

## National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series

David Frantz<sup>a,\*</sup>, Franz Schug<sup>a,b</sup>, Akpona Okujeni<sup>a</sup>, Claudio Navacchi<sup>c</sup>, Wolfgang Wagner<sup>c</sup>, Sebastian van der Linden<sup>d</sup>, Patrick Hostert<sup>a,b</sup>

<sup>a</sup> Earth Observation Lab, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

<sup>b</sup> Integrated Research Institute on Transformations of Human Environment Systems (IRI THESys), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

<sup>c</sup> Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8/E120, 1040 Vienna, Austria

<sup>d</sup> Institute of Geography and Geology, University of Greifswald, Friedrich-Ludwig-Jahn-Str. 16, 17489 Greifswald, Germany

### ARTICLE INFO

#### Keywords:

Germany  
Machine learning  
Copernicus  
SAR  
Optical  
Multi-sensor  
Data synergy  
Urbanization  
Quantitative remote sensing

### ABSTRACT

Urban areas and their vertical characteristics have a manifold and far-reaching impact on our environment. However, openly accessible information at high spatial resolution is still missing at large for complete countries or regions. In this study, we combined Sentinel-1A/B and Sentinel-2A/B time series to map building heights for entire Germany on a 10 m grid resolving built-up structures in rural and urban contexts. We utilized information from the spectral/polarization, temporal and spatial dimensions by combining band-wise temporal aggregation statistics with morphological metrics. We trained machine learning regression models with highly accurate building height information from several 3D building models. The novelty of this method lies in the very fine resolution yet large spatial extent to which it can be applied, as well as in the use of building shadows in optical imagery. Results indicate that both radar-only and optical-only models can be used to predict building height, but the synergistic combination of both data sources leads to superior results. When testing the model against independent datasets, very consistent performance was achieved (frequency-weighted RMSE of 2.9 m to 3.5 m), which suggests that the prediction of the most frequently occurring buildings was robust. The average building height varies considerably across Germany with lower buildings in Eastern and South-Eastern Germany and taller ones along the highly urbanized areas in Western Germany. We emphasize the straightforward applicability of this approach on the national scale. It mostly relies on freely available satellite imagery and open source software, which potentially permit frequent update cycles and cost-effective mapping that may be relevant for a plethora of different applications, e.g. physical analysis of structural features or mapping society's resource usage.

### 1. Introduction

Three quarters of humanity currently lives in cities and towns (Dijkstra et al., 2020), and it is estimated that this trend will continue unabatedly. While cities only cover a small portion of the Earth's land surface, their impact is far-ranging as they are, for example, accountable for up to 80% of total energy consumption and 75% of carbon emissions (United Nations, 2020). For understanding urban process regimes, knowing building heights is key (Zhu et al., 2019), and the vertical structure of settlements has been identified as a central parameter to systematize multi-dimensional urban form (Wentz et al., 2018). Many quantities scale linearly with building height. For example, building height – or related information like floor area – has

been shown to be an important indicator for estimating energy consumption (Resch et al., 2016), material stock allocation (Tanikawa et al., 2015), greenhouse gas emissions (Borck, 2016), human wellbeing and urban heat island effects (Perini and Magliocco, 2014), or the distribution of population (Alahmadi et al., 2013). The latter is also of increasing relevance as knowledge about the distribution and concentration of the population may help in better understanding the risk of spreading infectious diseases (Wu et al., 2017).

Building height can be derived with a broad variety of geoinformation or remote sensing-based approaches. Nonetheless, it is a parameter that is hard to quantify accurately and at regular intervals for larger areas such as complete countries and at a spatial resolution that approximates individual building footprints. The two-dimensional

\* Corresponding author.

E-mail address: [david.frantz@geo.hu-berlin.de](mailto:david.frantz@geo.hu-berlin.de) (D. Frantz).

<https://doi.org/10.1016/j.rse.2020.112128>

Received 30 June 2020; Received in revised form 5 October 2020; Accepted 8 October 2020

0034-4257/ © 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

building footprint from building cadasters can be extruded to three-dimensional cubes (often referred to as “Level-of-Detail 1” or LoD1), if databases include attributes like the number of floors, floor height or building function (Biljecki et al., 2017). The combination of cadastral building footprints with airborne laser scanning (ALS) represents an elaborated extension of the latter approach. This technique has become the quasi-standard for generating 3D building models at LoD2, wherein prototypical roof shapes are derived from the actual point cloud (Aringer and Roschlaub, 2014). Although open data policies are increasingly emerging in general (Nugroho Rininta et al., 2015), many datasets are still proprietary and come with considerable data purchase costs (e.g. Bayerische Vermessungsverwaltung, 2020). Open alternatives to cadastral data are mostly community-based (e.g. OpenStreetMap) and as such are regionally inconsistent (Neis et al., 2013) and spatially incomplete (Brovelli and Zamboni, 2018; Hecht et al., 2013). Alternatively, building height can be measured in very high resolution (VHR) images by (i) measuring the building's cast shadows in optical images (Wang et al., 2014), (ii) by photogrammetrically analyzing stereo pairs of optical satellite, aerial or drone images (Stal et al., 2013; Takaku et al., 2016; Unger et al., 2014), (iii) or by interferometric synthetic aperture radar (InSAR) techniques (Stilla et al., 2003). However, measuring shadows is less reliable when buildings shadows overlap each other (Biljecki et al., 2017), in which case also the SAR side-looking geometry causes adverse layover and shadow effects (Stilla et al., 2003). Photogrammetric height retrievals from aerial images are more accurate than their space-based counterparts (Sirmacek et al., 2012), which is related to an increasing error with flying altitude (Baltsavias, 1999). Although photogrammetry and ALS may be similarly accurate (Baltsavias, 1999), ALS is superior for measuring building height (Kaartinen et al., 2005), especially when paired with 3D city models where the footprints are coming from official cadasters. Notwithstanding, any acquisition technique without a global observation scenario, e.g. ALS, UAV, and airborne or spaceborne VHR imaging alike, ultimately faces challenges regarding actuality, continuity and regional consistency. The same applies for data affected by non-open data policies. For example, TerraSAR-X / TanDEM-X imagery for InSAR processing is freely available, but limited to an area of 100,000 km<sup>2</sup> and only available for scientists and on request (DLR, 2019).

Eventually, free and open, globally available satellite image archives with standardized and consistent data processing are required to allow building height estimation for large areas at fine resolution. Among the hundreds of EO satellite missions (Belward and Skøien, 2015), only few systems fulfill these requirements. The European Copernicus program (Aschbacher and Milagro-Pérez, 2012) encompasses the Sentinel-1 and Sentinel-2 constellations with a high potential for building height mapping. The Sentinel-1 constellation provides all-weather, day-and-night C-band Synthetic Aperture Radar (SAR) backscatter observations at 10 m spatial gridding (Torres et al., 2012). The primary observation mode over land surfaces is interferometric wide

(IW) swath, which covers a 250 km swath at two polarizations (VV and VH). The mission currently comprises two satellites, Sentinel-1A and Sentinel-1B, providing a same-orbit revisit frequency of 6 days. Due to lateral orbit overlaps, the actual revisit frequency increases with latitude; over Europe it is 1–3 days. Li et al. (2020b) have recently shown that Sentinel-1 backscatter is helpful for estimating 3D information in cities when mapping building height as an average elevation per raster cell, i.e. a mean value of building heights and non-built-up surfaces in-between. They exploited the fact that backscatter is in general positively correlated with building height (Koppel et al., 2017), due to specular reflectance by urban structures and a combination of single bounce, double bounce (dihedrals), and triple bounce (trihedrals) scattering mechanism (Dong et al., 1997). However, the relationship between backscatter and height varies with building orientation relative to the sensor, surface material, and roughness (Corbane et al., 2008; Kimura et al., 2005; Koppel et al., 2017; Li et al., 2016). Li et al. (2020b) therefore proposed an index from combined VV- and VH-polarized SAR data and in doing so increased the predictive strength of the statistical relationship with building height. Nevertheless, they report several urban layouts and structures where this relationship is weak or non-existent, including effects related to double bouncing, high reflectivity from metallic materials, as well as scattering related to tree canopy and building density. Therefore, it appears beneficial to add information from optical imagery that allows to complement ambiguous information in the radar signal and improve the signal's correlation with building height.

The Sentinel-2 satellite mission provides multi-spectral optical observations in 13 wavelengths at up to 10 to 60 m spatial resolution (Drusch et al., 2012). The mission comprises Sentinel-2A and Sentinel-2B, which are operated in the same orbit with a phase delay of 180°, providing a nadir revisit frequency of 5 days (Drusch et al., 2012). Due to the wide swaths of 290 km, lateral orbital overlaps can be considerable and partially increase revisit frequency to 2 to 3 days. Similar optical Earth observation missions have already proven their ability to contribute to generating 3D information, e.g. to map forest height with Landsat data (Helmer et al., 2010; Li et al., 2011; Wang et al., 2018). Li et al. (2020a) have recently shown that Landsat and Sentinel-1 (among others) are capable of predicting building height and volume at a coarse 1 km resolution, yet, their feature importance assessment indicates that their model is mostly driven by the radar data. Still, the contrast between roofs and shadows is easily visible in optical winter imagery (Fig. 1a). Taking into account the clear seasonality of this contrast (cf. summer imagery, Fig. 1b), the spatio-temporal information existing in time series of multi-spectral data may increase the value of the reflectance data.

Thus, we hypothesize that the synergistic usage of information from multi-temporal and multi-spectral acquisitions may well complement the capabilities of dual-polarized radar backscatter. In this study, we aim at creating a methodology to predict building height from free,

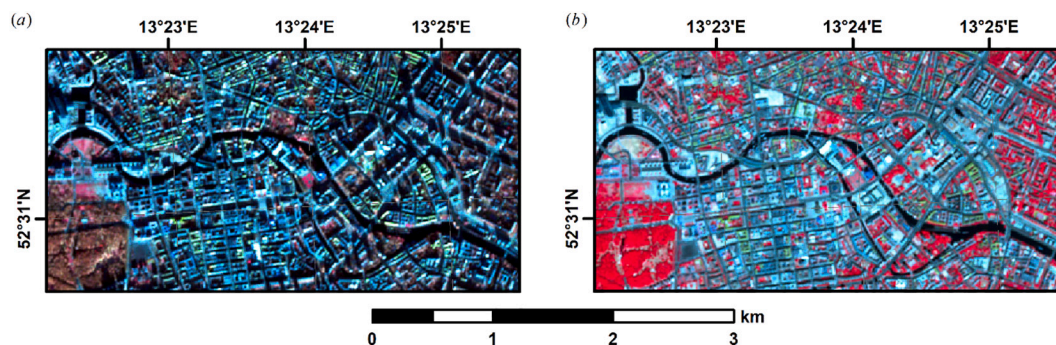


Fig. 1. Building-height induced shadows in Sentinel-2 imagery. (a): Sentinel-2A image, 05.12.2018; (b): Sentinel-2B image, 03.07.2018. The images depict the center of Berlin as false colour representation, i.e. R/G/B = near infrared/red/green. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

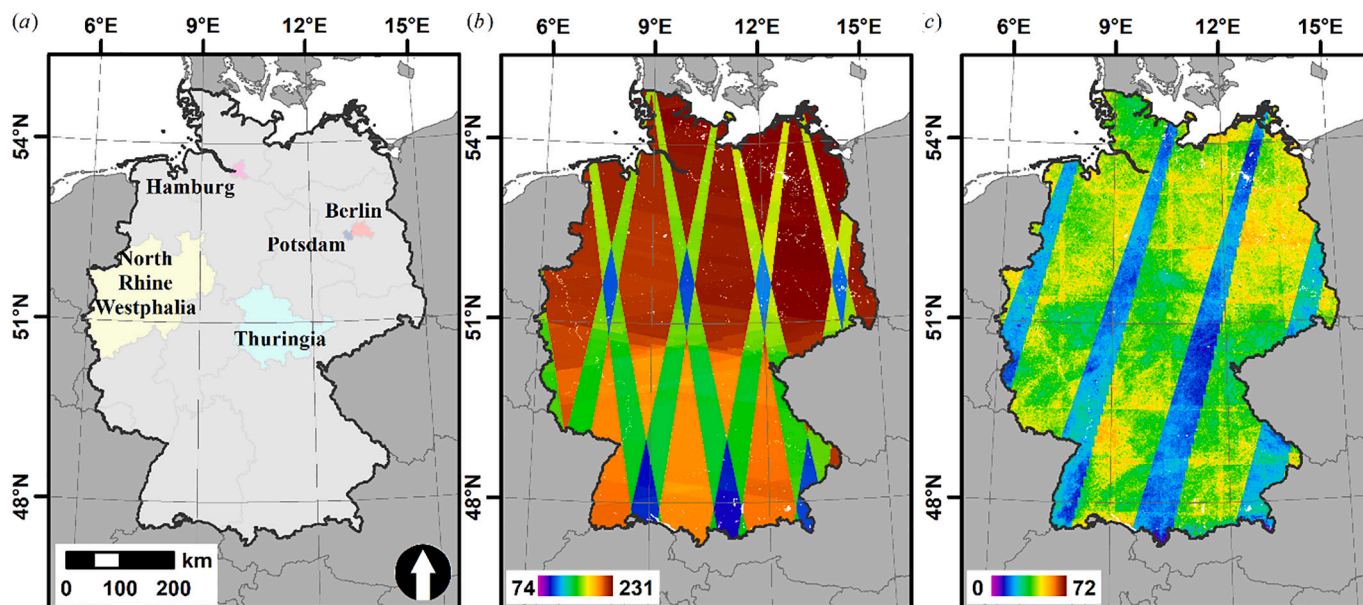


Fig. 2. Study area and data availability. (a): study area with locations of training and validation data (colored polygons); (b): number of Sentinel-1A/B acquisitions for 2017; (c): number of clear-sky, non-snow Sentinel-2A/B observations for 2018.

open, and globally available image archives using machine learning regression. We produce a building height map for entire Germany at unprecedented spatial resolution approximating building footprints. For this, we use dual-polarized Sentinel-1 and multi-spectral Sentinel-2 time series, and rely on highly accurate training and validation data derived from 3D building models. We particularly address the following research questions:

- Are synergistic models based on Sentinel-1 radar and optical Sentinel-2 data superior to using radar or optical data alone when predicting building height?
- How accurately can building height be predicted?
- Can the machine learning models be transferred to other regions without significant decrease in prediction accuracy?
- What regional patterns of building height distribution can be observed across our study area?

## 2. Study area

Our region of interest for this study was Germany, covering a total area of about 357,000 km<sup>2</sup> (Fig. 2a). Germany has about 83.1 Mio. inhabitants; population density is 232 cap km<sup>-2</sup>, varying from 69 cap km<sup>-2</sup> to 4090 cap km<sup>-2</sup> in the federal states (Statistisches Bundesamt (Destatis), 2020). About 9.6% of the area is covered by built-up infrastructure (Schug et al., 2020). The majority of buildings are residential (19,053,216 in 2018), half of them in the states of North Rhine Westphalia, Baden-Württemberg and Bavaria (Statistisches Bundesamt (Destatis), 2019b). The majority of these houses are single-family houses (82% one or two dwelling units); only 6% are tall multi-family houses (> 7 dwellings), however they make up for a third of all German dwellings (Landesamt für Statistik Niedersachsen, 2014). Information about the non-residential building stock is less reliable and conclusive; it is estimated that there are about 3 Mio. non-residential buildings (Deutsche Energie-Agentur (dena), 2016).

## 3. Data

### 3.1. Independent variables

#### 3.1.1. Sentinel-1

Sentinel-1 backscatter data for the year 2017 were accessed from the data archive of the Earth Observation Data Centre (EODC, Wagner et al., 2012). The EODC infrastructure couples an OpenStack cloud platform and the supercomputing resources of the Vienna Scientific Cluster (VSC) with a Petabyte-scale storage that contains among other satellite data the complete worldwide Sentinel-1 IW swath data record. The EODC data archive has been built up by continuously downloading Sentinel data from the Copernicus data hubs. To ensure the completeness of the data collection and short data latencies, a “Hubwatcher” is deployed to monitor and cross-check several Copernicus data hubs. In this study we used ground range detected (GRD) and high-resolution (HR) IW scenes, which contain backscatter intensities at 10 m pixel spacing in VV or VH polarization for 2017. As suggested by Koppel et al. (2017), Sentinel-1 data from both orbit directions were combined to enlarge the observation space for maximization of information content and reduction of areas in radar shadows. To allow efficient spatio-temporal analysis of the Sentinel-1 IW swath data, we pre-processed the original Level-1 data to generate a Sentinel-1 data cube based upon the Equi7 grid (<https://github.com/TUW-GEO/Equi7Grid>) developed by (Bauer-Marschallinger et al., 2014). The pre-processing was orchestrated and carried out using the SAR Geophysical Retrieval Toolbox (SGRT) developed by TU Wien, whereas the Sentinel Application Platform (SNAP) and the freely available 3-arc sec SRTM terrain model (Farr et al., 2007) were used for radiometric calibration and Range Doppler geometric terrain correction. Border noise effects, which affect a certain percentage of the original Sentinel-1 IW scenes, were eliminated with the bidirectional all-samples approach introduced by (Ali et al., 2018). The so-generated Sentinel-1 data cube holds backscatter values (sigma-naught in dB) sampled regularly at 10 m in space and irregularly in time (corresponding to the exact times of the Sentinel-1 data takes) along with local incidence angles and other relevant metadata.

Fig. 2b shows data availability; overlapping orbits and looking directions cover a large part of the study area and increase nominal revisit frequency to 1–2 days for a substantial part of the study area.

In a final post-processing step, all images were reprojected to EPSG:3035, and then split to 30 × 30 km data cubes (Frantz, 2019; Lewis et al., 2016) to match the characteristics of the optical data (see next subsection).

### 3.1.2. Sentinel-2

We downloaded all Sentinel-2 Level 1C products acquired in 2018 with cloud coverage < 70% from the ESA API Hub. The Sentinel-2 constellation was fully ramped up on 17 February 2018, thus providing full data density since this date (ESA, 2018). The study area is covered by 69 Military Grid Reference System (MGRS) tiles, and a total number of 7278 image products were available. Fig. 2c shows clear-sky data availability, where lateral orbital overlaps cover a large part of the study area and partially increase nominal revisit frequency to 2–3 days.

The data were processed to Level 2 Analysis Ready Data (ARD) through the Framework for Operational Radiometric Correction for Environmental monitoring (FORCE, Frantz, 2019) available from <https://github.com/davidfrantz/force>. Clouds and cloud shadows were identified using a modified version of the Fmask algorithm (Frantz et al., 2016a; Frantz et al., 2015; Zhu et al., 2015; Zhu and Woodcock, 2012), including a parallax-based extension to better separate clouds from bright targets such as built-up structures (Frantz et al., 2018). All images were radiometrically standardized as outlined by Frantz et al. (2016a). The atmospheric correction included radiative transfer modelling (Tanré et al., 1979), object-based Aerosol Optical Depth (AOD) estimation over dense dark vegetation (Kaufman and Sendra, 1988) and water (Royer et al., 1988), topographic correction using an enhanced C-correction as described in Buchner et al. (2020), adjacency effect correction (Bach, 1995), and Nadir BRDF-adjustment using a global set of MODIS-derived BRDF kernel parameters (Roy et al., 2016). Surface reflectance retrieval, as well as AOD estimation and cloud masking were assessed in the Atmospheric Correction and Cloud Masking Inter-comparison eXercises (ACIX/ACIX-II/CMIX, Doxani et al., 2018; ESA, 2019). The 20 m bands were improved to 10 m resolution using a data fusion approach, (Frantz et al., 2016b), wherein 10 m information from the local pixel neighborhood is used to weight the 20 m pixels using a weighted average approach. The 60 m bands were discarded. Eventually, all images were reprojected to a single coordinate system (EPSG:3035), and then split to 30 × 30 km image chips to form data cubes (Frantz, 2019; Lewis et al., 2016).

### 3.2. Dependent variable: building height

For building height training and validation, we used freely and openly available 3D building models (3DBM) in CityGML standard format. 3DBMs provided by geodetic surveys are optimal input for our purpose, as they are high quality compilations of multiple data sources. The building height information is typically based on official ALS campaigns of the local geodetic surveys, and is merged with building footprints from the official cadaster. Thus, in contrast to ALS point clouds, height information is provided for buildings only; no further disentanglement between buildings and other vertical structures was necessary (as e.g. required by the approaches of Geiß et al. (2019) and Esch et al. (2020)). As outlined in the introduction however, 3DBMs alone are insufficient to produce nation-wide building height maps, as they are only available for limited areas, where local administrations have decided on open data policies. We have acquired five 3DBMs, which cover entire cities to states (Table 1) and include nearly 15 million buildings. Those are mostly available as LoD1 or LoD2. In LoD1, each building is a block without considering its actual roof shape (accuracy of building height typically ± 5 m); LoD2 additionally contains a standard roof shape that best fits to the LiDAR point cloud (accuracy of building height typically ± 1 m). Due to the better accuracy, we only used LoD2 models in this study. The datasets were acquired between 2012 (Potsdam) and 2020 (Thuringia) with some within-dataset inconsistency (Table 1). Due to the fairly low net changes in the German

**Table 1**  
3D Building Models used for training and validation.

| Site                   | Year                                   | Accuracy in height | Buildings  | Data provider   | License   | Link  |
|------------------------|--|--------------------|------------|---|---|---|
| Berlin                 | 2014 (99.9%)<br>2015 (0.1%)            | unknown            | 540,172    | Berlin Partner für Wirtschaft und Technologie GmbH                        | Custom license: <a href="https://www.businesslocationcenter.de/berlin3d-downloadportal/documents/terms.en.html">https://www.businesslocationcenter.de/berlin3d-downloadportal/documents/terms.en.html</a> | <a href="https://www.businesslocationcenter.de/en/economic-atlas/download-portal/">https://www.businesslocationcenter.de/en/economic-atlas/download-portal/</a>   |
| Hamburg                | 2016                                   | ± 1 m              | 374,990    | Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung | Data license Germany – attribution – version 2.0  | <a href="http://suche.transparenz.hamburg.de/dataset/3d-stadtmodell-lod2-de-hamburg?forceWeb=true">http://suche.transparenz.hamburg.de/dataset/3d-stadtmodell-lod2-de-hamburg?forceWeb=true</a>   |
| Potsdam                | 2012                                   | unknown            | 44,832     | Landeshauptstadt Potsdam (LHP)  | Unspecified open data license   | <a href="https://opendata.potsdam.de/explore/dataset/3d-gebaudemodell-lod2-citygml/information">https://opendata.potsdam.de/explore/dataset/3d-gebaudemodell-lod2-citygml/information</a>   |
| North Rhine Westphalia | 2018 (39%)<br>2019 (61%)               | ± 1 m              | 11,498,734 | Bezirksregierung Köln, Geobasis NRW                                       | Data license Germany – attribution – version 2.0  | <a href="https://www.bezreg-koeln.nrw.de/brk_internet/geobasis/3d_gebaudemodelle/index.html">https://www.bezreg-koeln.nrw.de/brk_internet/geobasis/3d_gebaudemodelle/index.html</a>   |
| Thuringia              | 2018 (13%)<br>2019 (57%)<br>2020 (30%) | unknown            | 2,241,792  | Kompetenzzentrum Geodateninfrastruktur Thüringen (GDI-Th)                 | Data license Germany – attribution – version 2.0  | <a href="https://www.geoportal-th.de/de-de/Downloadbereiche/Download-Offene-Geodaten-Th%C3%BCrigen/Download-3D-Geb%C3%A4ude">https://www.geoportal-th.de/de-de/Downloadbereiche/Download-Offene-Geodaten-Th%C3%BCrigen/Download-3D-Geb%C3%A4ude</a> |

<sup>1</sup> Accuracy according to data publisher; errors may be larger for more complex roof types.

building stock (Statistisches Bundesamt (Destatis), 2019a), we consider these temporal differences negligible given the purpose of this study.

### 3.3. Other data

Similar to Li et al. (2020b) and Li et al. (2020a), we masked our final building height map to areas covered by residential and non-residential buildings only. Note that this was performed after the statistical validation. We used the European Settlement Map (ESM, Corbane and Sabo, 2019; Sabo et al., 2019), which is part of the Global Human Settlement Layer (GHSL) datasets made available by the European Commission (<https://ghsl.jrc.ec.europa.eu/data.php>). The dataset is based on Copernicus VHR imagery for the reference year 2015, and is generated at 2 m spatial resolution from Pleiades, Deimos-02, WorldView-2, WorldView-3, GeoEye-01 and Spot 6/7 images, which were acquired between 2014 and 2016. As net change of building stock is positive (+114,543 new buildings in 2018 (Statistisches Bundesamt (Destatis), 2019a)), our final map is representative for ca. 2015.

## 4. Methods

### 4.1. Theoretical background

As illustrated by Fig. 1, winter imagery in sun-synchronous optical imagery is much affected by shadows from vertical objects (e.g. buildings), whereas only the tallest objects cast a discernible shadow in summer imagery from Germany. In addition, there is also a distinct seasonal pattern in reflectance. Due to the constant overpass time of Sentinel-2, the sun elevation progressively increases towards the summer solstice, which results in decreasing building shadow lengths (Fig. 3a), which in turn results in a progressive increase of reflectance for flat areas next to buildings at 10 m spatial resolution (Fig. 3b, red line). Roof pixels that are sunlit throughout the entire year do not show this seasonal pattern (Fig. 3b, blue line).

These findings are further corroborated by the calculations shown in Fig. 4. Given a 10:30 local time overpass for 52.51°N (e.g. Berlin, Germany), the building shadow length  $d$  can be computed with

$$d = h / \tan \alpha, \quad (1)$$

where  $h$  is the building height and  $\alpha$  the sun elevation (Bivand and Lewin-Koh, 2019). For a large part of the year, i.e. between spring and fall equinox, which also coincides with highest optical data availability, shadow lengths for even tall buildings are usually below 50 m. Fig. 4b additionally demonstrates that the spatio-temporal effect found in Fig. 3b is also dependent on building height. In dependence of building height, the theoretical figure shows the number of days we are observing a shadow-free ground when being 10 m, 25 m, or 50 m away from the building.

### 4.2. Independent variables

Considering the theoretical background, we consequently developed a workflow that (1) uses radar backscatter and spectral reflectance, and (2) synergistically includes spatial context, multi-temporal, multi-spectral, and multi-polarized information. We generated an extensive feature space from the Sentinel-1 and Sentinel-2 time series, which include spectral and temporal information (subsection 1), as well as additional spatial information in a second step (subsection 2). Both steps were performed using FORCE (<https://github.com/davidfrantz/force>). For the sake of simplicity, we refer to “spectral” information throughout the remainder of the paper, although acknowledging that this is technically incorrect to some degree as e.g. different radar polarizations are measured at the same wavelength.

#### 4.2.1. Spectral-temporal features

As we identified temporal patterns in the satellite data, which are dependent on building height, we generated Spectral Temporal Metrics (STM, Mack et al., 2017; Müller et al., 2015; Potapov et al., 2009). These are statistical aggregations of all available high-quality observations within a specified time period (one year in this case), and as such provide rich information on spectral-temporal variability and data distribution (Frantz, 2019). Due to their spatial completeness and fairly high robustness against different observation densities, they are optimal features for machine learning applications, and have proven to be effective for a large variety of land cover mapping or quantitative variable estimation (e.g. Rufin et al., 2019; Schug et al., 2020). For the optical time series, only clear-sky, non-cloud, non-cloud shadow, non-snow observations were considered. For the radar time series, each observation was considered. All statistics available in FORCE (Frantz, 2019) were used: average, standard deviation, 0/10/25/50/75/90/100% quantiles, range, interquartile range, skewness, and kurtosis. The STMs were generated separately for each band of the SAR (2 bands, i.e. VV and VH polarization) and optical time series (10 land application bands covering VIS, NIR and SWIR domains). Additionally, six spectral indices derived from the optical time series were selected. The Normalized Difference Vegetation Index (NDVI, Tucker, 1979) and Tasseled Cap Greenness (Crist, 1985) were selected to account for urban vegetation, which was e.g. problematic in the Sentinel-1-only study by Li et al. (2020b); two “green” indices were chosen because NDVI alone is less suitable in Central Business Districts where shadow is dominating. The Normalized Difference Built-Up Index (NDBI, Zha et al., 2003) was used due to its sensitivity to urban areas. The Tasseled Cap Brightness (Crist, 1985) was chosen to capture brightness gradients of roofing materials. The Tasseled Cap Wetness (Crist, 1985) and modified Normalized Difference Water Index (mNDWI, Xu, 2006) were chosen to potentially account for water intermingled within the settlement, and as they emphasize shadows. The combination of 18 bands/indices and 13 statistical aggregations resulted in 234 spectral-temporal features.

#### 4.2.2. Spatial-spectral-temporal features

To capture the spatial characteristics originating from shadows cast off from nearby buildings, we computed texture metrics on top of the STMs, denoted as spatial STM (SSTM). We used morphological metrics

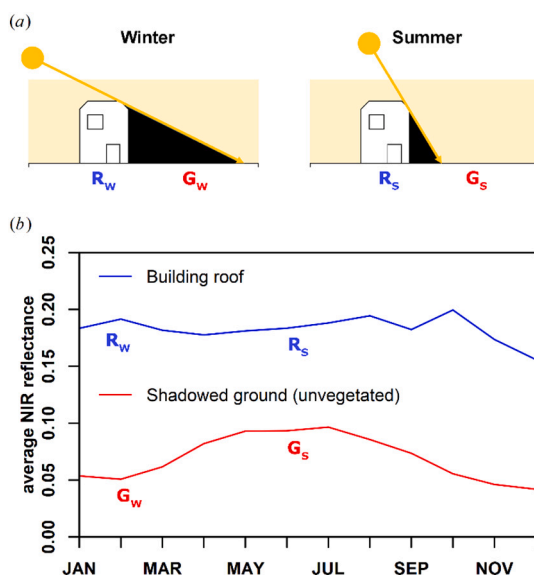


Fig. 3. Temporal effect of building shadows in Sentinel-2 imagery. (a): schematic representation of the shadowing effect for winter (left, subscript W) and summer (right, subscript S). (b): monthly near infrared average of all available clear-sky observations for a building roof pixel (blue), and the ground that is shadowed by this building (red); the two pixels were drawn from the subset shown in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

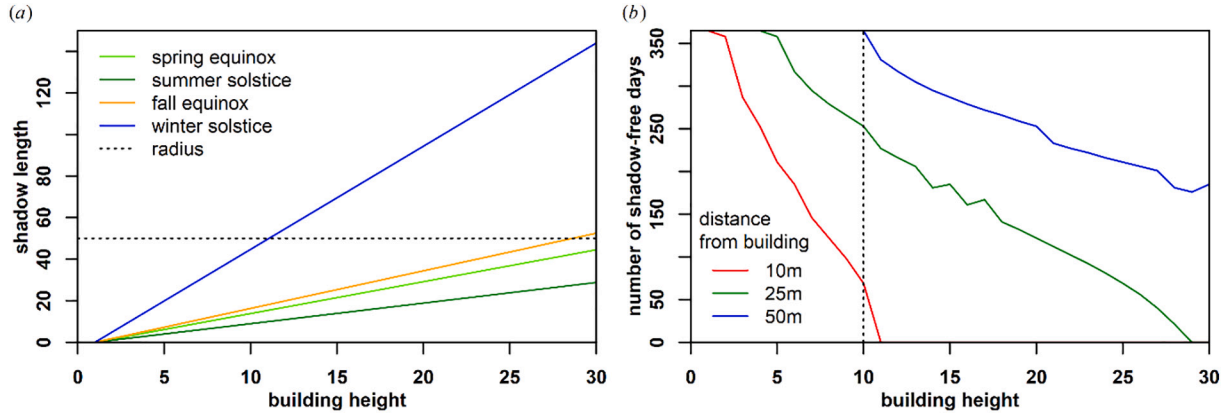


Fig. 4. Computations of building shadow lengths. (a): building shadow length in dependence on season, shown for four key astronomical dates. (b): number of shadow-free days when being 10 m, 25 m, or 50 m away from buildings with different height. The figure reads like this: for a 10 m high building (vertical line), we have 69, 253 and 365 shadow-free days a year if we step 10 m, 25 m, or 50 m away from the building.

as they are a standard way to capture spatial context in image processing (Mura et al., 2010), and can e.g. be used to extract the darkest pixel within a pixel's neighborhood. We computed erosion, dilation, opening, closing, morphological gradient, top hat and black hat using a circular structuring element with a 50 m radius. This radius was chosen as a compromise to capture most shadows (Fig. 4 shows that even taller buildings (30 m) cast a shadow smaller than 50 m for almost the entire year), while not losing too much spatial detail. This procedure resulted in a feature space with  $234 \times 7 = 1638$  features.

#### 4.3. Dependent variable: building height

##### 4.3.1. Data extraction

The 3DBMs available to us are of very high quality in terms of location, geometry, and height. However, without further processing, they are no adequate training (or validation) data for building height mapping with decimeter resolution EO data. In the CityGML format, buildings are further split into building parts, which are the geometries attributed with a building height value. However, these parts can range from complete buildings to small building extensions or garages. Sampling a garage, which often is located next to a taller building, would lead to undesired results when paired with satellite imagery, in which the garage and building are intermingled in one pixel. In addition, as outlined in the last section, a pure per-pixel estimate of building height would not work as we strive to incorporate surrounding information about the building's shadows in a 50 m radius. As such, all building parts in the local neighborhood need to be taken into account to match the prepared features, which was enabled by the spatial completeness of the official cadaster-based input datasets.

Fig. 5 displays a small portion of the 3DBM for Cologne as 2D representation, as well as the basic principle of our sample preparation. To speed up processing, we first parse the 3DBM, and store the height, footprint area and centroid of each building part in a table. Then, we overlay the 10 m pixel grid of the cubed EO data (black grid). For each pixel, e.g. the one marked with the “+” signature, it is checked whether a building part centroid (point signature) is within 5 m distance (small circle). If yes, all building parts with their centroids closer than 50 m (large circle) are included as

$$\bar{z} = \frac{\sum_p^n z_p \cdot A_p}{\sum_p^n A_p}, \quad (2)$$

wherein the area-weighted height  $\bar{z}$  is computed from the building height  $z_p$  of all nearby building parts  $p$ . The weights are the areas of the building parts  $A_p$  (as indicated by different point sizes in Fig. 5). Note that we only average the height of buildings – and not the ground in between as e.g. done by (Li et al., 2020b).

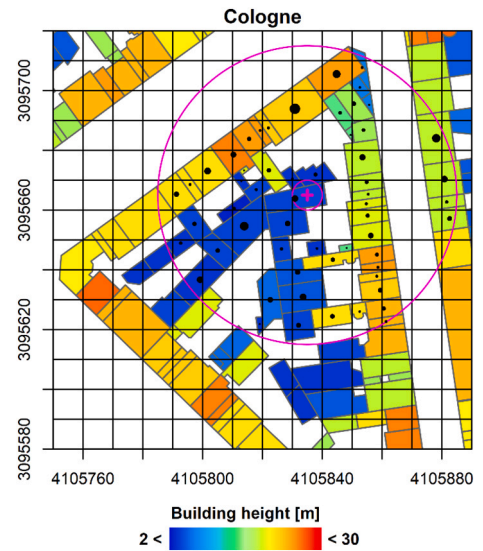


Fig. 5. 2D representation of a small part of a 3D Building model and basic principle of the sample generation. The footprint of each building part is drawn, and the building height is represented through colour. The black grid represents the pixel grid of the cubed EO images. The pink cross marks the current pixel; pink circles represent 5 m and 50 m buffers. The black points are building part centroids within the 50 m radius; the point size corresponds to the footprint area of the building parts. Coordinate system is EPSG:3035. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In the sampling process, selected building categories were removed, as we considered them structurally too specific and of rare occurrence, for example power poles of wind energy plants. Table 2 summarizes the number of potential samples per dataset; note that these numbers differ from Table 1 as the latter reports number of buildings, whereas the 3DBM processing is based on building parts.

##### 4.3.2. Sampling for training and validation

Within each 3DBM, we sampled sufficiently large, yet reasonably small datasets, which were used for training, as well as for hold-out and leave-one out cross-validation setups (see Table 2). This sample was drawn using constrained stratification of the building height: with 1 m resolution, the same number of samples (if available) was taken from all potential samples. For this, the dataset was iteratively sliced into meters, and within each slice, random samples were iteratively drawn. Due to their rare occurrence, all heights above 50 m were included in the

**Table 2**  
Number of retrieved samples per 3D Building Model.

| Site                                   | Berlin  | Hamburg | Potsdam | North Rhine Westphalia | Thuringia |
|--|---------|---------|---------|------------------------|-----------|
| Number of potential samples            | 435,610 | 445,729 | 36,592  | 7,651,841              | 960,712   |
| share of one to two floors (2–6 m)     | 15.49%  | 13.83%  | 22.75%  | 7.92%                  | 5.66%     |
| share of three to five floors (6–15 m) | 55.88%  | 74.37%  | 66.94%  | 89.41%                 | 90.77%    |
| share of five to ten floors (15–30 m)  | 26.56%  | 11.51%  | 9.71%   | 2.62%                  | 3.51%     |
| share of high-rise (30–150 m)          | 2.07%   | 0.29%   | 0.60%   | 0.05%                  | 0.05%     |
| share of sky-scrapers (> 150 m)        | 0.0012% | none    | none    | 0.0001%                | none      |
| Samples per height slice               | 1000    | 1000    | 1000    | 1000                   | 1000      |
| Number of retrieved samples            | 29,559  | 21,513  | 10,127  | 28,995                 | 20,805    |

50 m slice to avoid that each and every high-rise building was selected. If the sample candidate was located within 50 m of another confirmed sample, it was discarded. This was repeated until enough samples for each height slice were retrieved, if all potential samples were collected, or if we exceeded 100,000 iterations. In order to avoid a systematic rejection of tall or low buildings, the iterative procedure of height slicing was reversed after each iteration, i.e. in the first iteration samples were taken from the 3 m to 4 m height slice, then samples were taken from the 49 m to 50 m slice, then 4 m to 5 m, then 48 m to 49 m etc. Acting on the assumption that the 3DBMs are fully or at least almost spatially complete, we further generated background samples of zero building height in order to familiarize the machine learning models with non-building surfaces, which are not included in the 3DBMs. To achieve this, we first generated a random point dataset across the 3DBM bounding box using Poisson Disc sampling (Tulleken, 2008). We subsequently eliminated all points that were closer than 100 m to any building part centroid in the full 3DBM.

#### 4.3.3. Modelling and validation

Both model training and the subsequent building height prediction were performed using FORCE (<https://github.com/davidfrantz/force>).

**4.3.3.1. Preliminary analysis.** First, we inter-compared the performance of the radar features, the optical features, as well as a synergistic combination of both on the Berlin dataset. This dataset contains a large number of potential samples, as well as a good distribution among height classes (Table 2), which we attribute to Berlin being a diversely structured city with high-risers in the city center, very dense and compact neighborhoods in the inner city, and village-like settlements in the outskirts. We trained Support Vector Machine regression (SVR) models using 90% of the training data. The SVR hyper-parameters were tuned using grid search with 10-fold cross-validation. The remaining subsample was used for model inspection. A comparison of SVR with Random Forest regression can be found in the supplemental material section A1.

As the large feature space generated (1638 features) may result in overfitting and computational complexity related to the “curse of dimensionality” (Rust, 1997), we reduced dimensionality to 50 features using a two-step feature reduction approach as documented in supplemental material section A2.

**4.3.3.2. Final modelling.** For producing the wall-to-wall building height map for Germany, we combined the training data from all available 3DBMs. First, we evaluated the model performance with a left-out 30% sample of the training dataset. As we did not have access to training data for complete Germany, we further explored the model’s transferability and extrapolation capabilities by training models, where one training site was exclusively kept for model evaluation, whereas the remaining sites were used for training, respectively; this cross-validation approach yielded five models, i.e. each city was entirely left out once.

We generated density plots to visually compare the relationship

between building heights and the reference, complemented by a number of statistical descriptors. Ideally, this relationship is linear and oriented along the one-to-one line of the density plot. We therefore fitted a linear function to the point cloud to statistically evaluate this relationship.

As the stratified sampling scheme was designed to collect the same number of validation samples for each meter of height, whenever possible, we used ordinary least squares (OLS) regression, which is a good estimator of how well the model is able to predict buildings of different height classes. The Root Mean Square Error (RMSE) based on this sample estimates the height class uncertainty. However, in relative terms, there are few high-rise buildings (only 6% of all Germany buildings are multi-family houses with more than 7 dwellings (Landesamt für Statistik Niedersachsen, 2014)). Therefore, the OLS estimate is skewed towards higher building heights. Consequently, we additionally computed weighted least squares (WLS) regressions wherein each building height class was weighted with the frequency of its occurrence in the corresponding 3DBM. The WLS estimate is more representative of the areal accuracy, e.g. when reporting a mean building height for a given area (e.g. a city, district or state). Accordingly, the weighted RMSE is a measure of the areal height uncertainty. Complementary to this, we computed both OLS and WLS regression through the origin, as building height is a parameter with a well-defined lower boundary.

We further compared the building height distribution for all 3DBMs using histograms, and investigated potential saturation effects with increasing height by computing and visualizing mean predicted building height per building height class in the reference datasets (using the cross-validation approach).

After assessing prediction quality, we trained the model on the full training sample (i.e. 100%) before model deployment for the wall-to-wall building height map production. The final building height prediction was further pruned at the lower end, i.e. all predictions < 2 m were eliminated.

## 5. Results

### 5.1. Synergistic use of radar and optical data

We trained SVR models based on all radar, optical, and combined radar and optical features, respectively. The radar-only model achieved an uncertainty (RMSE) of 7.47 m, and frequency-weighted uncertainty of 4.32 m, respectively (Fig. 6a). The optical-only model (b) performed slightly better in all statistics. A visual inspection of the density plot reveals that the density is better concentrated on the one-to-one line and the absence of buildings (no building, i.e. reference height = 0 m) is better predicted. The combination of radar and optical features (c) outperforms both single-domain models in every statistic, e.g., both RMSE measures have been reduced by nearly one meter when compared to the radar-only model. The density plot clearly reveals less scatter, an improved density concentration on the one-to-one line, as well as an increase in regression slope, which especially in the weighted

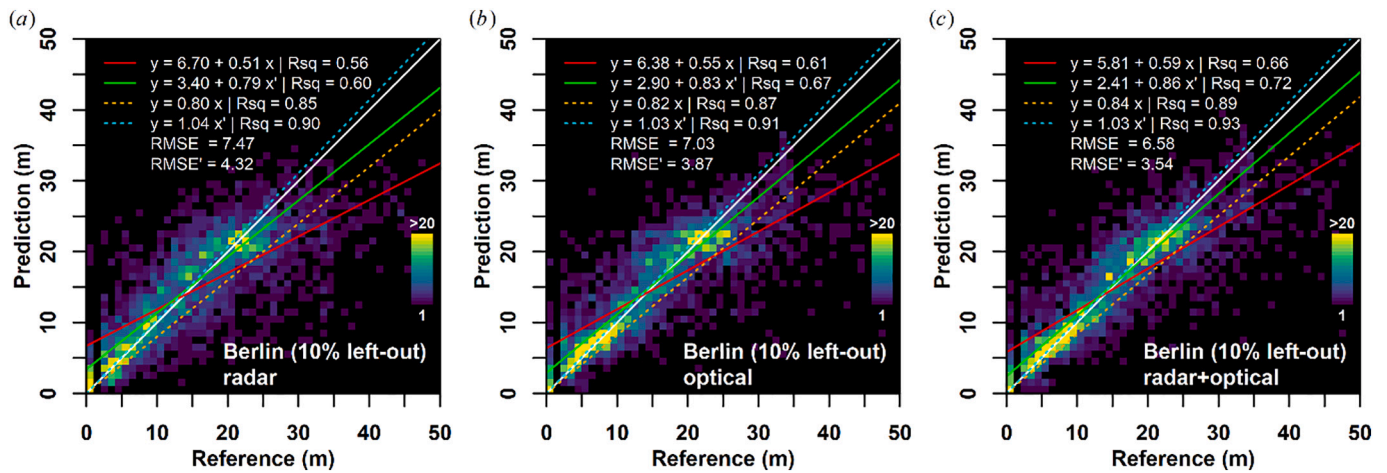


Fig. 6. Support Vector Regression model comparison using radar-only (a), optical-only (b), and both data sources combined (c). White Line = one-to-one; red line: ordinary least squares regression, orange line: weighted least squares regression through origin; green line: weighted least squares regression; cyan line: weighted least squares regression through origin; RMSE: Root Mean Squared Error, RMSE' = weighted RMSE; weights were obtained from the frequency of occurrence within the reference dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cases is very close to the one-to-one line. A comparison of SVR with Random Forest regression can be found in the supplemental material section A1. Results for the feature selection are documented in the supplemental material section A2.

### 5.2. Building height validation and model transfer

The performance of the global model (Fig. 7a) with a reduced feature set (trained and validated with a 70–30 data split with samples

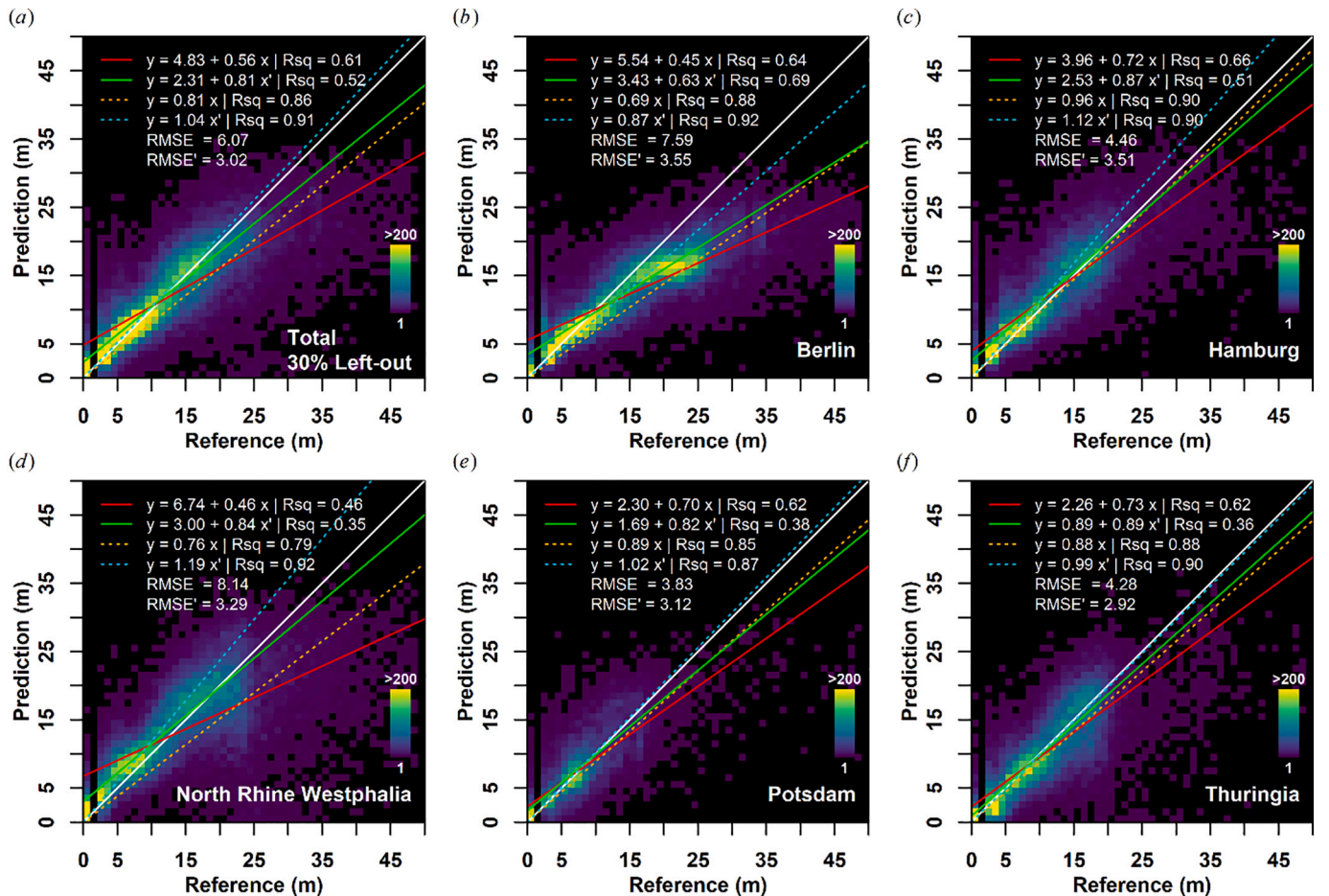


Fig. 7. Building height validation on a left-out 30% sample of the global training dataset (a), on independent validation datasets left out for testing model extrapolation and transferability (b-f). White Line = one-to-one; red line: ordinary least squares regression, orange line: ordinary least squares regression through origin; green line: weighted least squares regression; cyan line: weighted least squares regression through origin; RMSE: Root Mean Squared Error, RMSE' = weighted RMSE; weights were obtained from the frequency of occurrence within the reference dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



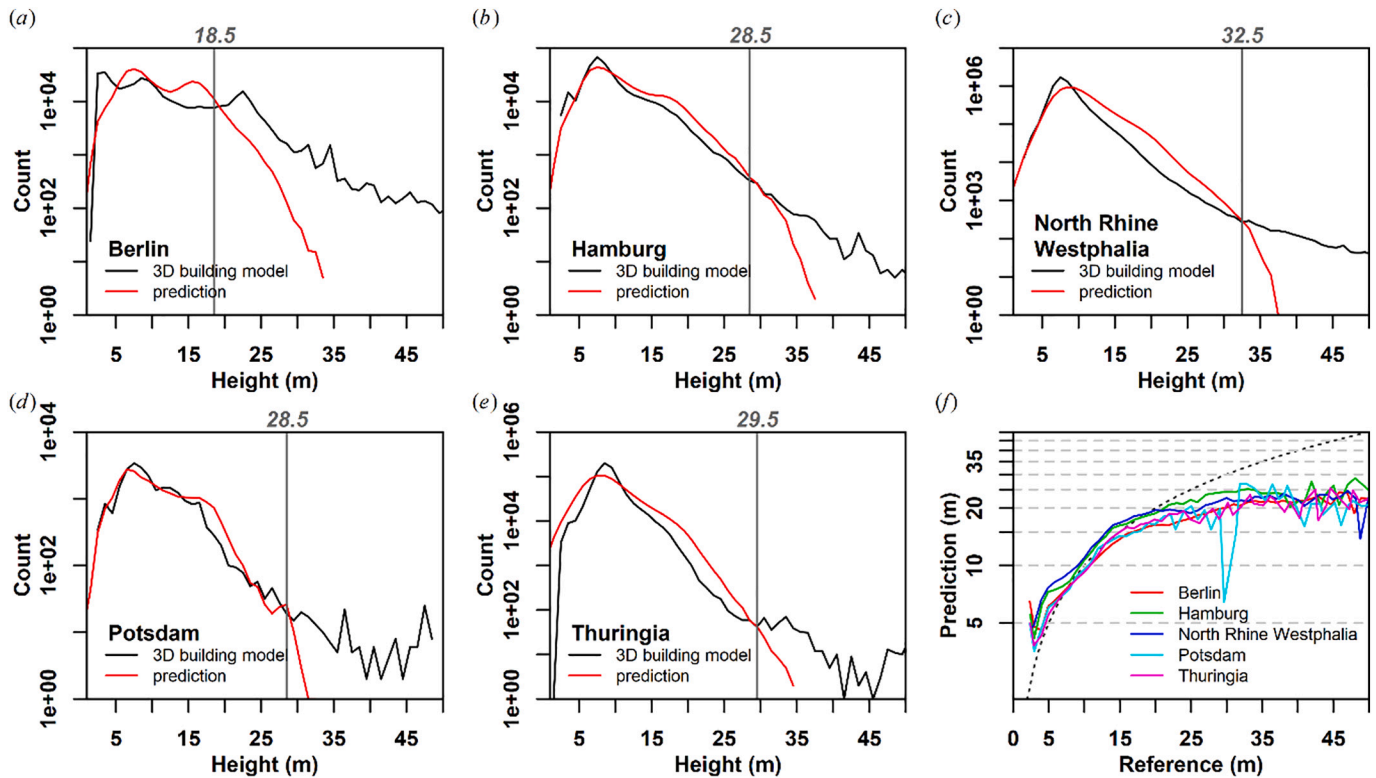


Fig. 8. Histograms (a-e) and saturation effect (f). (a-e): the histograms are not based on the validation sample, but on the complete dataset (“Number of potential samples” line in Table 2), thus, they represent the complete statistical population. The y-axes are drawn logarithmic as the building height distributions are right-skewed, e.g. more than 96% of all buildings in Thuringia are lower than 15 m (Table 2). (f): mean predicted building height per building height class of the reference datasets (in 1 m increments). The y-axis is drawn logarithmic and the dashed curve represents the one-to-one line.

from all sites) is very similar to the local model comprised of the complete feature space (Fig. 6c). A model degradation due to feature reduction was not observed. On the contrary, most performance indicators point to an increase in prediction quality.

Fig. 7b-e illustrate the performance of the model when transferred to different areas and independent 3DBMs (not used for training). The height class uncertainty (RMSE) ranges from 3.83 m to 8.14 m, and the areal height uncertainty (frequency-weighted RMSE) is between 2.92 m and 3.55 m. In terms of OLS  $R^2$ , the best correlation is found for the Hamburg dataset ( $R^2 = 0.66$ ). In other statistics, e.g. RMSE and RMSE', the best performance is achieved in Potsdam or Thuringia.

Histograms of the building height prediction and independent reference models (not used for training) are shown in Fig. 8a-e. For the city datasets, especially Hamburg and Potsdam, the height distribution is well matched. Almost all datasets are characterized by an overestimation of buildings above 10 or 15 m, which is most pronounced for the areal states of Thuringia and North Rhine Westphalia. Above a height of about 30 m however (vertical lines), the overestimation quickly turns into a substantial underestimation. This is also where the histogram becomes noisier due to decreasing sample sizes. In all datasets, a saturation effect is apparent at approximately 20 m reference height (Fig. 8f).

### 5.3. Building height mapping

The wall-to-wall building height map for Germany is shown in Fig. 9. The full dataset is openly available (Frantz et al., 2020), and can additionally be explored in this interactive map viewer: <https://ows.geo.hu-berlin.de/webviewer/building-height/>. Large cities like Berlin, Hamburg or Munich are readily visible in the overview map, as are large urban agglomerations, e.g. along the blue banana (Brunet, 1989) that covers the Ruhr District, Frankfurt Rhine-Main, Rhine-Neckar, and

Stuttgart metropolitan areas in Western Germany as indicated by the blue corridor. City centers are dominated by taller buildings, whereas city outskirts, smaller cities or the rural landscape are generally covered by smaller buildings.

Fig. 10a is characterized by taller buildings in the city center and decreasing building heights towards the city limits. The close-up, e.g. Central Berlin (a.1) reveal that our approach resolves fine-scale structures like building blocks. Buildings along larger and higher-grade streets are generally taller. Berlin-Reinickendorf (Fig. 10a.3) is a district largely dominated by single-family houses. However, it also contains one of Berlin's largest housing estates, the “Märkisches Viertel”, where high-rise buildings clearly contrast with the surrounding single-family houses. Hamburg (Fig. 10b) is characterized by a contrast between the circular city core, the port (south of the river) and a large allotment to the East (dark blue patch). Frankfurt (Fig. 10c) is the only German city, where a skyline-dominating CBD is present (red area in the city core north of the river). On the contrary, many German cities, such as Cologne (Fig. 10d) are rather dominated by historic buildings, where building height is more uniform in the city center. While the Ruhr district (Fig. 10e, Germany's largest metropolitan area) is characterized by transitions between different cities, clearly distinct nucleated settlements in the Stuttgart metropolitan area are shown in Fig. 10f, although they are formally administered as single city (Filderstadt). Dispersed settlements in rural Münsterland (Fig. 10g) are dominated by separate farmsteads with rather low buildings scattered throughout the area.

### 5.4. Regional building height distribution

The mean building height for three NUTS (Nomenclature of Territorial Units for Statistics) levels is shown in Fig. 11. NUTS-1 to 3 represent European states, government regions and districts,

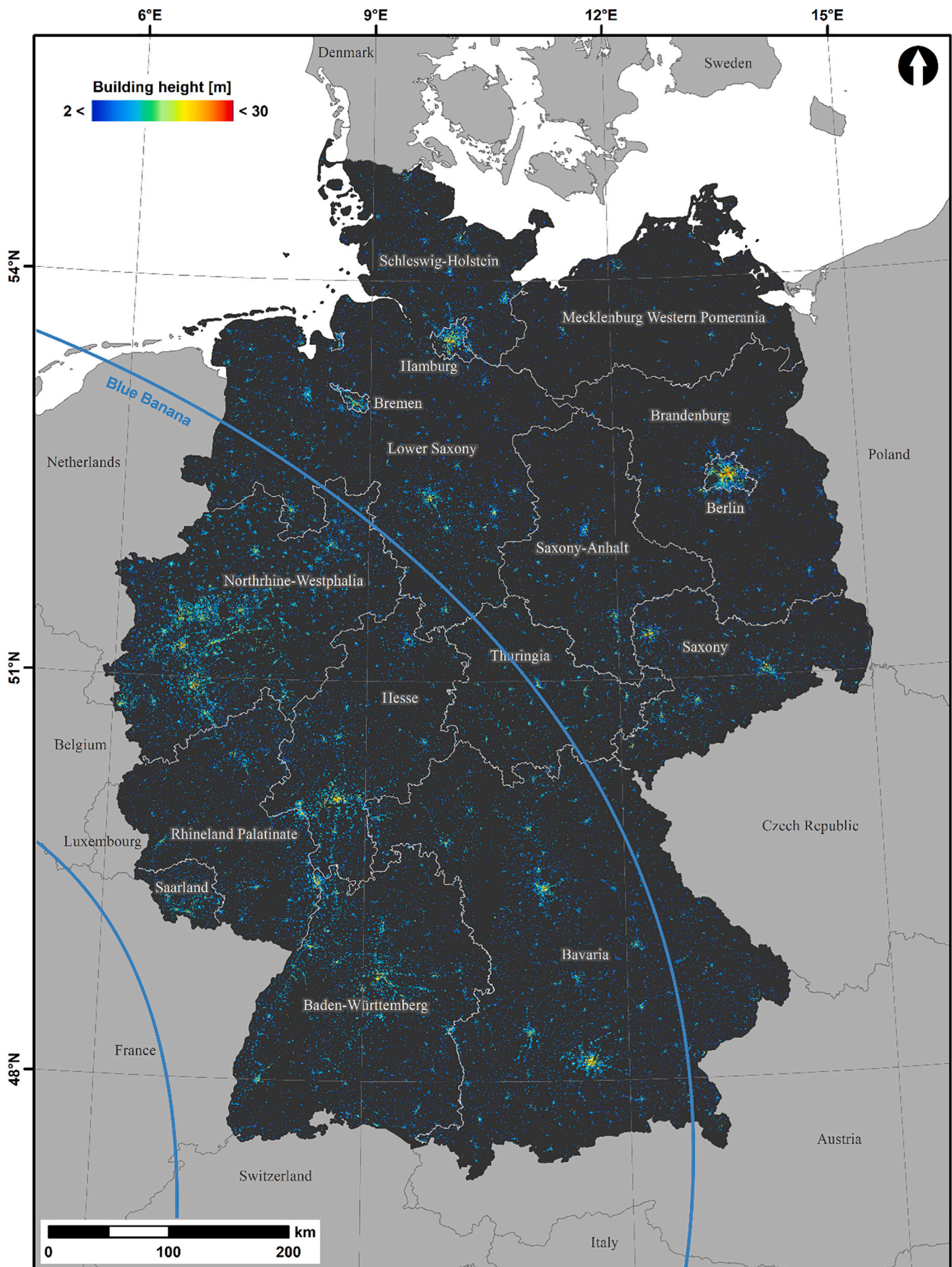
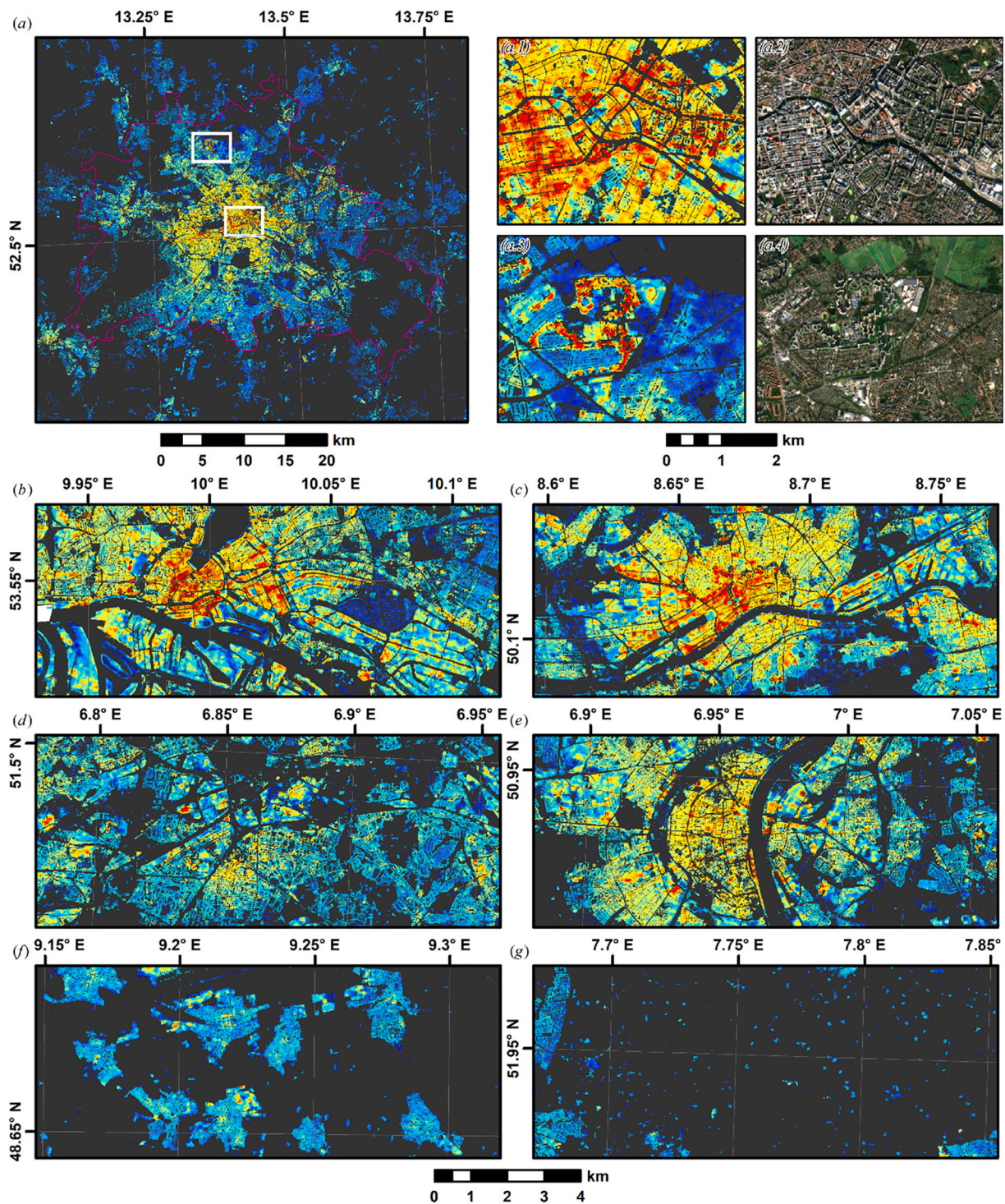


Fig. 9. Building height map for Germany. The full dataset is openly available (Frantz et al., 2020), and can additionally be explored in the interactive map viewer at <https://ows.geo.hu-berlin.de/webviewer/building-height/>.



**Fig. 10.** Building height close-ups. (a) Berlin with subsets depicting Berlin Center (a.1–2) and Berlin Reinickendorf (a.3–4) in the height map and in VHR true colour (Sentinel-2 Q50 STM); (b) Hamburg; (c) Frankfurt; (d) Ruhr District (Oberhausen); (e) Cologne; (f) nucleated settlements in Stuttgart metropolitan area; (g): dispersed settlements near Münster. The colour ramp corresponds to Fig. 9.

respectively. On the NUTS-1 level, the city-states Berlin, Hamburg and Bremen show highest mean building height. The areal states with highest and lowest buildings are North Rhine Westphalia and Brandenburg, respectively. On the NUTS-3 level, districts with large cities, e.g. Frankfurt, Munich, Stuttgart, Cologne, Düsseldorf and several cities in the Ruhr district exhibit even higher building heights than the city states. Frankfurt hosts Germany's highest buildings in its banking district with a mean building height across the entire city district of 13.7 m. Areas along the blue banana belt are also covered with higher

buildings than surrounding areas. In general, it appears that there is a trend towards lower buildings in East Germany, which may be a result of different planning policies before Germany's reunification. The more detailed NUTS levels indicate, however, that regions and districts with low building height are not exactly aligned with the former border between West and East Germany (red line). There are larger areas with low buildings in northern Lower Saxony and eastern Bavaria, too. In addition, building heights in southern Saxony and Thuringia do not differ substantially from other districts in West Germany.

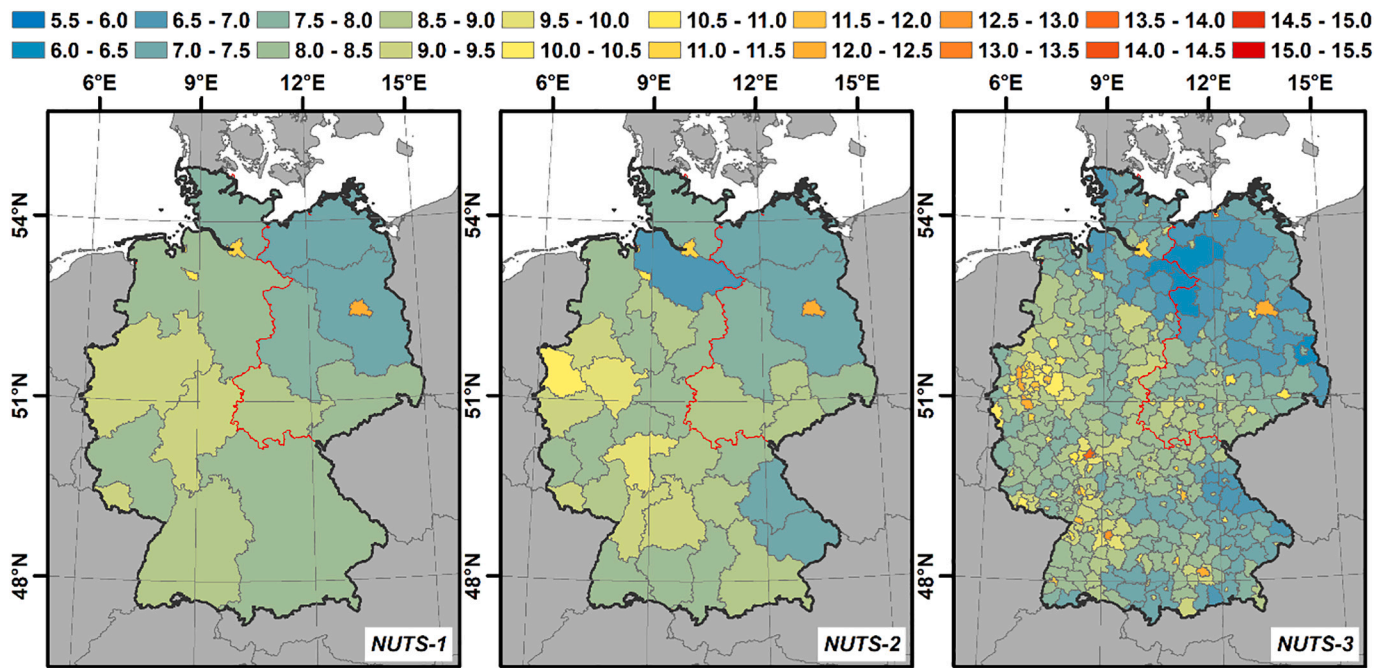


Fig. 11. Mean building height per NUTS unit; NUTS-1 to 3 represent states, government regions and districts, respectively. The red line represents the former border between West and East Germany (the former Inner-Berlin border is not drawn for graphical reasons). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## 6. Discussion

### 6.1. Synergistic use of radar and optical data

While radar is known to be sensitive to building height (Koppel et al., 2017), mapping from optical data was previously limited to VHR imagery, wherein either shadow measurement techniques or photogrammetric analysis was employed. However, as outlined in section 4.1, spatial and temporal patterns of building shadows are even observable in optical decameter resolution imagery. Our findings indicate that machine-learning based models are capable of predicting building height from both radar-only and optical-only data with similar performance, with a slightly better performance for optical data (Fig. 6). It appears that the optical-only results have less scatter in the density plot, and thus a smaller uncertainty (RMSE). In addition, the height in the absence of a building (reference height = 0 m) is better predicted in the optical case. Nevertheless, the synergistic combination of radar and optical data outperformed both single domain models.

It is known that the relationship between backscatter and building height is adversely affected by several effects. We aimed at reducing the effect of building orientation by using both polarizations (Li et al., 2020b), as well as by using data from both ascending and descending radar orbits (Koppel et al., 2017; Li et al., 2020a), which however may result in some uncertainty where viewing directions do not overlap (cf. Fig. 2b). Other geometric effects like urban layouts and structures causing double-bouncing effects, as well as overlay effects with increasing building density affect the results, too (Li et al., 2020b). Due to the side-looking view geometry, these effects cannot be fully compensated for when using radar data only, whereas such geometric effects are nearly non-existent in nadir-looking optical data. Unlike radar data, optical images are also not adversely affected by specific surface materials, e.g. specular reflectance over metal roofs that results in a loss of signal. On the contrary, the multi-spectral information likely provides explanatory power, as e.g. lower buildings in Germany are usually covered with roof tiles, whereas taller buildings often have flat roofs covered in other materials. Li et al. (2020b) further noted that tree canopies adversely affect building height predictions with radar data.

We consider that optical imagery has both advantages and disadvantages with regards to trees. On the one hand, multi-spectral and multi-temporal observations are expected to provide explanatory power as e.g. suggested by Li et al. (2020a), as e.g. low single-family buildings are often embedded in settlement forms with a high share of tree cover. On the other hand, we approximate shadowing effects through texture metrics. Thus, any strong image contrast can potentially result in the prediction of spurious heights, e.g. along avenues. To minimize this adverse effect, our approach is best paired with existing settlement extent layers, which was also done by (Li et al., 2020a; Li et al., 2020b). In our case, we masked our final map with the European Settlement Map. Despite its challenges (some residential streets classified as building and a large missing data patch of 94.7 km<sup>2</sup> in North Rhine Westphalia (51.497° N, 7.271° E), it is provided at optimal spatial resolution for our purpose, and is distributed with a free and open data license.

### 6.2. Prediction quality

#### 6.2.1. Validation of model

The quantitative validation on the left-out 30% global sample reveals a proportional relationship between predicted and reference height (Fig. 7a), which visually is close to the one-to-one line for a substantial part of the data range. The sampling scheme collected the same number of samples for each meter of height between 3 m and 50 m, wherever possible. Thus, OLS statistics between predicted and reference height are good estimators of prediction quality across different height classes. We observe a linear relationship ( $R^2 = 0.61$ ), thus our model is well capable of estimating building heights across most of the observed range of building height. However, the OLS estimates are skewed towards high-rise buildings, which in relative terms are rare in Germany. Thus WLS statistics, wherein each building height class was weighted with the frequency of its occurrence are more representative of the areal accuracy, and the corresponding RMSE is a measure of the areal height uncertainty. The WLS estimates are somewhat better than the OLS estimates: regression slope is closer to one, intercept closer to zero and RMSE is substantially lower (3.02 m vs 6.07 m), which all

suggests good prediction quality for the most frequently occurring building heights.

### 6.2.2. Model transfer (Transfer of model for national-scale mapping)

Validation on different areas with independent reference data (i.e. entirely left-out) suggests that the regional model transfer was successful and the resulting density plots (Fig. 7b-e) show a similar accuracy as the global validation (Fig. 7a). For some datasets, selected statistics are better, e.g. RMSE in Potsdam and Thuringia. Compared to all other datasets, both Potsdam and Thuringia have very few buildings taller than 20 m, which accordingly lowers RMSE, which is also reflected by smaller differences in the frequency-weighted RMSE between datasets.

The lowest transfer performance is observed for the state of North Rhine Westphalia with RMSE of 8.14 m,  $R^2$  of 0.46 and regression slope below 0.5. North Rhine Westphalia is the industrial core region of Germany; thus, many high-rising industrial buildings are located in this state, including tall factories but also many vertical structures like industrial smoke pipes, silos, or cooling towers. As we attempted to collect the same number of samples for each meter of height in the validation data, most of those tall infrastructures are included in the North Rhine Westphalia sample – although their relative proportion is comparably low (see Fig. 8). Accordingly, the OLS estimate is skewed to a larger degree as compared to the other datasets. This is readily visible in the density plots, wherein many buildings taller than 25 m are present, for which we additionally observe the most scatter (i.e. highest uncertainty). When applying the WLS-regression, the performance on the North Rhine Westphalia is much better, which suggests that the prediction of the most frequently occurring buildings is robust. The weighted RMSE is also more similar to the other datasets, which indicate that the areal height uncertainty is fairly uniform across Germany.

Our model shall be transferrable e.g. to neighboring countries that are characterized by similar building structures and material composition. From our results (especially the North Rhine Westphalia model transfer evaluation), we, however, expect that the model cannot be transferred without adaptations to regions that are too distinct from the settings in Germany, both in terms of climatic and structural properties (e.g. sky-scraper dominated cities in the U.S. or cities in deserts). When including such cases in the training sample (as e.g. by including samples from North Rhine Westphalia in the global model), we presume that the methodology might well be transferable to other parts of the globe. However, testing these hypotheses is out of scope of this paper, but merits future research.

### 6.2.3. Saturation

Despite the proportional relationship between predicted and reference building height, we generally observe a saturation effect above 20 m (Fig. 8f). Note however, that these are average numbers, higher values than 20 m do exist in our results, and Fig. 8a-e) indicates that in terms of building height frequency on the city to state level, underestimation does not occur until about 30 m. This confirms findings from Koppel et al. (2017), who report a saturation at 20 m due to the overlay effect at 41–46° incidence angle in Sentinel-1 data on Tallinn. It appears that the additional optical data cannot fully compensate for this. This may be due to physical limitations, as e.g. shadows do not have a multiplicative darkening effect when a shadow overlays on other building's shadows. However, the saturation may simply be an effect of an insufficient number of high-rise buildings in our training data. Machine learning regression usually works best if the predicted values are well inside the trained range. We hypothesize that our method would account better for high-risers by significantly increasing their occurrence in the training dataset. This however, would need to be tested in other study areas (e.g. in the U.S.) as these building types are rare in Germany (cf. Table 2). For reporting statistics for larger areas, e.g. for administrative units, we consider this effect to be negligible due to the

rare occurrence of this building category in Germany.

### 6.3. Inter-comparison with a recent 3D mapping product

Great advances have been made to map building height at spatial resolutions of 0.5 km to 1.0 km (Li et al., 2020a; Li et al., 2020b). Our maps presented in Fig. 9 and Fig. 10, as well as the online map (<http://ows.geo.hu-berlin.de/webviewer/building-height>) add to these advances by offering a further extension towards much higher spatial resolution. Our maps resolve fine scale built-up structures both in rural and urban contexts, e.g. different building blocks or even building footprints in the case of large buildings or farmsteads (see Fig. 10a.3 and g). However, it needs to be emphasized here that some care needs to be taken when interpreting this 10 m building height map. Our dependent variable is the average building height within a 50 m radius around the pixel of interest, thus, in dense and heterogeneous neighborhoods, we are still predicting a local mixture of different building heights. This is similar, although less pronounced, to previous approaches with coarser resolution. Our map can be readily aggregated to coarser spatial resolution to negate such issues if needed. Nevertheless, our approach allows for a comparative interpretation of highly detailed spatial patterns, which is not possible when directly mapping at coarse resolution.

We further inter-compared the final building height map with the 1 km map published by Li et al. (2020a), which was produced with a generalized, inter-continental model that has seen training data across the United States, China, and Europe. To homogenize the datasets as far as possible, we clipped the Li map to Germany, and aggregated our map to 1 km using average resampling (Fig. 12). Both approaches yield similar results for highly urbanized areas (see Fig. 12 top: Frankfurt Rhine-Main agglomeration). However, the predictions for rural and suburban areas are less congruent (see both Frankfurt Rhine-Main surroundings, as well as Berlin surroundings in the second row). Our map has fewer nodata values: this is a direct result of the different settlement masks and not related to the building height prediction. Still, outside of urban agglomerations, the Li map features very homogenous predictions, whereas our predictions are typically higher, and with more variability; see e.g. the differentiation between medium regional centers with ten thousands of residents (e.g. Eberswalde; green box) and small settlements.

A portion of the data is stretched along the one-to-one line (Fig. 13); however, the vast majority of pixels are located in a cluster with ca. 5 m buildings in the Li dataset, for which our map indicates a larger variability. This results in a rather low statistical relationship with  $R^2 = 0.27$ . Note that the disagreement at ca. 32 m in the Li map is a result of spurious estimates in surface mining areas, only affects a fairly small number of pixels, and is related to commission errors in both settlement masks. When stratifying building height predictions by the average built-up fraction within 1 km cells as derived from Schug et al. (2020), both predictions are fairly similar when built-up density is > 50% (Fig. 13b-c); the linear relationship between the two predictions for these pixels yields  $R^2 = 0.66$ , which is significantly higher than for the complete dataset (Fig. 13a). In the Li map, the variability of predicted building height decreases with decreasing built-up density, whereas the variability is fairly constant for all densities in our map. Also, the Li building height predictions are positively correlated with building density – as already reported by Li et al. (2020a). However, this relationship is much weaker in our case. We hypothesize that this difference is because of two main reasons:

- When predicting building height, uncertainty increases with resolution. This effect is well documented in Li et al. (2020b), and our findings quantify this uncertainty for our map (cf. Fig. 7a RMSE measures).
- There appears to be a loss of sensitivity for sparsely populated areas in the Li map. We hypothesize that this is because the “aggregate-

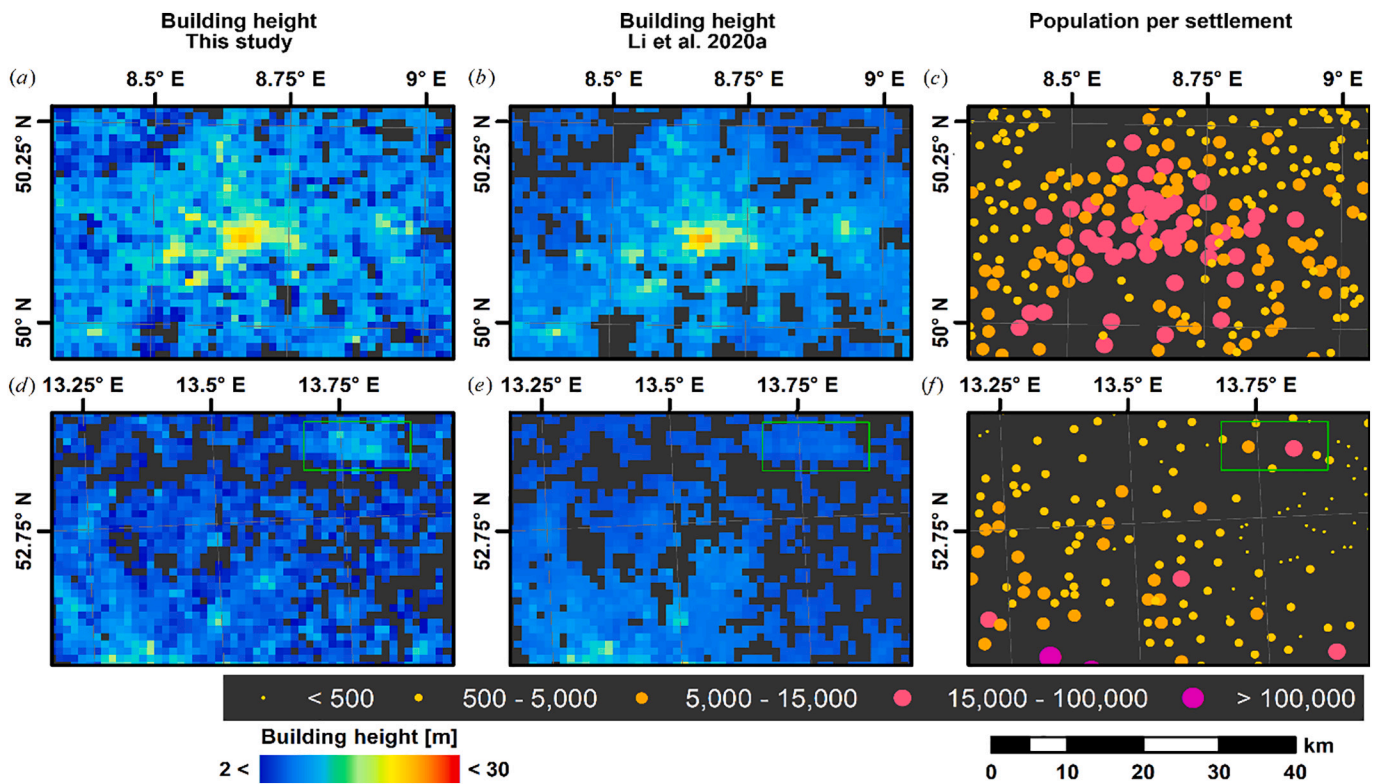


Fig. 12. Inter-comparison of building height prediction with Li et al. (2020a) for Frankfurt Rhine-Main (a-c) and Berlin surroundings (d-f). (a,d): aggregated building height prediction as presented in this study; (b,e) building height prediction of Li et al. (2020a); (c,f) population per settlement (GeoBasis-DE / BKG, 2020).

then-predict” method hits a spectral-mixture related limit when buildings become underrepresented in a pixel. This seemingly results in a dependency of the building height prediction on building density. In our “predict-then-aggregate” case, the 1 km average may be composed of a single building, for which our approach is sensitive enough. However, due to bullet one, a larger uncertainty applies as compared to the other strategy.

In summary, both 1 km products are similar for highly urbanized areas, whereas our map shows more variation in rather sparsely populated areas due to the absent dependency on building density when originally mapping building height at high resolution.

#### 6.4. Regional distribution

Across Germany, mean building height varies considerably (Fig. 11). Unsurprisingly, mean building height is higher for administrative units that are mostly comprised of large cities, e.g. Hamburg and Berlin on the state level, as well as other large cities like Frankfurt on the district level. The state level analysis implies that buildings are generally lower in Eastern Germany as compared to the West. The region and district analysis however suggests that this is unrelated to the former Inner-German border and thus not an effect of different socio-political systems. Instead, building height appears to be robustly dependent on population density (see Fig. A7 in the supplemental

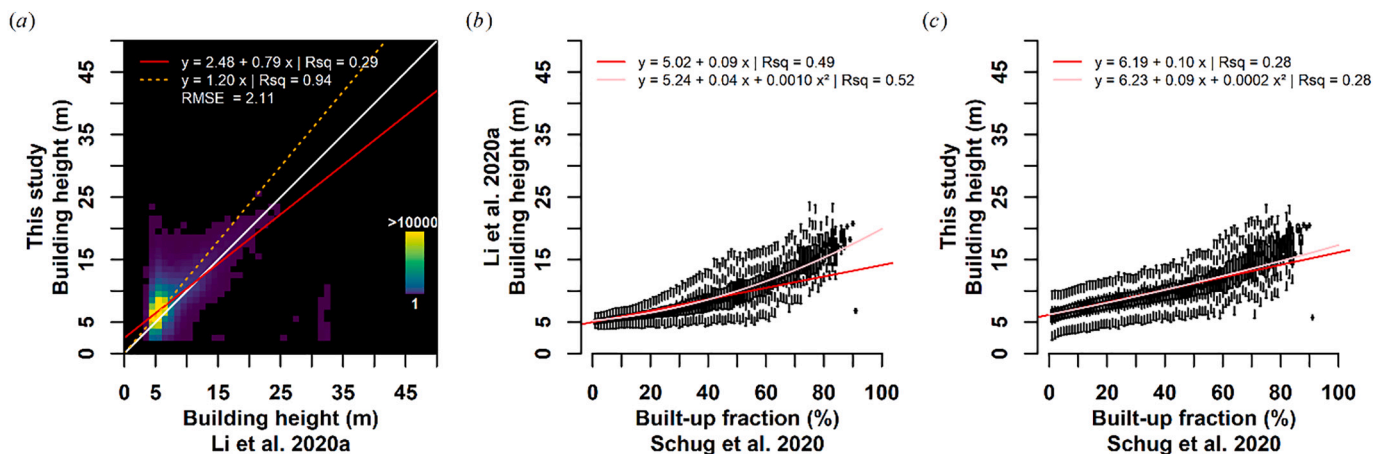


Fig. 13. Inter-comparison of building height prediction with Li et al. (2020a). a): density plot; white Line = one-to-one; red line: ordinary least squares regression, orange line: ordinary least squares regression through origin; RMSE: Root Mean Squared Error; note the non-linear colour ramp. (b-c): boxplot of the building height predictions per built-up density (Schug et al., 2020), as well as statistical relationship between building height and built-up density. Only pixels with valid predictions in both datasets were investigated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

material), which is probably due to a mutual interdependence of economic pull and building height.

## 7. Conclusion

This study presented a methodology to predict building height using a synergistic combination of dual-polarized Sentinel-1A/B and multi-spectral Sentinel-2A/B time series using a 10 m grid resolving local structures both in rural and urban contexts. Our findings confirm our hypothesis that the combined usage of optical and radar data excels the usage of one data source alone. We employed machine learning regression to predict building height using highly accurate training and validation data derived from 3D building models. We rigorously reported on prediction quality as well as on the accuracy of spatially transferring the model. We have inter-compared our map with a recent dataset and found that differences are mostly due to differences w.r.t. the data aggregation strategy. We further showed that building height varies considerably across Germany with lower buildings in less densely populated areas in Eastern and South-Eastern Germany. As demonstrated in this paper, we emphasize the straightforward applicability of this approach on the national scale as it relies on freely available satellite imagery and open source software, which potentially permit frequent update cycles and cost-effective mapping that may be relevant for a plethora of different applications. We conclude that the application of our method could be especially beneficial in countries, where information on building height is only available for smaller areas.

## Declaration of Competing Interest

None.

## Acknowledgements

We thank the European Space Agency and the European Commission for freely and openly sharing Sentinel imagery, as well as the European Settlement Mask. We thank Berlin Partner für Wirtschaft und Technologie GmbH, Freie und Hansestadt Hamburg / Landesbetrieb Geoinformation und Vermessung, Landeshauptstadt Potsdam, Bezirksregierung Köln / Geobasis NRW, and Kompetenzzentrum Geodateninfrastruktur Thüringen for freely and openly sharing 3D building models. We thank Mengmeng Li for freely and openly sharing the Continental-Scale 3D Building Structure Data. We thank Andreas Rabe and Benjamin Jakimow for discussions on feature selection. We thank Clement Atzberger for supporting this study by sharing disc space on EODC. We thank Mengmeng Li, the other three anonymous reviewers, and the editors for their very constructive feedback, which has substantially improved this paper. This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (MAT\_STOCKS, grant agreement No 741950).

## Appendix A. Supplementary information

Supplementary information to this article can be found online at <https://doi.org/10.1016/j.rse.2020.112128>.

## References

- Alahmadi, M., Atkinson, P., Martin, D., 2013. Estimating the spatial distribution of the population of Riyadh, Saudi Arabia using remotely sensed built land cover and height data. *Comput. Environ. Urban. Syst.* 41, 167–176. <https://doi.org/10.1016/j.compenvurbysys.2013.06.002>.
- Ali, I., Cao, S., Naeimi, V., Paulik, C., Wagner, W., 2018. Methods to remove the border noise from Sentinel-1 synthetic aperture radar data: implications and importance for time-series analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11, 777–786. <https://doi.org/10.1109/JSTARS.2017.2787650>.
- Aringer, K., Roschlaub, R., 2014. Bavarian 3D building model and update concept based on LiDAR, image matching and cadastre information. In: Isikdag, U. (Ed.), *Innovations in 3D Geo-Information Sciences*. Springer International Publishing, pp. 143–157.
- Aschbacher, J., Milagro-Pérez, M.P., 2012. The European earth monitoring (GMES) Programme: status and perspectives. *Remote Sens. Environ.* 120, 3–8. <https://doi.org/10.1016/j.rse.2011.08.028>.
- Bach, H., 1995. *Die Bestimmung hydrologischer und landwirtschaftlicher Oberflächenparameter aus hyperspektralen Fernerkundungsdaten*. Geobuch-Verlag, Munich, Germany.
- Baltsavias, E.P., 1999. A comparison between photogrammetry and laser scanning. *ISPRS J. Photogramm. Remote Sens.* 54, 83–94. [https://doi.org/10.1016/S0924-2716\(99\)00014-3](https://doi.org/10.1016/S0924-2716(99)00014-3).
- Bauer-Marschallinger, B., Sabel, D., Wagner, W., 2014. Optimisation of global grids for high-resolution remote sensing data. *Comput. Geosci.* 72, 84–93. <https://doi.org/10.1016/j.cageo.2014.07.005>.
- Belward, A.S., Skoien, J.O., 2015. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS J. Photogramm. Remote Sens.* 103, 115–128. <https://doi.org/10.1016/j.isprsjprs.2014.03.009>.
- Biljecki, F., Ledoux, H., Stoter, J., 2017. Generating 3D city models without elevation data. *Comput. Environ. Urban. Syst.* 64, 1–18. <https://doi.org/10.1016/j.compenvurbysys.2017.01.001>.
- Bivand, R., Lewin-Koh, N., 2019. *maptools: Tools for Handling Spatial Objects*. Available: <https://CRAN.R-project.org/package=maptools>.
- Borck, R., 2016. Will skyscrapers save the planet? Building height limits and urban greenhouse gas emissions. *Reg. Sci. Urban Econ.* 58, 13–25. <https://doi.org/10.1016/j.regsciurbeco.2016.01.004>.
- Brovelli, A.M., Zamboni, G., 2018. A new method for the assessment of spatial accuracy and completeness of openstreetmap building footprints. *ISPRS International Journal of Geo-Information* 7. <https://doi.org/10.3390/ijgi7080289>.
- Brunet, R., 1989. *Les villes européennes: rapport pour la DATAR, Délégation à l'Aménagement du Territoire et à l'Action. Régionale, La Documentation Française*.
- Buchner, J., Yin, H., Frantz, D., Kuemmerle, T., Askerov, E., Bakuradze, T., Bleyhl, B., Elizbarashvili, N., Komarova, A., Lewińska, K.E., Rizayeva, A., Sayadyan, H., Tan, B., Tepanosyan, G., Zazanashvili, N., Radeloff, V.C., 2020. Land-cover change in the Caucasus Mountains since 1987 based on the topographic correction of multi-temporal Landsat composites. *Remote Sens. Environ.* 248, 111967. <https://doi.org/10.1016/j.rse.2020.111967>.
- Corbane, C., Sabo, F., 2019. ESM R2019 - European Settlement Map from Copernicus Very High Resolution data for reference year 2015. Available: <http://data.europa.eu/89h/8bd2b792-cc33-4c11-afd1-b8dd60b44f3bhttps://doi.org/10.2905/8BD2B792-CC33-4C11-AFD1-B8DD60B44F3B>.
- Corbane, C., Faure, J.-F., Baghdadi, N., Villeneuve, N., Petit, M., 2008. Rapid urban mapping using SAR/optical imagery synergy. *Sensors* 8. <https://doi.org/10.3390/s8117125>.
- Crist, E.P., 1985. A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sens. Environ.* 17, 301–306. [https://doi.org/10.1016/0034-4257\(85\)90102-6](https://doi.org/10.1016/0034-4257(85)90102-6).
- Deutsche Energie-Agentur (dena), 2016. *Der dena-GEBÄUDEREPORT 2016: Statistiken und Analysen zur Energieeffizienz im Gebäudebestand*. Available: [https://www.dena.de/fileadmin/user\\_upload/8162\\_dena-Gebaeudereport.pdf](https://www.dena.de/fileadmin/user_upload/8162_dena-Gebaeudereport.pdf).
- Dijkstra, L., Hamilton, E., Lall, S., Wahba, S., 2020. How do we define cities, towns, and rural areas? Available: <https://blogs.worldbank.org/sustainablecities/how-do-we-define-cities-towns-and-rural-areas>.
- DLR, 2019. Proposal Submission Procedure. Available: <https://tandemx-science.dlr.de/cgi-bin/wcm.pl?page=TDM-Proposal-Submission-Procedure>.
- Dong, Y., Forster, B., Ticehurst, C., 1997. Radar backscatter analysis for urban environments. *Int. J. Remote Sens.* 18, 1351–1364. <https://doi.org/10.1080/014311697218467>.
- Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensens, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B., Vanhellemont, Q., 2018. Atmospheric correction inter-comparison exercise. *Remote Sensing* 10, 352. <https://doi.org/10.3390/rs10020352>.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution Mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>.
- ESA, 2018. Sentinel-2 images the globe every 5 days. Available: [https://earth.esa.int/web/sentinel/missions/sentinel-2/news/-/asset\\_publisher/Ac0d/content/sentinel-2-images-the-globe-every-5-days](https://earth.esa.int/web/sentinel/missions/sentinel-2/news/-/asset_publisher/Ac0d/content/sentinel-2-images-the-globe-every-5-days).
- ESA, 2019. CEOS-WGCV ACIX II - CMIX: Atmospheric Correction Inter-comparison Exercise – Cloud Masking Inter-comparison Exercise. Available: <https://earth.esa.int/web/sppa/meetings-workshops/hosted-and-co-sponsored-meetings/acix-ii-cmix>.
- Esch, T., Zeidler, J., Palacios-Lopez, D., Marconcini, M., Roth, A., Mönks, M., Leutner, B., Brzoska, E., Metz-Marconcini, A., Bachofer, F., Loekken, S., Dech, S., 2020. Towards a large-scale 3D modeling of the built environment—joint analysis of TanDEM-X, Sentinel-2 and open street map data. *Remote Sens.* 12. <https://doi.org/10.3390/rs12152391>.
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The shuttle radar topography Mission. *Rev. Geophys.* 45. <https://doi.org/10.1029/2005RG000183>.
- Frantz, D., 2019. FORCE—Landsat + Sentinel-2 analysis ready data and beyond. *Remote Sens.* 11, 1124. <https://doi.org/10.3390/rs11091124>.
- Frantz, D., Röder, A., Udelhoven, T., Schmidt, M., 2015. Enhancing the detectability of

- clouds and their shadows in multitemporal Dryland Landsat imagery: extending Fmask. *IEEE Geosci. Remote Sens. Lett.* 12, 1242–1246. <https://doi.org/10.1109/Lgrs.2015.2390673>.
- Frantz, D., Röder, A., Stellmes, M., Hill, J., 2016a. An operational radiometric Landsat preprocessing framework for large-area time series applications. *IEEE Trans. Geosci. Remote Sens.* 54, 3928–3943. <https://doi.org/10.1109/TGRS.2016.2530856>.
- Frantz, D., Stellmes, M., Röder, A., Udelhoven, T., Mader, S., Hill, J., 2016b. Improving the spatial resolution of land surface phenology by fusing medium- and coarse-resolution inputs. *IEEE Trans. Geosci. Remote Sens.* 54, 4153–4164. <https://doi.org/10.1109/TGRS.2016.2537929>.
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: separating clouds from bright surfaces based on parallax effects. *Remote Sens. Environ.* 215, 471–481. <https://doi.org/10.1016/j.rse.2018.04.046>.
- Frantz, D., Schug, F., Okujeni, A., Navacchi, C., Wagner, W., van der Linden, S., Hostert, P., 2020. Building height map for Germany (1.0). <https://doi.org/10.5281/zenodo.4066295>.
- Geiß, C., Leichte, 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, 2912–2927.
- GeoBasis-DE / BKG, 2020. Geographische Namen 1:250 000 (GN250). Available: <https://gdz.bkg.bund.de/index.php/default/open-data/geographische-namen-1-250-000-gn250.html>.
- Hecht, R., Kunze, C., Hahmann, S., 2013. Measuring completeness of building footprints in openstreetmap over space and time. *ISPRS International Journal of Geo-Information* 2. <https://doi.org/10.3390/ijgi2041066>.
- Helmer, E.H., Ruzicky, T.S., Wunderle Jr., J.M., Vogesser, S., Rufenacht, B., Kwit, C., Brandeis, T.J., Ewert, D.N., 2010. Mapping tropical dry forest height, foliage height profiles and disturbance type and age with a time series of cloud-cleared Landsat and ALI image mosaics to characterize avian habitat. *Remote Sens. Environ.* 114, 2457–2473. <https://doi.org/10.1016/j.rse.2010.05.021>.
- Kaartinen, H., Hyypä, J., Gülich, E., Vosselman, G., Hyypä, H., Matikainen, L., Hofmann, A.D., Mäder, U., Persson, Å., Söderman, U., 2005. Accuracy of 3D city models: EuroSDR comparison. *International archives of photogrammetry, remote sensing and spatial information sciences* 36, 227–232.
- Kaufman, Y.J., Sendra, C., 1988. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. *Int. J. Remote Sens.* 9, 1357–1381. <https://doi.org/10.1080/01431168808954942>.
- Kimura, H., Papathanassiou, K.P., Hajnsek, I., 2005. Polarization orientation effects in urban areas on SAR data. In: *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05*, pp. 4863–4867.
- Koppel, K., Zalite, K., Voormanskik, K., Jagdhuber, T., 2017. Sensitivity of Sentinel-1 backscatter to characteristics of buildings. *Int. J. Remote Sens.* 38, 6298–6318. <https://doi.org/10.1080/01431161.2017.1353160>.
- Landesamt für Statistik Niedersachsen, 2014. Zensus 2011 - Gebäude- und Wohnungsbestand in Deutschland. Available: [https://www.statistik.rlp.de/fileadmin/dokumente/gemeinschaftsveroeff/zen/Zensus\\_GWZ\\_2014.pdf](https://www.statistik.rlp.de/fileadmin/dokumente/gemeinschaftsveroeff/zen/Zensus_GWZ_2014.pdf).
- Lewis, A., Lymburner, L., Purss, M.B.J., Brooke, B., Evans, B., Ip, A., Dekker, A.G., Irons, J.R., Minchin, S., Mueller, N., Oliver, S., Roberts, D., Ryan, B., Thankappan, M., Woodcock, R., Wyborn, L., 2016. Rapid, high-resolution detection of environmental change over continental scales from satellite data – the earth observation data cube. *International Journal of Digital Earth* 9, 106–111. <https://doi.org/10.1080/17538947.2015.1111952>.
- Li, A., Huang, C., Sun, G., Shi, H., Toney, C., Zhu, Z., Rollins, M.G., Goward, S.N., Masek, J.G., 2011. Modeling the height of young forests regenerating from recent disturbances in Mississippi using Landsat and ICESat data. *Remote Sens. Environ.* 115, 1837–1849. <https://doi.org/10.1016/j.rse.2011.03.001>.
- Li, H., Li, Q., Wu, G., Chen, J., Liang, S., 2016. The impacts of building orientation on Polarimetric orientation angle estimation and model-based decomposition for multiloop Polarimetric SAR data in urban areas. *IEEE Trans. Geosci. Remote Sens.* 54, 5520–5532. <https://doi.org/10.1109/TGRS.2016.2567421>.
- Li, M., Koks, E., Taubenböck, H., van Vliet, J., 2020a. Continental-scale mapping and analysis of 3D building structure. *Remote Sens. Environ.* 245, 111859. <https://doi.org/10.1016/j.rse.2020.111859>.
- Li, X., Zhou, Y., Gong, P., Seto, K.C., Clinton, N., 2020b. Developing a method to estimate building height from Sentinel-1 data. *Remote Sens. Environ.* 240, 111705. <https://doi.org/10.1016/j.rse.2020.111705>.
- Mack, B., Leinenkugel, P., Kuenzer, C., Dech, S., 2017. A semi-automated approach for the generation of a new land use and land cover product for Germany based on Landsat time-series and Lucas in-situ data. *Remote Sensing Letters* 8, 244–253. <https://doi.org/10.1080/2150704X.2016.1249299>.
- Müller, H., Rufin, P., Griffiths, P., Barros Siqueira, A.J., Hostert, P., 2015. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens. Environ.* 156, 490–499. <https://doi.org/10.1016/j.rse.2014.10.014>.
- Mura, M.D., Benediktsson, J.A., Waske, B., Bruzzone, L., 2010. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Trans. Geosci. Remote Sens.* 48, 3747–3762. <https://doi.org/10.1109/TGRS.2010.2048116>.
- Neis, P., Zielstra, D., Zipf, A., 2013. Comparison of volunteered geographic information data contributions and community development for selected world regions. *Future Internet* 5, 282–300. <https://doi.org/10.3390/fi5020282>.
- Nugroho Rintina, P., Zuidervijk, A., Janssen, M., de Jong, M., 2015. A comparison of national open data policies: lessons learned. *Transforming Government: People, Process and Policy* 9, 286–308. <https://doi.org/10.1108/TG-03-2014-0008>.
- Perini, K., Magliocco, A., 2014. Effects of vegetation, urban density, building height, and atmospheric conditions on local temperatures and thermal comfort. *Urban For. Urban Green.* 13, 495–506. <https://doi.org/10.1016/j.ufug.2014.03.003>.
- Potapov, P., Hansen, M., Stehman, S.V., Pittman, K., Turubanova, S., 2009. Gross forest cover loss in temperate forests: biome-wide monitoring results using MODIS and Landsat data. *J. Appl. Remote Sens.* 3, 1–23. <https://doi.org/10.1117/1.3283904>.
- Resch, E., Böhne, R.A., Kvamsdal, T., Lohne, J., 2016. Impact of urban density and building height on energy use in cities. *Energy Procedia* 96, 800–814. <https://doi.org/10.1016/j.egypro.2016.09.142>.
- Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., Kovalsky, V., 2016. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* 176, 255–271. <https://doi.org/10.1016/j.rse.2016.01.023>.
- Royer, A., Charbonneau, L., Teillet, P.M., 1988. Interannual Landsat-MSS reflectance variation in an urbanized temperate zone. *Remote Sens. Environ.* 24, 423–446. [https://doi.org/10.1016/0034-4257\(88\)90017-X](https://doi.org/10.1016/0034-4257(88)90017-X).
- Rufin, P., Frantz, D., Ernst, S., Rabe, A., Griffiths, P., Özdoğan, M., Hostert, P., Patrick (2019). Mapping cropping practices on a National Scale Using Intra-Annual Landsat Time Series Binning. *Remote Sens.* 11, 232. DOI: <https://doi.org/10.3390/rs11030232>.
- Rust, J., 1997. Using randomization to break the curse of dimensionality. *Econometrica* 65, 487–516. <https://doi.org/10.2307/2171751>.
- Sabo, F., Corbane, C., Politis, P., Pesaresi, M., Kemper, T., 2019. Update and improvement of the European settlement map. In: *2019 Joint Urban Remote Sensing Event (JURSE)*, pp. 1–4.
- 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 Sens. Environ.* 246, 111810. <https://doi.org/10.1016/j.rse.2020.111810>.
- Sirmacek, B., Taubenböck, H., Reinartz, P., Ehlers, M., 2012. Performance evaluation for 3-D City model generation of six different DSMs from air- and Spaceborne sensors. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5, 59–70.
- Stal, C., Tack, F., De Maeyer, P., De Wulf, A., Goossens, R., 2013. Airborne photogrammetry and lidar for DSM extraction and 3D change detection over an urban area – a comparative study. *Int. J. Remote Sens.* 34, 1087–1110. <https://doi.org/10.1080/01431161.2012.717183>.
- Statistisches Bundesamt (Destatis), 2019a. Bautätigkeit und Wohnungen. Available: [https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Bauen/Publikationen/Downloads-Bautaeitaetigkeit/bautaeitaetigkeit-2050100187004.pdf?\\_blob=publicationFile](https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Bauen/Publikationen/Downloads-Bautaeitaetigkeit/bautaeitaetigkeit-2050100187004.pdf?_blob=publicationFile).
- Statistisches Bundesamt (Destatis), 2019b. Fortschreibung des Wohngebäude- und Wohnungsbestandes - Lange Reihen von 1969 bis 2018. Available: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Wohnen/Publikationen/Downloads-Wohnen/fortschreibung-wohnungsbestand-pdf-5312301.pdf>.
- Statistisches Bundesamt (Destatis). (2020). *Bevölkerungsstand*. Available: [destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/\\_inhalt.html](https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/_inhalt.html).
- Stilla, U., Soergel, U., Thoennessen, U., 2003. Potential and limits of InSAR data for building reconstruction in built-up areas. *ISPRS J. Photogramm. Remote Sens.* 58, 113–123. [https://doi.org/10.1016/S0924-2716\(03\)00021-2](https://doi.org/10.1016/S0924-2716(03)00021-2).
- Takaku, J., Tadono, T., Tsutsui, K., Ichikawa, M., 2016. VALIDATION OF "AW3D" GLOBAL DSM GENERATED FROM ALOS PRISM. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* III-4, 25–31. <https://doi.org/10.5194/isprs-annals-III-4-25-2016>.
- Tanikawa, H., Fishman, T., Okuoka, K., Sugimoto, K., 2015. The weight of society over time and space: a comprehensive account of the construction material stock of Japan, 1945–2010. *J. Ind. Ecol.* 19, 778–791. <https://doi.org/10.1111/jiec.12284>.
- Tanré, D., Herman, M., Deschamps, P.Y., de Lefte, A., 1979. Atmospheric modeling for space measurements of ground Reflectances, including bidirectional properties. *Appl. Opt.* 18, 3587–3594. <https://doi.org/10.1364/ao.18.003587>.
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 Mission. *Remote Sens. Environ.* 120, 9–24. <https://doi.org/10.1016/j.rse.2011.05.028>.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).
- Tulleken, H., 2008. Poisson disk sampling. *Dev. Mag* 21, 21–25.
- Unger, J., Reich, M., Heipke, C., 2014. UAV-based photogrammetry: monitoring of a building zone. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XI-5, 601–606. <https://doi.org/10.5194/isprsarchives-XI-5-601-2014>.
- United Nations, 2020. Sustainable Development Goal 11: Make cities inclusive, safe, resilient and sustainable. Available: <https://www.un.org/sustainabledevelopment/cities/>.
- Bayerische Vermessungsverwaltung. (2020). Gebühren- und Preisliste für Geobasisdaten der Bayerischen Vermessungsverwaltung Available: [ldbv.bayern.de/file/pdf/1269/Preisliste\\_aktuell.pdf](https://ldbv.bayern.de/file/pdf/1269/Preisliste_aktuell.pdf).
- Wagner, W., Triebnig, G., Wotawa, G., Nagler, T., Schardt, M., Hoffmann, C., Rott, H., Walli, A., 2012. Feasibility Study for an Earth Observation Data Centre for Water Resources Monitoring (EODC-Water).
- Wang, X., Yu, X., Ling, F., 2014. Building heights estimation using ZY3 data — A case study of Shanghai, China. In: *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 1749–1752).
- Wang, M., Sun, R., Xiao, Z., 2018. Estimation of Forest canopy height and aboveground



- biomass from Spaceborne LiDAR and Landsat imageries in Maryland. *Remote Sens.* 10. <https://doi.org/10.3390/rs10020344>.
- 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. *Landsc. Urban Plan.* 179, 55–62. <https://doi.org/10.1016/j.landurbplan.2018.07.007>.
- Wu, T., Perrings, C., Kinzig, A., Collins, J.P., Minter, B.A., Daszak, P., 2017. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: a review. *Ambio* 46, 18–29. <https://doi.org/10.1007/s13280-016-0809-2>.
- Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* 27, 3025–3033. <https://doi.org/10.1080/01431160600589179>.
- Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* 24, 583–594. <https://doi.org/10.1080/01431160304987>.
- Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* 118, 83–94. <https://doi.org/10.1016/j.rse.2011.10.028>.
- Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images. *Remote Sens. Environ.* 159, 269–277. <https://doi.org/10.1016/j.rse.2014.12.014>.
- 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 Sens. Environ.* 228, 164–182. <https://doi.org/10.1016/j.rse.2019.04.020>.