Applied Earth Observation and Geoinformation* **66:**37-45. https://doi.org/10.1016/j.jag.2017.11.002

Yuan, F., Bauer, M. E., 2007, Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery, *Remote Sensing of Environment* **106**(3)**:**375-386. https://doi.org/10.1016/j.rse.2006.09.003

Yue, W., Liu, Y., Fan, P., 2013, Measuring urban sprawl and its drivers in large Chinese cities: The case of Hangzhou, *Land Use Policy* **31:**358-370. https://doi.org/10.1016/j.landusepol.2012.07.018

Yue, W., Qiu, S., Xu, H., Xu, L., Zhang, L., 2019, Polycentric urban development and urban thermal environment: A case of Hangzhou, China, *Landscape and Urban Planning* **189:**58-70. https://doi.org/10.1016/j.landurbplan.2019.04.008

Zabel, F., Delzeit, R., Schneider, J. M., Seppelt, R., Mauser, W., Vaclavik, T., 2019, Global impacts of future cropland expansion and intensification on agricultural markets and biodiversity, *Nature Communications* **10**(1)**:**2844. https://doi.org/10.1038/s41467-019-10775-z

Zambon, I., Serra, P., Salvati, L., 2019, The (Evolving) urban footprint under sequential building cycles and changing socio-demographic contexts, *Environmental Impact Assessment Review* **75:**27-36. https://doi.org/10.1016/j.eiar.2018.11.003

Zhang, W., Li, W., Zhang, C., Hanink, D. M., Liu, Y., Zhai, R., 2018, Analyzing horizontal and vertical urban expansions in three East Asian megacities with the SS-coMCRF model, *Landscape and Urban Planning* **177:**114-127. https://doi.org/10.1016/j.landurbplan.2018.04.010

Zhang, X., Liu, L., Zhao, T., Gao, Y., Chen, X., Mi, J., 2022, GISD30: global 30 m impervious-surface dynamic dataset from 1985 to 2020 using time-series Landsat imagery on the Google Earth Engine platform, *Earth System Science Data* **14**(4)**:**1831-1856. https://doi.org/10.5194/essd-14-1831-2022

Zhao, W., Duan, S.-B., Li, A., Yin, G., 2019, A practical method for reducing terrain effect on land surface temperature using random forest regression, *Remote Sensing of Environment* **221:**635-649. https://doi.org/10.1016/j.rse.2018.12.008

Zhao, Y., Ovando-Montejo, G. A., Frazier, A. E., Mathews, A. J., Flynn, K. C., Ellis, E. A., 2017, Estimating work and home population using lidar-derived building volumes, *International Journal of Remote Sensing* **38**(4)**:**1180-1196. https://doi.org/10.1080/01431161.2017.1280634

Zhong, T., Qian, Z., Huang, X., Zhao, Y., Zhou, Y., Zhao, Z., 2018, Impact of the top-down quota-oriented farmland preservation planning on the change of urban land-use intensity in China, *Habitat International*. https://doi.org/10.1016/j.habitatint.2017.12.013

Zhou, B.-B., Aggarwal, R., Wu, J., Lv, L., 2021, Urbanization-associated farmland loss: A macro-micro comparative study in China, *Land Use Policy* **101**. https://doi.org/10.1016/j.landusepol.2020.105228

Zhou, B., Rybski, D., Kropp, J. P., 2017, The role of city size and urban form in the surface urban heat island, *Scientific Reports* **7**(1)**:**1-9. https://doi.org/10.1038/s41598-017-04242-2

Zhou, C., Wang, Z., Chen, Q., Jiang, Y., Pei, J., 2014, Design optimization and field demonstration of natural ventilation for high-rise residential buildings, *Energy and Buildings* **82:**457-465. https://doi.org/10.1016/j.enbuild.2014.06.036

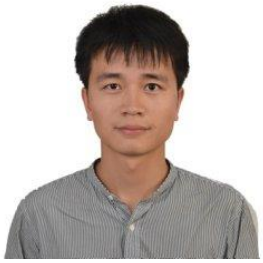
Zhou, Y., Smith, S. J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., Asrar, G. R., Zhang, X., He, C., Elvidge, C. D., 2015, A global map of urban extent from nightlights, *Environmental Research Letters* **10**(5). https://doi.org/10.1088/1748-9326/10/5/054011

Zhu, X. X., Bamler, R., 2010, Very High Resolution Spaceborne SAR Tomography in Urban Environment, *IEEE Transactions on Geoscience and Remote Sensing* **48**(12)**:**4296-4308. https://doi.org/10.1109/tgrs.2010.2050487

Zhu, X. X., Qiu, C., Hu, J., Shi, Y., Wang, Y., Schmitt, M., Taubenböck, H., 2022, The urban morphology on our planet – Global perspectives from space, *Remote Sensing of Environment* **269**. https://doi.org/10.1016/j.rse.2021.112794

Zhu, Z., Zhou, Y., Seto, K. C., Stokes, E. C., Deng, C., Pickett, S. T. A., Taubenböck, H., 2019, Understanding an urbanizing planet: Strategic directions for remote sensing, *Remote Sensing of Environment* **228:**164-182. https://doi.org/10.1016/j.rse.2019.04.020

About the author



Mengmeng Li was born in June 1992 at a remote village of Henan province, where he started receiving elementary education in 1998. At the age of 12 when Mengmeng just finished the primary education, he had to cycle a 12-km slimy road to middle school every two weeks. In the beginning, he came close to dropping out because of homesickness. After finishing the middle-school education in 2010, Mengmeng went to Henan Polytechnic University, where he spent four years pursing his bachelor degree in geography, and developed his fascination in spatial analysis and remote sensing. Right after the graduation in 2014, Mengmeng started his master education free of entrance examination at Beijing Forestry University. For his master dissertation, Mengmeng assessed vegetation degradation using satellite time-series in Xilingol, Inner Mongolia, and quantified soil loss due to wind erosion therein.

In 2017, Mengmeng was awarded a personal grant that enabled the four-year PhD research at the Institute for Environmental Studies (IVM), VU University Amsterdam. During the stay at IVM, he attended a series of international conferences (including EARSeL, CONCUR, and GLP working group meeting) to present his research outputs. Prior to the public defense of this thesis, Mengmeng has been acting as a short-term research assistant since January 2022 at Zhejiang University.

After this public defense, Mengmeng will be joining the land-use systems group of Swiss Federal Research Institute WSL in April 2023, where he will continue exploring the nuanced process of human settlement change with digital plan data using a scenario-based simulation approach.

mengbjfu@126.com

https://twitter.com/mengbjfu

https://scholar.google.com/citations?user=TwTgEzwAAAAJ

List of publications

Publications on which this thesis is based:

Li, M.*, van Vliet, J., Ke, X., & Verburg, P. H. (2019). Mapping settlement systems in China and their change trajectories between 1990 and 2010. *Habitat International*, 94, 102069. https://doi.org/10.1016/j.habitatint.2019.102069

Li, M.*, Koks, E., Taubenböck, H., & van Vliet, J. (2020). Continental-scale mapping and analysis of 3D building structure. *Remote Sensing of Environment*, 245, 111859. https://doi.org/10.1016/j.rse.2020.111859

Li, M., Wang, Y, Roiser, J.F., Verburg, P. H., & van Vliet, J.* (2022). Global maps of 3D built-up patterns for urban morphological analysis. *International Journal of Applied Earth Observation and Geoinformation*, 114, 102048. https://doi.org/10.1016/j.jag.2022.103048.

Li, M.*, Verburg, P.H., & van Vliet, J. (2022). Global trends and local variations in land take per person. *Landscape and Urban Planning*, 218, 104308. https://doi.org/10.1016/j.landurbplan.2021.104308

Other publications:

van Vliet, J.*, Birch-Thomsen, T., Gallardo, M., Hemerijckx, L. M., Hersperger, A. M., Li, M., Tumwesigye, S., Twongyirwe, R., & Van Rompaey, A. (2020). Bridging the rural-urban dichotomy in land use science. *Journal of Land Use Science*, 15(5), 585-591. https://doi.org/10.1080/1747423X.2020.1829120

Wei, J., Yue, W.*, Li, M., & Gao, J. (2022). Mapping human perception of urban landscape from street-view images: A deep-learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102886. https://doi.org/10.1016/j.jag.2022.102886

Guo, A., Yue, W.*, Yang, J., He, T., Zhang, M., & Li, M. (2022). Divergent impact of urban 2D/3D morphology on thermal environment along urban gradients. *Urban Climate*. 45, 101278. https://doi.org/10.1016/j.uclim.2022.101278

Yue, W., Feng, B., Zhou, Q., Xu, R., & Li, M.* (2023). An assessment of the Ecological Conservation Redline: Unlocking priority areas for conservation. Journal of Environmental Planning and Management. https://doi.org/10.1080/09640568.2022.2145939

Acknowledgements

When I was still a kid, my dad went to visit an esteemed fortune-teller, who foretold that China would be too small to host my growth. In that time, this was nothing but a wild fantasy for a rural boy growing up in China. My father himself wasn't a faithful religious believer, but he followed the dogma and didn't tell me this prophecy until it came true in 2017. Thanks to many, I was so lucky to be able to finish my PhD at an institute that I couldn't have imagined before.

First and foremost, I would like to express my deepest appreciation to my supervisors: Jasper van Vliet and Peter Verburg. Jasper, I don't know if you have any relations with that fortune-teller, but I am really grateful for your life-changing email (you know which one I refer to), and your professional and personal advice has helped much along the way. We are both interested in everything urban, but we also talked about carbon-rich peatland during our weekly meetings. In my deep memory, all beers taste alike, but I really love the beer you recommended at the Basket. Whether you admit it or not, I also learnt from you that some beers (that some Dutch people are proud of) are just "yellow water". Parallel appreciation goes to my other promotor – Peter, you showed me how to critically assess research by comparing it to satellite observations on your mobile phone. During our first meeting, I was so culturally shocked by you - eating a fresh carrot from somewhere. You were also an excellent tour guide and home-party host. Your knowledge and experience will encourage me in my research and daily life.

I was so fortunate to spend four years at Vrije Universiteit Amsterdam with a fantastic group of people. Thank you, Marjolijn and Cecilia, for showing me around the first day I spent at IVM, and the endless supports thereafter. I also enjoyed sharing conversations, lunches, coffees, beers and journal clubs with Alex, Bep, Bohan, Brian, Camille, Cecilia, Christian, Claudia, Floris, Franzi, Harun, Job, Jonas, Katharina, Kina, Lanping, Michelle, Niels, Nynke, Perrine, Pierre, Qiangyi, Rebecca, Thales, Tim, Vita, Yue, Ziga, and everyone else that I may forget to mention.

This thesis benefited greatly from the input of my collaborators I worked with during this PhD: Elco Koks, Job Rosier, Yuan Wang, Hannes Taubenböck, and Xinli Ke. Thank you for the inspiring discussions, timely feedback, and technical support.

During the COVID pandemic, my neighbours at Laan van Spartaan were very supportive. I enjoyed our small-group parties and daily conversations, especially with my floormates Nana and Chenhui.

After four years in Amsterdam, I was lucky to continue my research in the group of prof. Wenze Yue at Zhejiang University. Prof. Yue's invaluable advice and global vision will keep shaping my academic career in a long run. I would also like to thank Andong, Bi'ou, Jingxian, Qiushi, Ronghua, Youpeng, and everyone else for a cherished time spent together in the group, and in social settings.

I would like to extend my sincere gratitude to my committee members: prof. Philip Ward, dr. Eric Koomen, prof. Yifang Ban, prof. Karin Pfeffer, and prof. Chunyang He. I feel deeply privileged to have such an eminent group of experts to critically read my dissertation and challenge me with professional questions.

Finally, special thanks go to all my family for their encouragement and support all through my studies. As a Chinese old saying goes, "*a man should establish himself in the society at the age of thirty*" (三十而立). My family is always there whenever I feel concerned myself about this.