https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Open Access Earth System Science Data Discussions

U-Surf: A global 1km spatially continuous urban surface property dataset for kilometer-scale urban-resolving Earth system modeling

Yifan Cheng[1], Lei Zhao[1,2,3], Tirthankar "TC" Chakraborty[4], Keith Oleson[5], Matthias Demuzere[6], Xiaoping Liu[7], Yangzi Che[7], Weilin Liao[7], Yuyu Zhou[8], Xinchang "Cathy" Li[1]

[1] Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA
[2] Institute for Sustainability, Energy, and Environment (iSEE), University of Illinois Urbana-Champaign, Urbana, IL, USA
[3] National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, IL, USA
[4] Atmospheric, Climate, and Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA
[5] Climate and Global Dynamics Laboratory, NSF National Center for Atmospheric Research, Boulder, Colorado, USA
[6] B-Kode VOF, Ghent, Belgium
[7] Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China
[8] Department of Geography, The University of Hong Kong, 999077, Hong Kong, China

Correspondence to: Lei Zhao (leizhao@illinois.edu), TC Chakraborty (tirthankar.chakraborty@pnnl.gov), and Weilin Liao (liaoweilin@mail.sysu.edu.cn)

Abstract. High-resolution urban climate modeling has faced substantial challenges due to the absence of a globally consistent, spatially continuous, and accurate dataset to represent the spatial heterogeneity of urban surfaces and their biophysical properties. This deficiency has long obstructed the development of urban-resolving Earth System Models (ESMs) and ultra-high-resolution urban climate modeling, particularly at large scales. Here, we present a first-of-its-kind 1km-resolution present-day (circa-2020) global continuous urban surface parameter dataset – U-Surf. Using the urban canopy model (UCM) in the Community Earth System Model as a base model for developing dataset requirements, U-Surf leverages the latest advances in remote sensing, machine learning, and cloud computing to provide the most relevant urban surface biophysical parameters, including radiative, morphological, and thermal properties, for UCMs at the facet- and canopy-level. Our high-resolution U-Surf dataset significantly improves the representation of the urban land heterogeneity both within and across cities globally. U-Surf provides essential, high-fidelity surface biophysical constraints to urban-resolving ESMs, enables detailed city-to-city comparisons across the globe, and supports the next-generation kilometer-resolution Earth system modeling across scales. U-Surf parameters can be easily converted or adapted to various types of UCMs, such as those embedded in weather and regional climate models, as well as air quality models. The fundamental urban surface constraints provided by U-Surf are also relevant as features for machine learning models and can have other broad-scale applications for socioeconomic, public health, and urban planning contexts. We expect U-Surf to promote the research frontier on urban systems science, climate-sensitive urban design, and coupled human-Earth systems in the future. The dataset is publicly available at https://doi.org/10.5281/zenodo.11247599 (Cheng et al., 2024).

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

1 Introduction

Urban areas are the global hotspots of climate hazards (Intergovernmental Panel On Climate Change, 2023; Robinson et al., 2021; Tabari, 2020; van der Wiel and Bintanja, 2021; Zhao et al., 2021), exposure (Chen et al., 2023; Li et al., 2019; Yang et al., 2023), and vulnerability (Ajjur and Al-Ghamdi, 2021; Lobo et al., 2023), due to the uniqueness of local urban climates (Baklanov et al., 2018; Cao et al., 2016; Chakraborty et al., 2023; Li and Bou-Zeid, 2013; Zhan et al., 2023, p.201; Zhao et al., 2014, 2018), concentration of population, infrastructure, and capital assets (Gao and Bukovsky, 2023; Masson et al., 2020; Shu et al., 2023), and mixture of diverse communities and socio-ethnic groups (Islam and Winkel, 2017; Kim et al., 2021). With an additional 2.5 billion people projected to reside in urban areas by 2050 (United Nations, 2018), these climate-driven risks are expected to be exacerbated in future warmer climates (Intergovernmental Panel On Climate Change, 2014). This inevitable urbanization coupled with climate change will expose cities and their residents to greater risks across the world (Feng et al., 2021; Gao and Bukovsky, 2023; Scheuer et al., 2017; Sjöstrand, 2022), but also presents a historic and time-sensitive opportunity to mitigate and adapt to the negative climate impacts (Krayenhoff et al., 2018; Zhao, 2018; Zhao et al., 2017). To address this grand challenge, it is urgent to better understand urbanization and its complex two-way interactions with climate across spatiotemporal scales. Achieving this goal, however, requires advanced data and tools that realistically resolve urban land in models such as mesoscale weather models, Earth system models (ESMs), and Earth System Digital Twins (Li et al., 2023a), both to better understand cities and their impacts and for planning effective climate adaptation and mitigation strategies (Krayenhoff et al., 2021).

In light of the increasingly recognized importance of urban climatic implications, substantial efforts on representing urban landscapes in local-to-regional climate models have been reported in the past decade, including improved urban-scale process representations (Chen et al., 2011; Conigliaro et al., 2021; Jongen et al., 2024; Langendijk et al., 2024; Lipson et al., 2024) and surface input datasets (Ching et al., 2018; He et al., 2023; Qi et al., 2024; Sun et al., 2021). Urban representation in global-scale models, however, is significantly lagging. This is because an urban canopy model (UCM) is largely missing in most state-of-the-art ESMs or global climate models (Hertwig et al., 2021; Zhao et al., 2021; Zheng et al., 2021). This omission will become an even more critical issue in the future, as next-generation ESMs are expected to run at kilometric scales (Schär et al., 2020, 2021; Wang et al., 2022; Yuan et al., 2023), at which resolving urban areas, their unique biophysical properties, changes over time, and interactions with broader-scale systems will inevitably be required (Chakraborty and Qian, 2024; Grimmond et al., 2011; Sharma et al., 2021). One primary roadblock that has prevented the development of "urban-resolving" ESMs and accurate global urban climate modeling for decades is the lack of globally consistent estimates of urban surface properties, which are critical model inputs, especially at fine resolutions.

Currently, there is no global and spatially continuous urban dataset that can provide all relevant biophysical parameters for UCMs that can be used in state-of-the-art ESMs across scales (Masson et al., 2020). Unlike local- and regional-scale studies using mesoscale UCMs, for which the urban surface parameters usually rely on either simple look-up tables or user-supplied locally-defined physical description of the study area, common UCMs embedded in ESMs

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

need a complete, fine-resolution, globally and internally consistent, and spatially explicit urban surface parameters. These parameters are required at the facet and canopy level, and therefore, are dramatically challenging to produce at the global scale.

An urban surface data created by Jackson et al., 2010 (hereafter referred as J2010) is, to our knowledge, the only available global dataset to date that can provide the required UCM parameters for Earth system modeling in a globally consistent manner. It was developed by synthesizing population density estimates, satellite data, existing literature, building codes, and municipal documentation. This dataset and its updated version (Oleson and Feddema, 2020) serve as the default urban surface property input for the Community Earth System Model version 2 (CESM2) (Danabasoglu et al., 2020) and Energy Exascale Earth System Model (E3SM) (Golaz et al., 2022). Compiled at a time when fine-resolution geospatial data were very scarce, J2010 is coarse-grained, spatially discontinuous, and somewhat outdated (valid for circa-2000), and hence poorly constrains the spatial heterogeneity of urban properties within cities and across the world. J2010 clusters the global urban areas into 33 distinct regions of similar climates, socio-economic characteristics, and architectural practices (Figure S1), with properties defined within each region for up to four urban density classes: low density (LD), medium density (MD), high density (HD), and tall building district (TBD). The dataset then prescribes uniform surface properties to each density type within a region. These simplistic, coarse-grained, and region-based urban property constraints impede its application in resolving the true heterogeneity of cities and their interactions with background climate, especially relevant for high-resolution urban climate modeling.

Recent development of the Local Climate Zone (LCZ) typology framework (Stewart and Oke, 2012) provides another potential means to supply spatially explicit urban parameters to regional and global models. LCZ standardizes a common descriptive methodology to classify land surfaces into 10 built and 7 natural land cover types, each associated with some prescribed ranges of values for a subset of (mostly morphological) parameters. Compared to the widely used conventional land cover maps, LCZs are a step forward for representing the spatial heterogeneity of urban landscapes (Demuzere et al., 2022a). Many high-resolution regional (Demuzere et al., 2021; Huang et al., 2021; Qi et al., 2024) and global (Demuzere et al., 2022a) LCZ raster maps have been produced in recent years, greatly advancing the description of urban areas at large scales in a “universal” way. However, a critical gap that remains is how to determine the urban canopy parameters based on the LCZ raster maps. A common approach currently relies on referring to the predefined value ranges from the original LCZ typology (Demuzere et al., 2022b; Stewart and Oke, 2012; Sun et al., 2021), which essentially remains a look-up-table method. Similar to J2010, the inherent assumption of uniformity within each zone, i.e. cities located in different countries and diverse climate regimes will be assigned the same set of parameters if classified as one LCZ type, oversimplifies the complexity and heterogeneity of urban surfaces. In addition, LCZs, by their very nature, largely describe typologies of urban morphology, but other characteristics such as radiative properties and construction materials are less well defined and subject to large uncertainties (Hidalgo et al., 2019; Masson et al., 2020). More importantly, these properties can be frequently decoupled from that morphology, meaning the complete set of parameters used as model inputs are not internally consistent.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

To address this long-standing urban representation challenge at large scales and to facilitate next-generation kilometer-scale (k-scale) urban-resolving Earth system modeling, we develop a first-of-its-kind global high-resolution (1 km) urban surface dataset, namely U-Surf, to support urban climate modeling across scales. The development of U-Surf is enabled by latest advances in high-resolution remote sensing measurements from recent satellite missions, new algorithms to derive satellite-derived products, building footprint estimates from global scale image segmentation methods, and advancements in hybrid cloud supercomputing. We use the urban scheme in CESM2's land model (Community Land Model or CLM) as the base model to develop the dataset, as it is one of the very few state-of-the-art ESMs with an urban canopy representation. Nevertheless, the derived parameters in U-Surf can be easily adapted to other mesoscale weather or global climate models such as The Weather Research and Forecasting Model (WRF) and E3SM, the latter using a UCM identical to that in CLM version 4.5. The U-Surf data does not rely on any coarse-graining (clustering), but instead estimates the facet- and canyon-level surface properties in a spatially continuous manner at 1 km resolution. Therefore, the final U-Surf product provides a global, internally consistent and comprehensive set of urban surface input for UCMs, captures the fine-scale spatial heterogeneity both within and across cities, and markedly advances the potential for urban representation in weather and climate models across scales. In addition to its applications in climate modeling, U-Surf could be used directly as input features for machine learning models, and can also be leveraged for other non-climatic modeling exercises, analyses, or applications in the energy, geography, and socioeconomic fields.

This paper is organized as follows. Section 2 details the data sources and methodology employed in developing the dataset. Section 3 presents the spatial distributions of the newly created 1km-resolution dataset, highlighting selected parameters across various scales. Sections 4 and 5 discuss the broad implications of the dataset, the current limitations, and potential future work. Section 6 provides information and links on accessing the dataset in different formats and associated Google Earth Engine (GEE) web application, while section 7 provides concluding remarks.

2 Data and methods

2.1 Urban representation in CESM2

The current version of the U-Surf dataset is based on the urban parameterization scheme embedded in the CESM2 for two reasons. First, CESM2 is one of the very few state-of-the-art ESMs with a physically based UCM – the Community Land Model Urban (CLMU) (Lawrence et al., 2019; Oleson and Feddema, 2020). CLMU has been evaluated against site observations and satellite measurements across the world with consistently reasonable agreement (Demuzere et al., 2008, 2013, 2014, 2017; Fitria et al., 2019; Lin et al., 2016; Oleson et al., 2008a, b; Zhang et al., 2023a; Zhao et al., 2014, 2021) and has also demonstrated high credibility among various UCMs in the recent multi-model comparison project Urban-PLUMBER (Lipson et al., 2024). Second, the urban canopy concept that CLMU uses is widely adopted in various UCMs embedded in weather models and RCMs. Therefore, a dataset developed based on this conceptual representation can be easily extended to other UCMs or climate models.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Open Access Earth System Science Data Discussions

The urban canopy representation used in CLMU and many other UCMs is called an "urban canyon" schema where the urban landscape at a given location is conceptualized as an infinite "urban canyon" (Figure 1). This canyon hypothesis assumes a geometry of an infinitely long street bordered by two building walls with identical height. An "urban canyon" consists of five facets: building roof, impervious (e.g., roads, parking lots, sidewalks) and pervious (e.g., lawns, street trees, parks) canyon floors, and sunlit and sun-shaded walls (Oleson et al., 2008a). This conceptual representation reduces the considerable complexity of urban surfaces into a single urban canyon, and yet provides an essential base to represent key urban biogeophysical processes effectively. The UCMs using this approach therefore require sets of properties at both facet- and canopy-level to represent urban landscapes and model their interactions with the lower atmosphere in climate/weather simulations. These properties can be generally grouped into three categories: morphological (e.g., canyon height-to-street width ratio, roof fraction, average building height, and pervious canyon floor fraction), radiative (e.g., facet-level albedo and emissivity), and thermal (e.g., heat capacity and thermal conductivity) (Figure 1 and Table 1). These surface properties characterize the "urban areas" and are critical for constraining their surface energy budget, and thus the near-surface microclimate, in weather and climate models. More details about the CLMU parameterization scheme and its evolution over the years can be found in Jackson et al., (2010), Lawrence et al., (2019), Li et al., (2024), Oleson et al., (2008a, b), Oleson and Feddema, (2020).

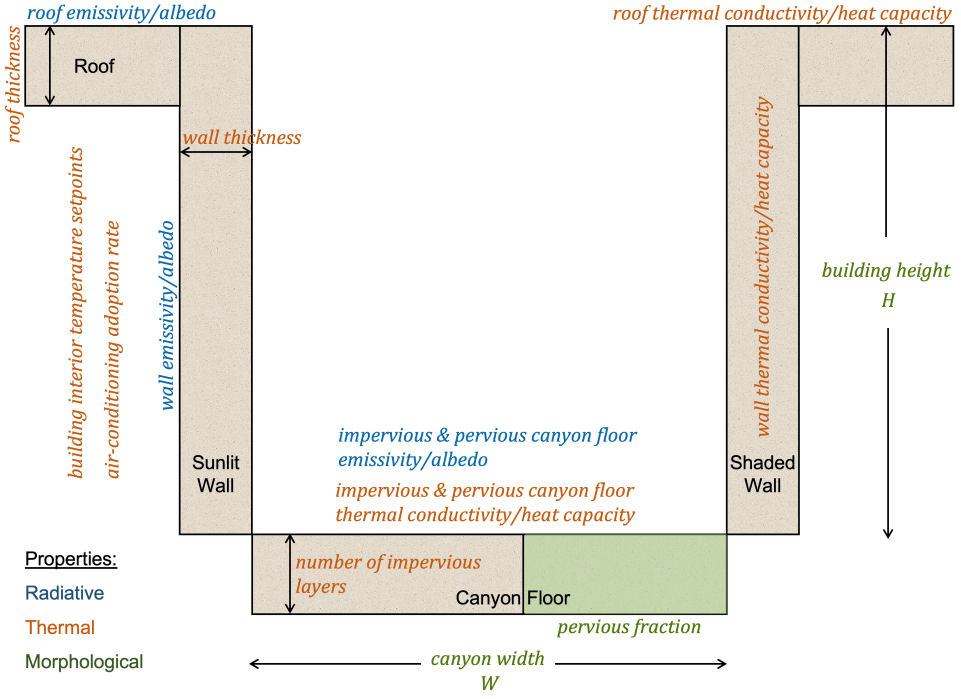

Figure 1. Conceptual schematic of an urban canyon to represent urban landscape in CLMU (adapted from Oleson et al., 2008a). Properties are color-coded: blue for radiative, orange for thermal and green for morphological. Note that roof and wall thickness (despite being related to urban morphology) are considered thermal properties, as they are primarily used as weighting factors to calculate conduction fluxes into and out of canyon surfaces in CLMU (Lawrance et al., 2018; Oleson et al., 2010).

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Open Access Earth System Science Data Discussions

Table 1. Data sources and retrieval methods for each urban parameter in U-Surf and the CLMU dataset.

Category	Urban Parameters	U-Surf	CLMU
Radiative	Roof Emissivity	• Source: 100m ASTER v3 emissivity product (Hulley et al., 2015) • Time span: static, representing 2000-2008 • Spatial resolution: 1km	• Source: local building codes, municipal documentation, literature, satellite imageries (Jackson et al., 2010; Oleson and Feddema, 2020) • Time span: 1966-2007 • Spatial resolution: Regional-level, density-class-specific
	Impervious Canyon Floor Emissivity		
	Pervious Canyon Floor Emissivity		
	Wall Emissivity*		
	Roof Albedo	• Source: 10m Sentinel2 albedo product (Lin et al., 2022) and narrow-to-broadband algorithm (Bonafoni & Sekertekin, 2020) • Time span: 2019-2022 • Spatial resolution: 1km	
	Impervious Canyon Floor Albedo		
	Pervious Canyon Floor Albedo		
	Wall Albedo*		
Morphological	Building Height	• Source: 3D-GloBFP (Che et al., 2024), 3D building structure (Li et al., 2022) • Time span: 2014-2021 • Spatial resolution: 1km	• Source: local building codes, municipal documentation, literature, satellite imageries (Jackson et al., 2010; Oleson and Feddema, 2020) • Time span: 2000-2007 • Spatial resolution: Regional-level, density-class-specific
	Canyon Height-to-width Ratio	• Source: infinite canyon street model (Masson et al., 2020) • Time span: 2014-2021 • Spatial resolution: 1km	
	Roof Fraction	• Source: Microsoft global building footprints (Microsoft, 2024), East Asia building footprints (Shi et al., 2024) • Time span: 2014-2021 • Spatial resolution: 1km	
	Pervious Canyon Floor Fraction	• Source: 10m ESA Worldcover v200 (Zanaga et al., 2022) • Time span: 2021-2022 • Spatial resolution: 1km	
	Urban Percentage	• Source: building footprints (Microsoft, 2024, Shi et al., 2023) and ESA (Zanaga et al., 2022) • Time span: 2014-2021 • Spatial resolution: 1km	• Source: LandScan global population database (Bright et al., 2005) • Time span: 2004 • Spatial resolution: 1km
Thermal	Air Conditioning Penetration Rate	• Source: global AC penetration rate (Li et al., 2024) • Time span: present-day, loosely defined as 2010-2020 • Spatial resolution: national and sub-national level	AC penetration rate is not explicitly modeled in CLMU as of CLM5 (Oleson and Feddema, 2020)
	Number of Impervious Canyon Layers	• Source: local building codes, municipal documentation, literature, satellite imageries (Jackson et al., 2010; Oleson and Feddema, 2020) • Time span: 1966-2007 • Resolution: Regional-level, density-class-specific**	
	Roof Thickness		
	Wall Thickness		
	Minimum Interior Building Temperature		
	Maximum Interior Building Temperature		
	Roof Thermal Conductivity		
	Impervious Canyon Thermal Conductivity		
	Wall Thermal Conductivity		
	Roof Volumetric Heat Capacity		
	Impervious Canyon Volumetric Heat Capacity		
	Wall Volumetric Heat Capacity		

* Wall emissivity and albedo are derived leveraging the remote sensing data and CESM2 default J2010 radiative data which was based on building materials.
** Although thermal properties in U-Surf are provided at a 1 km resolution, the values are derived from regional-level and density-class-specific propertiesfrom Oleson and Feddema (2020).

170

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Open Access Earth System Science Data Discussions

2.2 Development of the U-Surf urban parameters

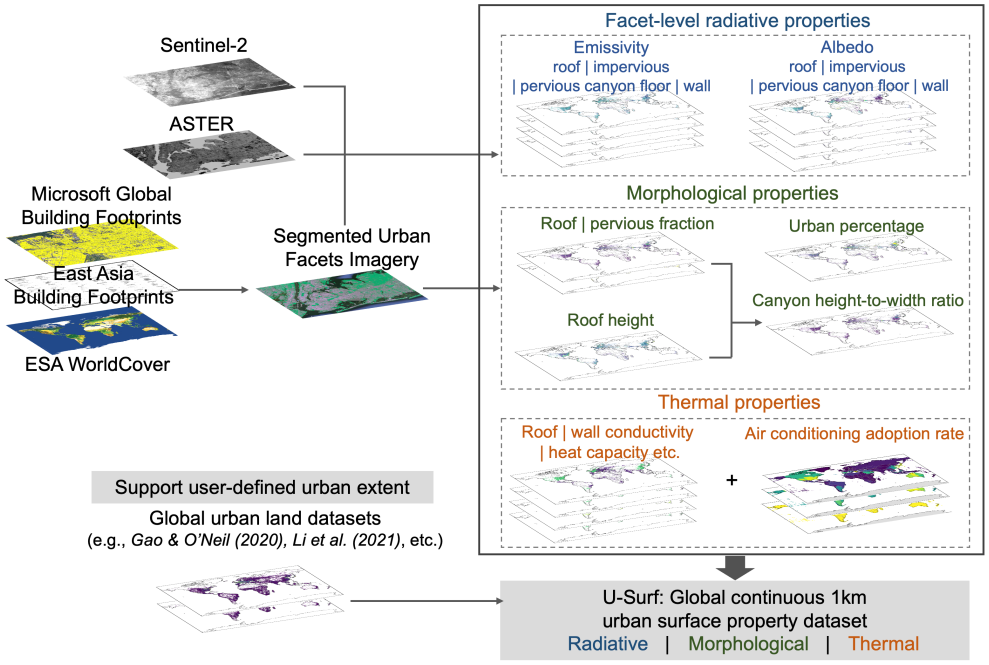

Figure 2. Overview of data synthesis workflow, including individual data sources and examples of final data product layers. The satellite images are from © Google Earth Engine.

175

The new urban surface parameter dataset, U-Surf, describes the global urban areas in a spatially continuous manner, providing all the required parameters in three categories (radiative, morphological, and thermal) that are compatible with the urban canyon representation in CLMU and potentially also for other UCMs. We developed a multi-step federated workflow that leverages four key categories of data: segmented land use/land cover map, 3D building
180 footprints, high-resolution satellite observations, and thermal properties of construction materials. Utilizing these, we first generated segmented urban imagery, which distinguishes among different facets. Then we integrated this imagery with satellite observations to derive facet-level radiative properties and fractional parameters. From there we synthesized multiple data sources to construct the 3D urban canyon morphological attributes. Finally, we incorporated existing databases to produce thermal properties (Figure 2).

185 #### 2.2.1 Radiative parameters

To derive facet-level radiative properties, we first needed to identify individual facets such as building roofs, impervious and pervious ground within each 1 km grid. Here in this study, we use the open-source vector-based Microsoft Global Building Footprints dataset (hereafter referred to as MS-BFP; Microsoft, 2022) in conjunction with the East-Asia building footprints from Shi et al., (2024) and Che et al., (2024) to identify building roofs. The additional
190 East-Asia dataset is necessary due to the insufficient building vectors in the current version of MS-BFP for that region.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

We then combined this data with the European Space Agency (ESA) WorldCover (Zanaga et al., 2022), a 10 m resolution global land cover product based on Sentinel-1 and 2 data, to characterize impervious and pervious canyon floors. We choose the ESA WorldCover instead of other available global 10 m land cover products since its 'built-up' class is definitionally consistent with the impervious surfaces in CLMU (Chakraborty et al., 2024). Accordingly, the impervious canyon floor was estimated by subtracting the roof pixels (derived from the MS-BFP vectors) from areas classified as "built-up" class; whereas pervious surfaces are estimated by aggregating the "tree cover", "shrubland", "grassland", and "bare or sparsely vegetated" areas identified in the ESA WorldCover. This process results in a segmented global urban facet image that serves as the foundation for our subsequent derivation of facet- and canopy-level radiative and morphological parameters.

This facet-segmented image was then applied to the Advanced Spaceborne Thermal Emission and Reflection (ASTER) Global Emissivity Dataset 100 m V003 product (hereafter referred as ASTER GEDv3; Hulley et al., 2015) and the Sentinel-2 land surface albedo data (Lin et al., 2022) to extract the emissivity and albedo of building roof, and impervious and pervious ground. For emissivity, we use a linear spectral-to-broadband algorithm (Malakar et al., 2018) to estimate the broadband emissivity from ASTER GEDv3 bands (Eq. 1):

$$\varepsilon_b = 0.128 + 0.014\varepsilon_a^{10} + 0.145\varepsilon_a^{11} + 0.241\varepsilon_a^{12} + 0.467\varepsilon_a^{13} + 0.004\varepsilon_a^{14} \quad (1)$$

where ε_b is the broadband emissivity, ε_a^{10} to ε_a^{14} denote the ASTER mean emissivity of bands 10 to 14, respectively, respectively, which are the five thermal infrared bands with 90 m resolution.

Note that the 100 m resolution of ASTER GEDv3 could be too coarse for certain small individual facets (e.g., small roof tops, narrow roads between buildings) and thus a potential source of uncertainty. However, given the relatively narrow range of emissivity values (i.e., near blackbody) of typical materials and natural surfaces (Oke et al., 2017), this uncertainty is likely small.

For albedo, we used a 10 m land surface blue-sky albedo product retrieved from Sentinel-2 which covers nearly 2,300 major cities across the globe (Lin et al., 2022). For the rest of the global urban areas, we applied the narrow-to-broadband conversion method (Bonafoni and Sekertekin, 2020) to estimate the 10m-resolution albedo based on Sentinel-2 surface reflectance (Eq. 2). Cloud-contaminated pixels were masked using the Cloud Score+ (CS+) dataset (Pasquarella et al., 2023), where pixels with a CS+ quality assessment score below 0.8 were excluded.

$$\alpha = 0.2266\rho_{B2} + 0.1236\rho_{B3} + 0.1573\rho_{B4} + 0.3417\rho_{B8} + 0.1170\rho_{B11} + 0.0338\rho_{B12} \quad (2)$$

where α is the broadband surface albedo, ρ_{B2} to ρ_{B12} represent the surface reflectance for bands B2 to B12 of Sentinel-2 Multispectral Instrument (MSI) respectively. The 10 m resolution Sentinel-2 albedo data provides the fine granularity to differentiate between roof, impervious, and pervious canyon floor.

The derived facet-level emissivity and albedo were then aggregated to the 1km grids using an area-weighted approach:

$$\varepsilon_f^{1km} = \frac{\sum_{i=0}^{99} w_f^i \cdot \varepsilon_i^{100m}}{\sum_{i=0}^{99} w_f^i} \quad (3)$$

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

$$\alpha_f^{1km} = \frac{\sum_{i=0}^{9999} w_f^i \cdot \alpha_i^{10m}}{\sum_{i=0}^{9999} w_f^i} \quad (4)$$

where ε_f^{1km} and α_f^{1km} is 1 km emissivity and albedo, respectively, for a certain facet (roof/impervious canyon floor/pervious canyon floor), w_f^i is the area fractions of a certain facet within each 100m or 10m grid cell. The subscript f stands for each individual facet.

Because satellites mostly sample roofs (canopy top) and canyon floors, wall emissivity and albedo can hardly be measured from passive satellite remote sensing. To address this issue, we leveraged the CLMU radiative data which was based on building materials. Specifically, for wall emissivity, we assume it shares the same/similar emissivity as its building roof, since the wall surfaces within the same building could either have analogous material compositions in interior layers with roofs or not deviate much in terms of emissivity value given the nature of its narrow range. This will be further discussed in Results and Discussion (Sect. 3.1 and 3.2). For wall albedo, we assume that the ratio of material-based roof albedo to wall albedo in J2010 approximately holds for our data. We then applied this roof-to-wall albedo ratio calculated from J2010 to our new satellite-based roof albedo to derive the wall albedo at each 1 km grid.

2.2.2 Morphological parameters

The morphological parameters in the "urban canyon" conceptual model (particularly the fractional parameters) are normally defined with respect to its corresponding "urban" landscape. In U-Surf, we combined the Global Urban Boundaries (GUB, Li et al., 2020) and the ESA WorldCover data to identify and sufficiently preserve "urban" or "built-up" landscapes as much as possible. Developed based on the Global Artificial Impervious Area data (Gong et al., 2020), the GUB dataset provides a collection of physical boundaries of global urban extents. We first overlaid the GUB polygons with the ESA WorldCover map to identify all "urban" surfaces recognized by GUB. For the grids falling outside of the urban boundaries, we applied a 10×10 (i.e., 100m × 100m) window on the ESA WorldCover data and calculated its "built-up" fraction (i.e., the sum of roof and impervious canyon floor fraction) within the window. If the "built-up" fraction is larger than 10%, we define the window as "urban". We chose a threshold of 10%, which is at the lower end of the typical thresholds used in the literature (10% - 30%), to preserve as many "urban" grids as possible. The urban fraction was then calculated based on the proportional areas of roof, impervious and pervious canyon floor, following the GUB-defined thresholds. Although this will likely result in an inclusion of some "sub-urban" landscapes in the U-Surf raw data, users have the flexibility to apply stricter criteria (larger built-up thresholds) to extract "urban" grids according to their own definition. This approach is designed to maximize the retention of grids, ensuring U-Surf's adaptability to various user-defined urban extents (Figure 2; e.g., Gao and O'Neill, 2020; Li et al., 2021; Zhao et al., 2022; Zhou et al., 2015). The roof fraction is then defined as the ratio between roof area and urban horizontal surface area, where roof area is calculated from the building footprints polygons. Consistent with the definition in CLMU, the pervious fraction is defined as the ratio of pervious canyon floor to the sum of impervious and pervious canyon floors.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

The building height (H) was obtained primarily from the 3D-GloBFP data (Che et al., 2024) and supplemented by another building height dataset by Li et al., (2022) to maximize the spatial coverage. The 3D-GloBFP is a global building height data at a building footprint scale recently developed by leveraging a combination of Synthetic Aperture Radar (SAR), optical imagery, terrain, population, nighttime light data, and XGBoost machine learning approach. We aggregated the vector-level height to 1 km grids using area-weighted averages. The second global building height data (Li et al., 2022) is a raster map at 1 km spatial resolution that also utilizes radar and optical satellite imagery, along with additional geographical information. To comply with the CLMU requirement, we calculated another building height-related parameter: the height at which wind speed in urban canyons is computed. This parameter is simply set at half the building height in the current version of CLMU, providing a standardized reference point for wind calculations in urban environments.

Canyon height-to-width ratio (H/W; i.e., the ratio of building height to canyon width) is another critical morphological parameter that is widely used in most UCMs including CLMU. It is a proxy parameter that implies the structural layout and compactness of the built area. Consistent with the urban canyon geometry in UCMs, the H/W in this study is estimated using the 2D infinite street canyon model with two recommended primary parameters, building fraction (or plan area density; λ_p) and wall surface density (λ_w) (Masson et al., 2020):

$$H/W = \frac{\lambda_w}{2(1-\lambda_p)} \quad (5)$$

where λ_w is calculated as the ratio between the surface wall area that is in direct contact with the atmosphere (i.e., external wall surfaces, A_w) and the horizontal urban surfaces, as represented in the building footprints (MS-BFP; Che et al., 2024; Shi et al., 2024); and λ_p is building fraction (i.e., roof fraction) as described above. The external wall surfaces area is estimated by

$$A_w = N \cdot P_b \cdot H_b \quad (6)$$

where N, P_b and H_b are the number of buildings, the average perimeter of buildings, and the height of buildings within each 1km grid, respectively.

2.2.3 Thermal parameters

Requirements of thermal parameters are relatively diverse compared to radiative and morphological parameters among various UCMs, depending not only on the UCM's parameterization itself but also on whether and what type of a building energy model is in place s (Reinhart and Cerezo Davila, 2016; Sezer et al., 2023). The thermal parameters required in CLMU include volumetric heat capacity and thermal conductivity of roofs, impervious canyon floors and walls, thickness of roofs and walls, minimum/maximum building interior temperature, as well as the penetration rate of air conditioning (AC). These parameters are exceptionally challenging to acquire on a large scale, as they cannot be detected by satellite remote sensing. The most feasible way to estimate these parameters by far is still from information about dominant construction materials combined with local surveys and building code, which is largely the approach used in J2010 data.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024


Open Access Earth System Science Data Discussions

However, because of the coarse resolution of previous versions of CLMU, the capability of the J2010 thermal parameters data has not yet been taken full advantage of. Here we adapt the thermal parameters from J2010 raw data to U-Surf, aiming to better leverage its material-based estimates. J2010 compiled a comprehensive look-up table based on the thermal properties of 49 types of construction materials from imagery, construction data and documentations by country (Jackson et al., 2010). This table includes thickness, thermal conductivity and volumetric heat capacity of up to 10 layers for common types of roofs, walls, roads (layers with identical materials are allowed) (Oleson and Feddema, 2020). Based on the percentiles of canyon height-to-width ratio from J2010, we classified 1km U-Surf urban grids into four nominal density classes: TBD (0.016% of the pixels), HD (3.83%), MD (41.98%), LD (54.17 %) (Figure S2). We then applied the corresponding thermal parameters from the lookup table to each class to ensure it covers all possible materials used in 33 regions (Figures S19-S26).

The AC adoption rate (P_{AC}) is a new thermal parameter added to the latest version of CLMU/CESM because of the introduction of a new explicit-AC-adoption scheme in the building energy model of CLMU (Li et al., 2024). Along with this new scheme, Li et al., (2024) also created a present-day, global, survey-based, and spatially explicit AC adoption rate dataset at country and sub-country level. The AC adoption rate data are created by leveraging U.S. EIA data, literature reports, national surveys, government documentation, the AC units per household data from the International Energy Agency (IEA). More details on this new P_{AC} data are discussed in Li et al., (2024). We incorporated this P_{AC} dataset into our new U-Surf dataset by producing the density-class-weighted averages at 1 km resolution.

All the source data, estimation and/or processing methods, and the comparison with CLMU urban surface data are summarized in Table 1.

2.3 Masking, gap filling and quality control

After estimating all the required parameters as described above, we took several additional steps to ensure the accuracy, coherence, and transparency of our U-Surf data product, including masking, gap filling, and quality control. First, we only retain the grids containing all three facets – roofs, impervious and pervious canyon floors in U-Surf, because a complete "urban canyon" can only be formed when all the three facets are present. This constraint helps make U-Surf more consistent with the conceptual definition of physical urban land in an UCM and is an improvement over the J2010 dataset, which used urban density classes from population estimates (LandScan), leading to large over- and under-estimations of physical urbanization depending on region (Chakraborty et al., 2024). Second, we masked out the grids with extremely high canyon height-to-width ratios (>12) yet low building heights (<40 m). These grids are actually very sparsely built suburban or rural landscapes instead of densely built areas.

The last step is to gap-fill the missing values caused by synthesizing multiple datasets with different spatial coverage. For example, the emissivity product from ASTER GEDv3 has missing pixels in certain regions due to cloud coverage. These missing values were gap-filled using a simple approach. We combined two classification data: the Koppen

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

climate zones (Beck et al., 2018) and the 33 urban regions defined in J2010, both at 1 km-resolution. The average parameter values for each combined class were then used to fill the missing values of the corresponding parameters. Note that only a small proportion of grids needs to be gap-filled, accounting for less than 3.5% of the total among all parameters. To keep the aforementioned data source, processing, and gap-filling information accessible and to make it easier for users to track changes in future version releases, each parameter comes with an additional quality control (QC) band using a 4-digit code (Table 2). The first and second digits differentiate algorithms and single/multi source data, respectively, while the last two digits indicate whether the parameter was directly derived or gap filled. These QC codes are consistent across the entire dataset and will be updated accordingly in later versions.

Table 2. Quality control flags for U-Surf dataset. Note that percentages in the parentheses represent the percentage of grid cells with the corresponding QC flag.

Category	Radiative								Morphological			
Variable	Emissivity				Albedo				Fraction		Building height	Canyon height-to-width ratio
	Roof	Impervious canyon floor	Pervious canyon floor	Wall	Roof	Impervious canyon floor	Pervious canyon floor	Wall	Roof	Pervious canyon floor		
1st digit (Algorithm)	1	1	1	2	1	1	1	2	1	1	3	2
	1: Processing based on observation products, 2: Processing based on model/assumptions, 3: Regridding of existed products w/o further change											
2nd digit (Source)	0	0	0	0	1	1	1	1	0	0	1	1
	0: Single source, 1: Multiple source											
3rd digit (Gapfill)	00, 99	00, 99	00, 99	00, 99	00, 99	00, 99	00, 99	00, 99	00	00	00, 99	00, 99
	00: Direct derivation, 99: Gapfilled values											
QC_flag	1000 (98.51%)*, 1099 (1.49%)*	1000 (96.63%), 1099 (3.37%)	1000 (98.71%), 1099 (1.29%)	2000 (98.51%), 2099 (1.49%)	1100 (98.35%), 1199 (1.65%)	1100 (96.85%), 1199 (3.15%)	1100 (98.60%), 1199 (1.40%)	2100 (98.34%), 2199 (1.66%)	1000 (100.00%)	1000 (100.00%)	3100 (96.59%), 3199 (3.41%)	2100 (96.59%), 2199 (3.41%)

* percentages in parentheses represent the percentage of grid cells with the corresponding QC flag.

2.4 Dataset validation

Validating urban surface parameters on the global scale is extremely challenging primarily due to the lack of globally consistent measurement networks. This challenge is exacerbated by the scarcity of long-term urban observational sites, especially in diverse urban environments. The inherent variability within urban areas further complicates validation efforts, as data from one site may not represent the broader urban landscape. U-Surf is composed of extraction of satellite measurement, satellite-derived products (i.e., land cover data and building footprints), and our own derived parameters. The satellite measurements and derived products have been validated and quality-controlled, parameters derived based on which are therefore subject to their inherent uncertainties. These validation and uncertainty quantification are summarized in Table 3 in Sect. 3.4. We have further conducted a thematic validation on the derived morphological parameters against the 3D World Settlement Footprint (WSF-3D, Esch et al., 2022) observational site data and site meta information from Urban-PLUMBER at 1 km resolution (see Sect. 3.4 for detailed discussion).

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

3 Results and discussion

3.1 Global distribution of 1km urban surface property parameters

U-Surf demonstrates significant improvements over the default CLMU parameters. Although U-Surf provides spatially continuous urban surface parameters without any density or land use classification, here just for illustrative purpose and easier comparison, we plot the distributions of the urban surface parameters in both U-Surf and J2010 data at the TBD, HD, MD, and LD locations defined in J2010 (Figure 3). The overall distribution of U-Surf raw data is also shown in the figure. As most of the thermal properties in U-Surf are adapted from J2010, the discussion here will be mainly focused on radiative and morphological parameters and the comparison of thermal properties can be found in Figure S3.

Retrieved from direct remote sensing measurements, the radiative properties exhibit physically more reasonable ranges compared to J2010 data. As described above, urban surface properties in J2010 were estimated on the basis of building materials sampled in a predefined region and then generalized to the entire region. This clearly leads to unreasonable values, such as abnormally low emissivity and high albedo, across entire regions for certain countries. The former issue is not only true for CLMU, but also for urban emissivity constraints in regional models like WRF (Chakraborty et al., 2021). For instance, the minimum roof emissivity in J2010 is as low as 0.04 in regions like Mongolia, Kazakhstan, France and Germany (Figure 3, Figure S6) and the roof albedo can be as high as 0.61 for Chili, Argentina, Mongolia and Kazakhstan (Figure 3, Figure S10). These values were derived from specific low-emissivity and high-albedo materials (e.g. zinc/galvanized steel coating), which might be possible for individual buildings, but are highly unlikely for all urban areas in a large region (Chakraborty et al., 2021). Broadcasting to an entire region from sampled material estimates results in an oversimplified representation of urban surfaces. In contrast, the "effective" emissivity retrieved from ASTER GEDv3 (Hulley et al., 2015) in U-Surf is generally higher and more narrowly concentrated, typically between 0.95 and 1.0 across urban facets, with exceptions in a few specific areas. This pattern also aligns with urban canyon characteristics, where the effective emissivity of an urban canyon is slightly higher than the weighted-average values from all individual components due to the "canyon trapping" effects (i.e., increased absorption from reflections between facets) (Harman et al., 2004; Oke et al., 2017). Likewise, we can observe a narrower spread of roof albedo values concentrated between 0.1 and 0.3 across countries, which align with the aggregated values from the commonly-used urban roof materials such as tiles (0.10 - 0.35), shingles (0.05 - 0.25), and slate (0.08 - 0.18) (Oke et al., 2017). Our results confirm that the blue/clear sky albedo (actual albedo) calculated in U-Surf, an interpolation between white- and black-sky albedo (Liang et al., 1999), represents the real-world conditions more accurately.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

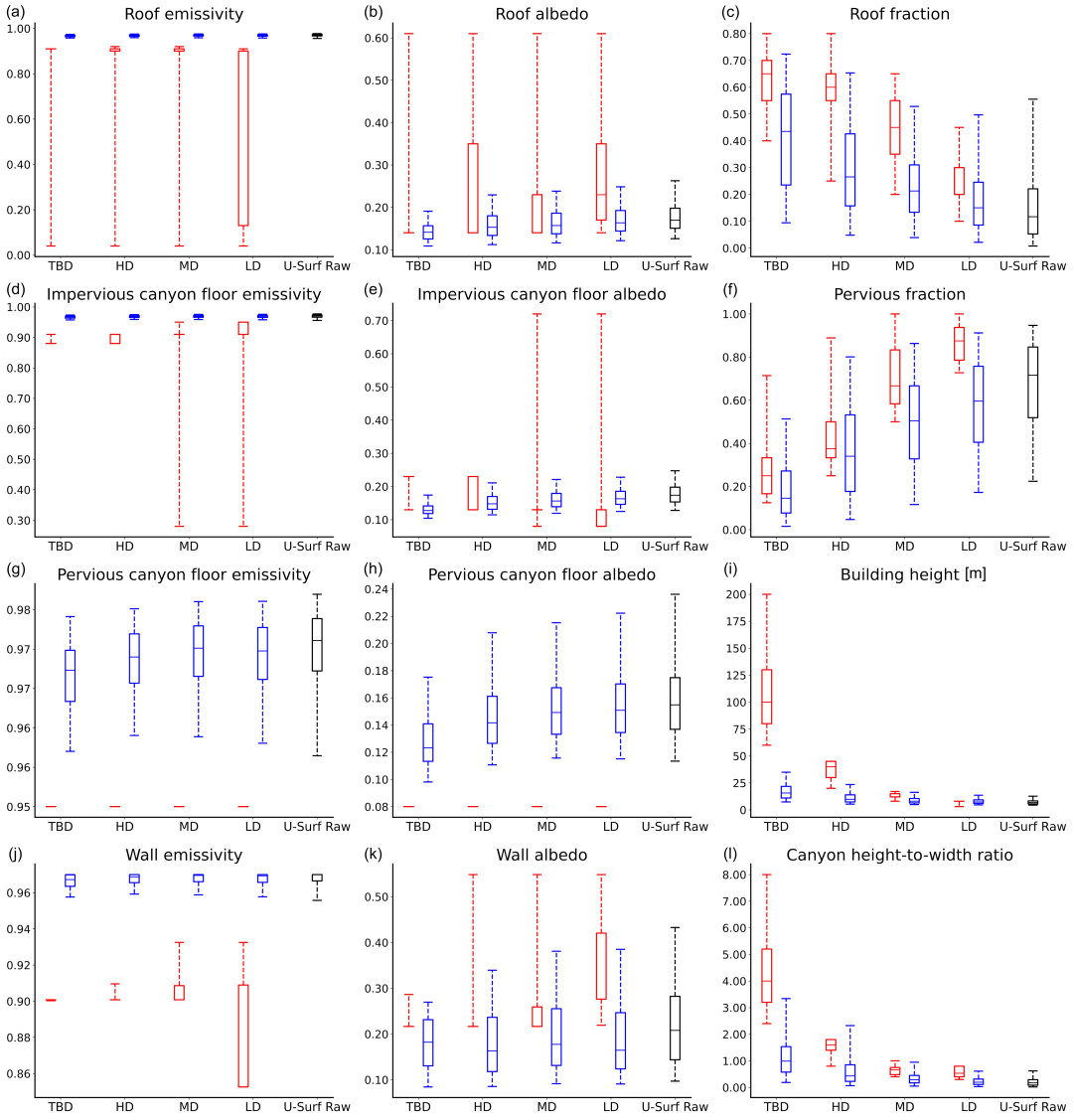

Figure 3. Distribution of urban surface properties at four density classes locations (Oleson and Feddema, 2020)**: Tall Building District (TBD), High Density (HD), Medium Density (MD), and Low Density (LD), compared with raw U-Surf data**. Red bars represent CLMU values (discrete, 33 regions), and blue bars show new U-Surf values (continuous, 1 km) extracted from grids identified as TBD, HD, MD and LD per J2010's definition. The black bars show the distribution of 1 km U-Surf raw data. Box and whisker plots show the 25th, median and the 75th (the bottom, middle and top horizontal bars), and extend to the 5th and 95th quantiles.

The morphological parameters in our dataset also provide more reasonable estimates of both mean values and variability. The four morphological parameters follow similar trends with J2010 in their variations across urban density types. For example, the roof fractions (pervious fractions) are generally higher (lower) in TBD locations identified in J2010, and decrease (increase) as the built density decreases (i.e., HD, MD and LD). However, U-Surf

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

captures much larger variabilities in these parameters compared to J2010, reflecting a more diverse urban morphology. This is again because of the J2010's approach of applying uniform parameter values based on selected representative buildings in a region. This approach not only fails to represent the granular spatial variability in a region, but also easily skews the estimates. For example, J2010 reported an unrealistically high roof fraction of 0.8 for the MD class over Brazil; whereas U-Surf presents a more realistic roof fraction predominantly ranging between 0.06 and 0.22, with a median value of 0.17 for this region, which aligns more closely with observations. Note that the median values of the four morphological parameters in U-Surf raw data (black boxes in Figure 3) are generally lower (or higher in "Pervious Fraction") than TBD, HD, MD and LD categories (blue boxes). This is because U-Surf raw covers more pixels than the location identified as TBD, HD, MD and LD in J2010, most of which are sparsely built landscapes. In fact, the less densely built urban areas dominate the global urban landscapes. The four density classes TBD, HD, MD, and LD in J2010, for example, account for 0.022%, 5.85%, 23.76%, and 70.37% of all urban grids, respectively.

The H/W values in U-Surf are concentrated within ranges of 0.6-1.6, 0.3-0.9, and 0.2-0.5 for TBD, HD, and MD locations identified in J2010, respectively. These values are close to real-world observations which typically vary between 0.5 and 2 at the neighborhood scale (Vardoulakis et al., 2003). Note that high H/W values are very rare at 1-km resolution in real cases. Only very densely built central metropolitan areas (such as the lower Manhattan area in New York City, US) exhibit ratios exceeding 1. These occasions, however, usually only constitute a small proportion. This explains why the overall raw U-Surf H/W values are mostly concentrated between 0.03 and 0.5. We note that in very rare cases, there are some very high roof fraction numbers (> 0.9) in U-Surf which nevertheless are not located in central metropolitan areas. These outliers are places with relatively large roof cover but very small "urban" impervious areas (such as, near the edge of a suburban area).

3.2 Enhanced urban surface properties

In this section, we present selected radiative and morphological parameters as illustrative examples to demonstrate the improvement of the new urban surface dataset in spatial heterogeneity, granularity, accuracy, and broader applicability from global to city scale. The global maps of the complete list of parameters can be found in SI (Figures S6-S27).

U-Surf exhibits significant advancements in capturing the spatial heterogeneity and granularity, which is a crucial improvement over traditional categorical urban classifications, such as density classes used in CLMU and LCZs used in state-of-the-art mesoscale models. The dataset's fine-scale resolution reveals detailed variations in both urban radiative and morphological parameters (Figure 4, Figures S6-S18). For instance, in J2010, the emissivity of pervious canyon floors is uniformly set at 0.95 globally to represent a typical value for vegetation using a bulk parameterization scheme (Oleson et al., 2010), and roof albedo is limited to 11 distinct values (Figures 4a and 4b). These discrete values lead to oversimplifications that fail to represent the critical variations in urban areas, potentially affecting the accuracy of climate predictions. Conversely, U-Surf data shows clear variability both within and across regions (Figures 4c and

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

4d). In general, albedo exhibits greater variability than emissivity across different facets (Figures S6d-S13d). The albedo of impervious canyon floor is comparable with the pervious one, while roof and wall albedo tend to be higher, especially in the city center with densely-built tall buildings (Figures S10d-S13d). In New York City, for example, the mean albedo values are 0.13 for both impervious and pervious canyon floors, while roof albedo averages 0.16 and wall albedo is even higher at 0.22. This pattern is consistent with the fact that commonly used road pavement materials, such as asphalt and concrete, exhibit similarly low albedo values when compared to urban vegetated surfaces like parks and lawns (Oke et al., 2017). Moreover, roofing materials in metropolitan areas often feature higher reflectivity to reduce heat absorption by buildings (Jia et al., 2024), further contributing to the observed differences in albedo. These variations not only reflect differences in materials used but also adaptation strategies to local climate conditions, thereby providing more insights into local climate-sensitive urban design practices.

On the global scale, U-Surf also reveals high-level distinct spatial patterns that correspond to the varying stages of urban development across regions (Figure 5). In the Global North, particularly in Europe and North America, urban areas typically exhibit higher roof fractions, greater building heights, and higher canyon height-to-width ratios. These characteristics are indicative of more developed urban form and well-established infrastructure, often driven by the need to accommodate growing populations in limited spaces. For instance, metropolitan regions in these areas frequently exceed 30-40% roof coverage (e.g., Figure S4b), with average building heights surpassing 30 meters. In contrast, the Global South (Latin America, Africa, and parts of Asia) generally shows lower values for these parameters and higher pervious surface fractions. However, this trend is rapidly changing in emerging economies, including India and Brazil, where cities are experiencing swift urban growth. Additionally, regions such as East Asia exhibit urbanization patterns that are more akin to those in North America and Europe, characterized by high roof fractions (e.g., Figure S4a) and significant vertical development, reflecting rapid industrialization and economic growth that have rapidly transformed the urban landscape over the past few decades (Cai et al., 2022). These observations further demonstrate the fidelity of U-Surf to reveal globally comparable yet regionally nuanced urbanization representations, which are essential for understanding geographical disparities and advancing region-specific sustainable urban development.

At a more localized level, U-Surf uncovers intriguing patterns within countries and even individual cities, offering insights into the complex interactions between urban morphology and local climate conditions. For instance, in the Southwestern United States (e.g., California, Arizona), Northern Africa countries like Egypt and Tunisia, along with Northeastern China, U-Surf captures lower pervious canyon floor emissivity (below 0.93) and higher roof albedo (above 0.25) (Figures 4c and 4d), reflecting the potential impact of arid conditions and the use of high-albedo materials for heat adaptation in hot climates. Furthermore, U-Surf highlights regional differences in surface morphological properties (Figure 5), which play crucial roles in determining local urban climates. Building heights are notably higher along the coasts and in southern regions of Contiguous United States (CONUS), with cities like New York, Chicago, and Miami showing exceptionally high values (> 100m) and corresponding high canyon height-to-width ratios (> 2) in city cores. These cities also exhibit high roof fractions, showing more clustered building patterns in city centers,

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

with density decreasing outwards (Figure S4b). In densely populated developing countries like India and China, high roof fractions exceeding 50% are observed, particularly in regions such as the Indo-Gangetic Plain and the area spanning from the Bohai Economic Rim to the Yangtze River Basin (Figure S4a). In underdeveloped regions of South America and Africa (e.g., Bolivia, Chad) with widely dispersed urban areas, buildings are more sparsely distributed, typically concentrated within fewer metropolitan areas. It is interesting to note that the high-resolution U-Surf data even captures the very densely populated informal settlements (such as the slum areas in Delhi) where buildings are tightly packed and often overcrowded (characterized by high roof fraction and population density) (Figures S4c and S4d). This illustrates the potential use of U-Surf as a valuable tool to better inform socioeconomic disparities in environmental and climate hazards within cities, currently difficult to do using process-based models (Chakraborty et al., 2023; Zhao et al., 2021).

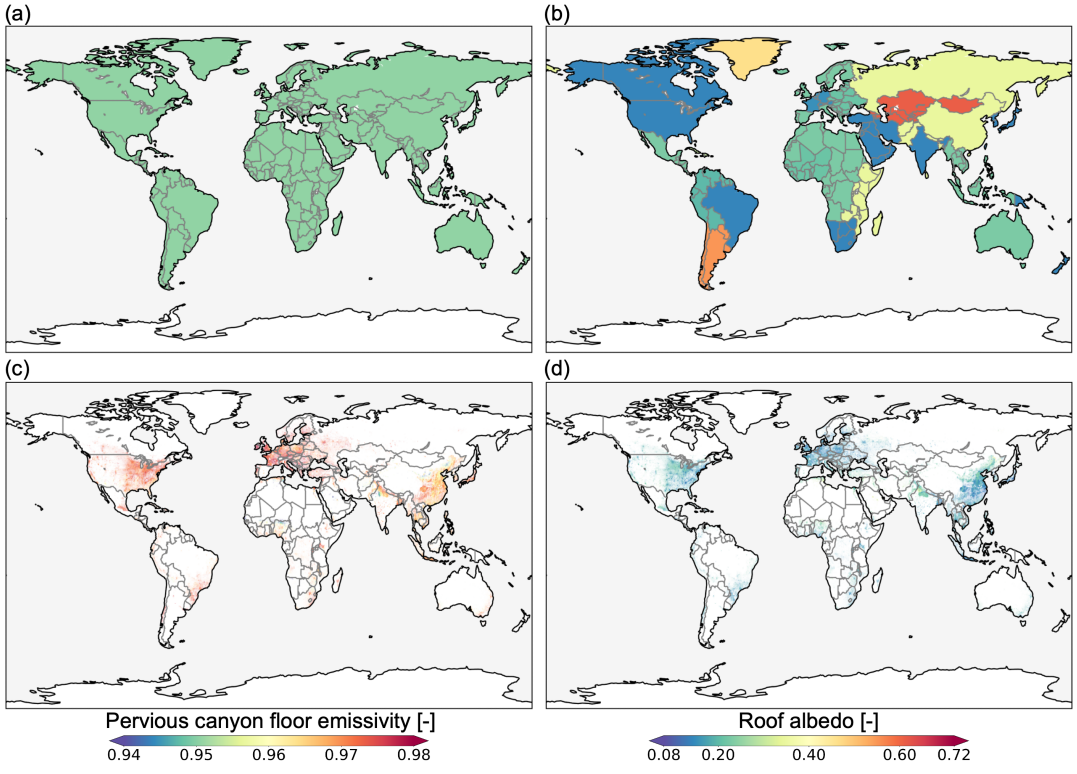

Figure 4. Global-scale comparison between the default CLMU and U-Surf parameters. (a, b) Categorical pervious canyon floor emissivity [unitless] and roof albedo [unitless] over 33 regions (area-weighted averages across TBD, HD and MD) in CLMU; (c, d) 1-km continuous pervious canyon floor emissivity and roof albedo in U-Surf. Each column shares the same color scale range but note that default CLMU parameters only have categorical values over 33 regions and 3 density classes.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

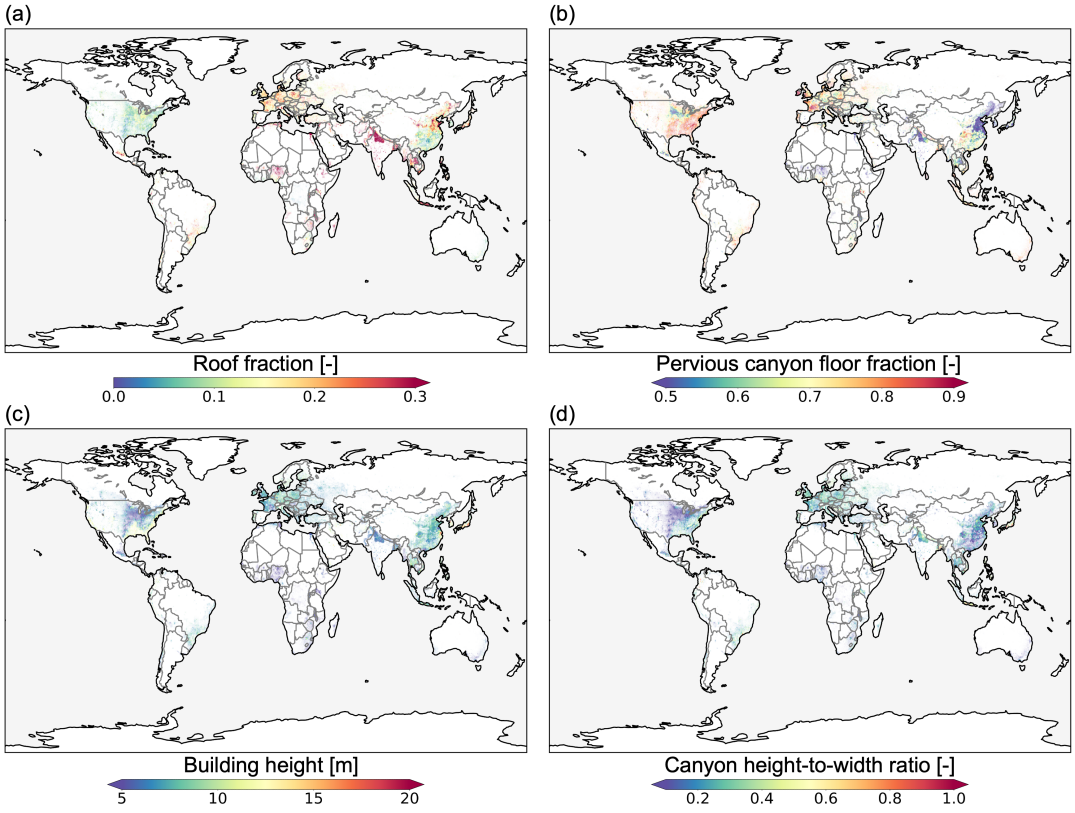

Figure 5. Global spatial distribution of U-Surf morphological parameters: (a) roof fraction [unitless], (b) pervious fraction [unitless], (c) building height [m], (d) canyon height-to-width ratio [unitless].

3.3 Improved urban representation across scales

The high-resolution U-Surf data enables intra- and inter- city comparison in global-scale urban climate modeling, in an unprecedented way. To illustrate this point, we identified two cities with similar background climates: Chicago, IL, USA, and Seoul, South Korea (Figure 6), both of which are classified as *Dfa* under Koppen-Geiger climate classification (humid continental climate with hot summers) (Beck et al., 2018). Because of the coarse-resolution urban surface input in J2010, these two cities share the exact same roof-specific parameters of the MD class. However, U-Surf reveals distinct contrasts in their radiative and morphological properties. Chicago, which has a history of applying heat mitigation strategies such as cool roofs (Mackey et al., 2012; Zhao et al., 2014), demonstrates higher roof emissivity and albedo, with average values of 0.972 and 0.175, respectively, compared to Seoul's 0.955 and 0.114. As to intra-city variations, Chicago's urban form is characterized by a higher concentration of buildings, with an average roof fraction of 0.284, in the northern part of the city. High-rise buildings or skyscrapers are predominantly clustered around Lake Michigan and the Chicago Loop area. On the contrary, Seoul exhibits a more dispersed urban structure, with buildings spread more widely across the city. Such detailed representation facilitates comprehensive

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

attribution and sensitivity analyses, permitting the examination of how individual parameters, such as emissivity, albedo and building height, can alter the city microclimate and further influence the role of cities in local to global climate change scenarios (Krayenhoff et al., 2018, 2021; Zhao et al., 2017) and potentially informs more actionable climate adaptation and mitigation strategies.

U-Surf demonstrates a remarkable ability to capture the spatial heterogeneity and textural details of global urban landscapes across scales. To demonstrate this point, we aggregated the 1km U-Surf data to coarser resolutions of $0.125°$ and nominal 1° (a typical resolution that ESMs are run at) to compare with J2010 side by side. For illustrative purposes, only the comparisons of H/W are shown here. Our results demonstrate that U-Surf represents the detailed urban form considerably well, even at much coarser resolutions. The spatial variability and urban texture are well preserved at global (Figure 7d), national (Figure 7e), and city (Figure 7f) scales. This further demonstrates the adaptability and application of U-Surf in multi-scale, cross-scale urban modeling studies, with potential usage in regionally refined models (RRM) such as E3SM-RRM (Tang et al., 2023a) as well as the variable-resolution models (Huang et al., 2016) like Multi-Scale Infrastructure for Chemistry and Aerosols (MUSICA; Pfister et al., 2020), where seamless transitions between different spatial scales are crucial for comprehensive and coherent analysis.

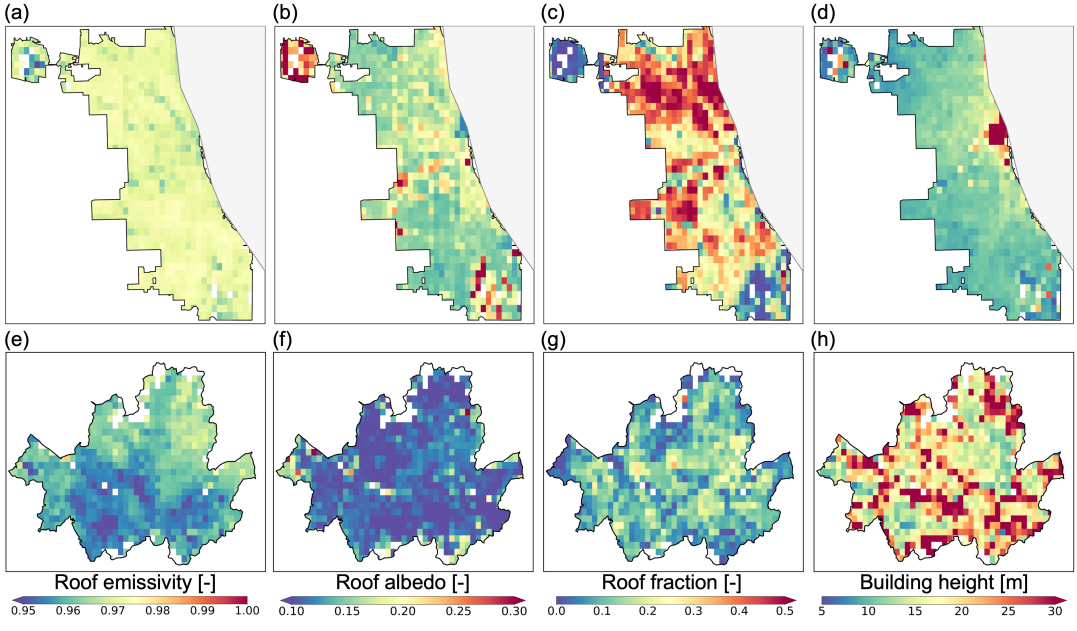

Figure 6. Spatial distributions of roof emissivity [-], roof albedo [-], roof fraction [-], building height [m] in (a-d) Chicago, USA and (e-h) Seoul, South Korea. Each pair of panels within the same column shares a consistent color scale.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

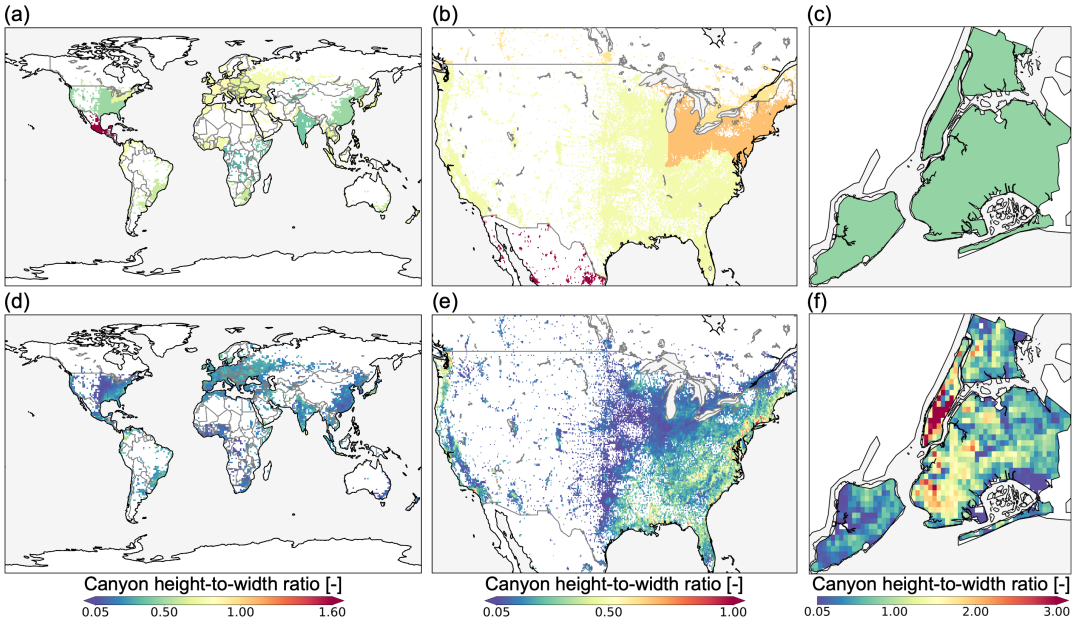

Figure 7. Spatial variability of (a-c) area-weighted averages from CLMU and (d-f) U-Surf surface dataset. The parameter shown here is canyon height-to-width ratio. The spatial resolution from left to right: global nominal 1-deg, 0.125-deg over CONUS, and 1km in New York City, US. Panels in the same column share a colorbar at the bottom of the column. Different colorbar ranges are used to help visualize the distributions across scales.

3.4 Accuracy assessment and uncertainty

As discussed briefly in Sect. 2.4, U-Surf's parameters are inherently influenced by the uncertainties embedded within the synthesized data sources. The validation approaches and associated uncertainties with these sources, as conducted by the development teams, are summarized in Table 3. Specifically, three datasets used to differentiate roofs, impervious and pervious canyon floors demonstrate high global classification accuracy. The 10m-resolution ESA land cover (Zanaga et al., 2022) was validated using updated Copernicus Global Land Service-Land Cover Validation (CGLS-100) dataset. The global overall accuracy across all land cover types is 76.7±0.5%, with the highest accuracy of 82.1±1.0% in Asia and the lowest accuracy of 72.5±1.3% in Oceania. The MS-BFP data (Microsoft, 2022) were evaluated using building polygon labels from Bing Maps, including Maxar and Airbus data. The precision of semantic segmentation (i.e., building pixel detection) showed regional variations with the highest precision of 97.17% in Central Asia and lowest precision of 92.2% in the Caribbean. Correspondingly, Europe had the highest recall of 85.9% while Africa had the lowest recall of 70.9%. The accuracy of polygonization (i.e., conversion from building pixels to polygons) spanned from 63.1% in South Asia to 68.2% in Central Asia. The East Asia building footprints (Shi et al., 2024) were validated in sampled Chinese cities with manual annotation, compared against OSM building data and regional roof vectors (Zhang et al., 2022). It has an overall average accuracy of 89.63% and F1 score of 82.55%. The primary data source of building height underwent rigorous validation against various reference height datasets and

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

selected cities from Google Earth Pro. The validated results showed R^2 values ranging from 0.66 to 0.96 and Root Mean Squared Error (RMSE) from 1.9m to 14.6m across different subregions. The supplementary dataset was also validated and compared against WSF-3D, with a global RMSE of 2.56m.

All remote sensing products and algorithms used to derive radiative properties were validated against ground measurements with high credibility. ASTER GEDv3 (Hulley et al., 2015) was compared with MODIS Collection 4 & 5 Emissivity and validated against lab measurements at four large sand dune fields, yielding relatively low RMSEs. The 10m land blue-sky albedo (Lin et al., 2022), retrieved from Sentinel-2 surface reflectance, was validated against local flux tower measurements, achieving an overall R^2 of 0.94 and RMSE of 0.03 across five land cover types. In addition, the narrow-to-broadband algorithm (Bonafoni and Sekertekin, 2020) demonstrated a R^2 of 0.77 and RMSE of 0.023 when compared against the ground measurements at six Surface Radiation Budget Network (SURFRAD) stations. It also showed a R^2 of 0.98 and RMSE of 0.021 when compared against albedometer measurements at eighteen Perugia sites.

The primary source of uncertainty in the AC adoption rate (Li et al., 2024) stems from the linear model that correlates AC adoption rate with the number of AC units per household. The linear model with saturation effect has an R^2 of 0.9 ($p < 0.001$), RMSE of 11.5 and MAE of 8.5 (both in the unit of %).

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Table 3. Validation and uncertainty analysis of synthesized data products.

Dataset	Source	Validation	Uncertainty
ESA Landcover 2021 v200	Zanaga, D. et al., 2022	Validated using Copernicus Global Land Service-Land Cover Validation dataset	Global accuracy of 76.7 ± 0.5, with the highest accuracy of 82.1 ± 1.0 in Asia and the lowest accuracy of 72.5 ± 1.3 in Oceania and Australia. Treecover: 91.9%, bare/sparse vegetation: 82.5%, grassland: 73.2%, built-up: 66.7%, shrubland: 46.9%
Microsoft Global Building Footprints*	Microsoft, 2022	Evaluated on a set of building polygon labels for each region based on Bing Maps including Maxar and Airbus between 2014 and 2021	Precision: 92.2% (Carribbean) -- 97.17% (Central Asia) Recall: 70.9% (Africa) -- 85.9% (Europe) IoU: 63.1% (South Asia) -- 68.2% (Central Asia) Rotatioin err: 5.67° (Africa) -- 10.28° (Europe)
East Asia Building Footprints	Shi et al., 2024	Validated in sampled Chinese cities with manual annotation, compared against OSM building data and regional roof vectors	Accuracy: 89.63% (average) F1 score: 82.55% Recall: 84.52% Precision rate: 81.06%
ASTER Global Emissivity Dataset v3	Hulley et al., 2015	Validated against lab measurements & MODIS C4, C5 emissivity (2000-2008) over selected 4 sites	RMSE of 0.41%, 0.84%, 0.87%, 0.95% at four sites: Algodone Dunes, Namib, Senegal Basin and Rub Al Khali
Sentinel-2 Albedo	Lin et al., 2022	Validated against ground meaurements & MODIS satellite product at local flux sites	Overall across 5 land cover types: $R^2 = 0.94, RMSE = 0.030$ Cropland: $R^2 = 0.97, RMSE = 0.036$ Decidious broadleaf forest: $R^2 = 0.58, RMSE = 0.027$ Evergreen needleleaf forest: $R^2 = 0.72, RMSE = 0.028$ Grassland: $R^2 = 0.95, RMSE = 0.032$ Open shrubland: $R^2 = 0.92, RMSE = 0.026$
Sentinel-2 NTB Albedo	Bonafoni & Sekertekin, 2020	Validated against ground measurements at selected sites	$R^2 = 0.77, RMSE = 0.023$ when compared against six Surface Radiation Budget Network stations measurements during 2018-2019 $R^2 = 0.98, RMSE = 0.021$ when compared against albedometer measurements at 18 Perugia sites, summer 2016
Building Height	Che et al., 2024	Validated against various reference dataset and selected cities from Google Earth Pro	R^2: 0.66 (Europe) -- 0.96 (South America) RMSE: 1.92m (South America) -- 14.60m (Japan, North and south Korea)
	Li et al., 2022	Evaluated on the validation set, compared against WSF-3D	RMSE 2.56m, MAE 1.39m, SE -0.05m
AC Penetration Rate	Li et al., 2024	35 countries/regions were directly collected; additional linear model was built to map other 34 regions/countries and sub-country level data	Linear model R^2 = 0.9, RMSE=11.5%, MAE=8.5%

*currently there is no official documentation on the validation method, evaluation results were collected from https://github.com/microsoft/GlobalMLBuildingFootprints

For the derived morphological parameters, we conducted a thematic validation based on two recently available, observation-based datasets, Urban-PLUMBER and WSF-3D. WSF-3D is a high-resolution (~ 90 m at the equator) global dataset that provides detailed three-dimensional information on building fraction, height, and volume, derived from satellite imageries, offering crucial insights into urban structures and their spatial distribution across the globe (Esch et al., 2022). We compared the roof fraction and height at 1 km resolution across WSF-3D's 18 validation sites. The Urban-PLUMBER project primarily aims to enhance the understanding of the quality of current urban climate models and has also produced a harmonized dataset of quality controlled and gap-filled observations from 21 diverse urban flux tower sites across different climate zones and urban built environments (Lipson et al., 2022). We compared all four morphological parameters across these sites by using neighboring pixels around the flux towers to evaluate against the site-specific information.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Open Access Earth System Science Data Discussions

The roof fraction showed strong agreement across the reference sites in both WSF-3D and Urban-PLUMBER, with low mean absolute errors (MAEs) of 0.076 and 0.081 (Figure 8a). Similarly, the pervious fraction also aligned well at most Urban-PLUMBER sites, with a mean MAE of 0.124 (Figure S5a). Some discrepancies were observed in building height (Figures 8c and 8d) and canyon height-to-width ratio (Figure S5b). These discrepancies are primarily attributed to the disparity between the neighborhood-scale values captured by flux towers, typically representing areas within several hundreds of meters, and the 1 km-resolution averaged values.

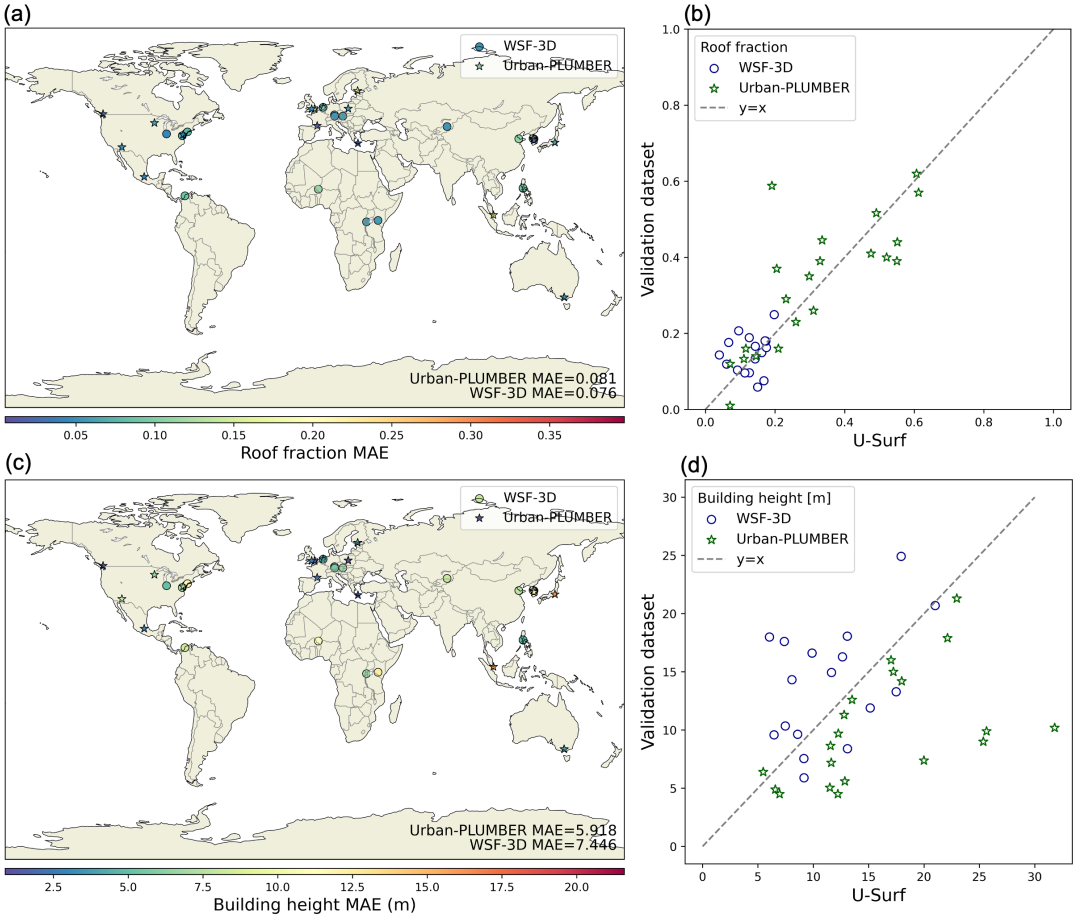

Figure 8. Comparison of two morphological parameters: (a-b) roof fraction, and (c-d) building height, evaluated against 21 Urban-PLUMBER and 18 WSF-3D sites. The numbers labeled on the bottom right corner of (a) and (c) indicate the average mean absolute errors (MAEs) across sites.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

4 Broader implications of U-Surf

The U-Surf dataset significantly advances the development of ultra-high resolution urban-resolving process-based ESMs and RCMs. Its high-resolution capabilities allow for a detailed and refined representation of urban areas, breaking away from the limitations of previous models that relied on coarse regional divisions and outdated classifications. By integrating the latest global data sources, U-Surf provides global continuity and local granularity in urban surface representation. Moreover, the remote-sensing-based methodology offers a unique capability to track the quantitative evolution of urban canopy parameters (UCPs) over time, a level of detail that is difficult to extract from traditional classification methods.

While developed with the architecture of Earth system models in mind (namely CLMU and its versions used in various ESMs), U-Surf can be adapted to other UCMs, such as those embedded in RCMs like WRF, and atmospheric chemistry models such as MUSICA (Pfister et al., 2020; Tang et al., 2023b). Its scalability enables its use in studies ranging from local-scale high-resolution applications to regional and global-scale analyses. Incorporating detailed fine-resolution UCPs (e.g., plan area fraction λ_p, frontal area index λ_F), as demonstrated in WRF studies, is essential for accurately modeling urban climate dynamics (Best and Grimmond, 2014; Georgescu, 2015; Sharma et al., 2017). U-Surf's application in the next-generation kilometer-scale models could help resolve fine-resolution processes such as convection and advection, further advancing the high-fidelity climate and air quality simulations.

Finally, the implications of U-Surf extend beyond the realm of climate or Earth system modeling. This comprehensive dataset provides essential urban informatics and properties on the global scale that can be directly used as key input features for machine learning models (Chajaei and Bagheri, 2024; Furuya et al., 2023; Li et al., 2023b), making U-Surf a valuable resource for both process-based and data-driven modeling. U-Surf can potentially serve as a powerful tool for researchers, policymakers, and urban planners across multiple disciplines. In socio-economic studies, for instance, the dataset can be utilized to explore correlations between urban morphology and economic indicators, potentially revealing relationships between building density, green space distribution, and neighborhood income levels (Chakraborty et al., 2022; Wang et al., 2024). In the public health sector, U-Surf could be used to investigate links between urban structural design and air quality (Zhang et al., 2023b; Zhang and Gu, 2013). Moreover, the dataset could be beneficial for emergency management and disaster preparedness, enabling more accurate risk assessments in densely built areas (Li et al., 2020b; Ma and Mostafavi, 2024).

5 Limitations and future work

We note several limitations, which also present opportunities for future improvements. The accuracy of U-Surf is inherently linked to the uncertainties of the synthesized data sources. For instance, the use of Microsoft global building footprints (mostly 2014-2021, with additional updates up to 2023) may result in missing roofs or land cover misclassification within certain pixels. One primary challenge arises from integrating datasets with varying spatiotemporal coverage. Most of the datasets we utilized reflect urban surface properties from 2014 to 2021. Although

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

the temporal discrepancy among different data sources may introduce additional uncertainties, given the small changes in built surfaces within this short time span, these uncertainties are likely small. Additionally, the spatial resolution of ASTER GEDv3 data is 100 meters, which could be too coarse to accurately distinguish small individual facets, potentially resulting in mixed facet representation. There is also room for improvements in the remote sensing algorithms used to derive some of the raw surface properties incorporated into U-Surf since they are not always calibrated for urban areas (Chakraborty et al., 2021; Chen et al., 2016), though this is beyond the scope of this study. Lastly, we note that certain U-Surf morphological parameters are constructed on the basis of the 2D infinite-street urban canyon conceptual model. Direct application of those parameters should follow the same conceptual assumptions of the urban geometry. Caution should be given if they are used in more complex representations of the real-world urban landscapes.

We plan to continue improving the U-Surf in future versions through multiple aspects. For example, we anticipate that ongoing efforts and continuing endeavor of urban scientific and remote sensing communities will lead to the emergence of more datasets with higher spatial resolutions and accuracies (e.g., more comprehensive building footprints) to be incorporated or updated in U-Surf. We will also adjust the parameter list to reflect advancements in urban parameterization within RCMs and ESMs. For instance, while urban vegetation is not explicitly represented in the current version of CLMU and most operational RCMs, we can follow similar data pipelines and set of constraints (same land cover data, building footprint estimates, etc.) to develop internally consistent global urban vegetation estimates, known to strongly modulate global inter-sample variability in urban climate signals (Chakraborty and Lee, 2019), for the next-generation UCMs in the future. Lastly, depending upon the availability of data sources and new downscaling approaches, we plan to provide temporally varying urban surface properties, which are important for capturing changes in various urban climate signals over time (Chakraborty and Qian, 2024; Fang et al., 2023; Wu et al., 2024).

6 Data availability

The global 1km continuous urban surface property dataset (U-Surf) is publicly available at https://doi.org/10.5281/zenodo.11247599 (Cheng et al., 2024). In addition to the raw dataset at 1km resolution, we also provided the CESM2/E3SM-compatible version at standard resolution ($0.9375° \times 1.25°$) as the ready-to-use input surface dataset for CESM2 simulations. The U-Surf dataset will be incorporated as part of a future release in the high-resolution branch of CTSM (https://github.com/ESCOMP/CTSM).

To facilitate interactive data visualization, query, download, and location-specific analysis, we have further developed a web application using Google Earth Engine (GEE). This interactive platform allows users to explore various urban areas by zooming in on the map, toggle between different parameter layers for comprehensive analysis and extract precise values for all parameters by simply clicking on points of interest. The GEE web application is publicly available at https://ycheng1891.users.earthengine.app/view/global-1km-urban-surface-property-dataset.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

665 The code and intermediate data layers are available from authors upon request.

7 Conclusion

Despite recent advances in urban climate model development across scales, one long-standing critical barrier remains: the absence of a complete, fine-resolution, globally consistent, and spatially explicit urban surface property dataset. Existing products relying on broad categorization underscores the challenge of developing an urban representation
670 that can balance global consistency and local precision. This has been preventing the development of urban-resolving Earth system models for decades, as well as the ultra-high-resolution urban modeling across scales. To address this challenge, we develop a first-of-its-kind global 1km continuous urban surface property dataset – U-Surf. Leveraging recent advancements in remote sensing technologies and machine learning algorithms, U-Surf provides a comprehensive, present-day dataset of urban surface properties that can be used in state-of-the-art ESMs and RCMs.
675

The high-resolution U-Surf data significantly enhances the urban representation in terms of both spatial heterogeneity and accuracy on the global scale, enables detailed city-to-city comparisons in Earth system modeling, and facilitates high-resolution urban climate modeling across scales. By breaking the constraints of predefined urban density classes, the new dataset provides a more nuanced and accurate representation of urban environments worldwide. The remote-
680 sensing-based approach captures the actual surface properties as observed from space, accounting for the complex mixture of materials and structures in urban areas that are difficult to illustrate through traditional bottom-up material-based approaches, which provides more effective and accurate urban canopy parameterization compared to the generalization of material-based values used in previous dataset.

685 The dataset represents a key step forward in advancing the development of ultra-high resolution Earth system modeling. While developed consistent with common ESM architecture in mind, U-Surf can be adapted to quite easily for other models such as weather and regional climate models, and air pollution models, and may be useful inputs for machine learning algorithms. U-Surf also provides useful urban informatics for research and applications across multiple disciplines such as socioeconomics, public health, and urban planning, making it a powerful tool for
690 addressing contemporary challenges in urban development, disaster preparedness, and sustainable city planning. As climate change and urbanization continue to reshape the planet, toolkits like this dataset have increasingly vital roles for understanding future climate-change- and urbanization-driven risks and impacts, further opening up new avenues for research into context-specific guidance for climate-sensitive urban planning and actionable climate adaptation strategies.

695 **Author contributions**

L.Z. and T.C. proposed and designed the study. Y.Cheng performed the data processing, modeling and analysis, and developed the final data product. L.Z., T.C., K.O., M.D., W.L. and Y.Z. contributed ideas to the data collection, modeling, validation, and analysis. Y.Che and W.L. constructed and trained the global building height model, and

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

contributed to building height data curation and analysis. X.L. developed the global AC penetration rate dataset. Y.Cheng and L.Z. drafted the manuscript. All authors edited and revised the manuscript.

Competing interests

One co-author is a member of the editorial board of Earth System Science Data.

Acknowledgements

L.Z. acknowledges the support by the U.S. National Science Foundation (CAREER Award Grant No. 2145362) and the Institute for Sustainability, Energy, and Environment at the University of Illinois Urbana-Champaign. T.C.'s contribution was supported by the U.S. Department of Energy (DOE), Office of Science, Biological and Environmental Research program through the Early Career Research Program. Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. Contributions from K.O. are based upon work supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977. M.D. acknowledges the support of the European Union's HORIZON Research and Innovation Actions under grant agreement No 101137851, project CARMINE (Climate-Resilient Development Pathways in Metropolitan Regions of Europe). We acknowledge the high-performance computing support provided by NSF NCAR's Computational and Information Systems Laboratory, sponsored by the U.S. National Science Foundation. We also thank Justin Braaten of the Google Earth Engine team for helping debug our initial workflow.

References

Ajjur, S. B. and Al-Ghamdi, S. G.: Evapotranspiration and water availability response to climate change in the Middle East and North Africa, Climatic Change, 166, 28, https://doi.org/10.1007/s10584-021-03122-z, 2021.

Baklanov, A., Grimmond, C. S. B., Carlson, D., Terblanche, D., Tang, X., Bouchet, V., Lee, B., Langendijk, G., Kolli, R. K., and Hovsepyan, A.: From urban meteorology, climate and environment research to integrated city services, Urban Climate, 23, 330–341, https://doi.org/10.1016/j.uclim.2017.05.004, 2018.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Sci Data, 5, 180214, https://doi.org/10.1038/sdata.2018.214, 2018.

Best, M. J. and Grimmond, C. S. B.: Importance of initial state and atmospheric conditions for urban land surface models' performance, Urban Climate, 10, 387–406, https://doi.org/10.1016/j.uclim.2013.10.006, 2014.

Bonafoni, S. and Sekertekin, A.: Albedo Retrieval From Sentinel-2 by New Narrow-to-Broadband Conversion Coefficients, IEEE Geosci. Remote Sensing Lett., 17, 1618–1622, https://doi.org/10.1109/LGRS.2020.2967085, 2020.

Cai, Z., Demuzere, M., Tang, Y., and Wan, Y.: The characteristic and transformation of 3D urban morphology in three Chinese mega-cities, Cities, 131, 103988, https://doi.org/10.1016/j.cities.2022.103988, 2022.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Cao, C., Lee, X., Liu, S., Schultz, N., Xiao, W., Zhang, M., and Zhao, L.: Urban heat islands in China enhanced by haze pollution, Nat Commun, 7, 12509, https://doi.org/10.1038/ncomms12509, 2016.

Chajaei, F. and Bagheri, H.: Machine Learning Framework for High-Resolution Air Temperature Downscaling Using LiDAR-Derived Urban Morphological Features, Urban Climate, 57, 102102, https://doi.org/10.1016/j.uclim.2024.102102, 2024.

Chakraborty, T. and Lee, X.: A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability, International Journal of Applied Earth Observation and Geoinformation, 74, 269–280, https://doi.org/10.1016/j.jag.2018.09.015, 2019.

Chakraborty, T. and Qian, Y.: Urbanization exacerbates continental- to regional-scale warming, One Earth, 7, 1387–1401, https://doi.org/10.1016/j.oneear.2024.05.005, 2024.

Chakraborty, T., Lee, X., Ermida, S., and Zhan, W.: On the land emissivity assumption and Landsat-derived surface urban heat islands: A global analysis, Remote Sensing of Environment, 265, 112682, https://doi.org/10.1016/j.rse.2021.112682, 2021.

Chakraborty, T., Biswas, T., Campbell, L. S., Franklin, B., Parker, S. S., and Tukman, M.: Feasibility of afforestation as an equitable nature-based solution in urban areas, Sustainable Cities and Society, 81, 103826, https://doi.org/10.1016/j.scs.2022.103826, 2022.

Chakraborty, T. C., Newman, A. J., Qian, Y., Hsu, A., and Sheriff, G.: Residential segregation and outdoor urban moist heat stress disparities in the United States, One Earth, 6, 738–750, https://doi.org/10.1016/j.oneear.2023.05.016, 2023.

Chakraborty, T. C., Venter, Z., Demuzere, M., Zhan, W., Gao, J., Zhao, L., and Qian, Y.: Large disagreements in estimates of urban land across scales and their implications, https://doi.org/10.21203/rs.3.rs-3958909/v1, 18 March 2024.

Che, Y., Li, X., Liu, X., Wang, Y., Liao, W., Zheng, X., Zhang, X., Xu, X., Shi, Q., Zhu, J., Yuan, H., and Dai, Y.: 3D-GloBFP: the first global three-dimensional building footprint dataset, Earth System Science Data Discussions, 1–28, https://doi.org/10.5194/essd-2024-217, 2024.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., Loridan, T., Manning, K. W., Martilli, A., Miao, S., Sailor, D., Salamanca, F. P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A. A., and Zhang, C.: The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems, International Journal of Climatology, 31, 273–288, https://doi.org/10.1002/joc.2158, 2011.

Chen, F., Yang, S., Su, Z., and Wang, K.: Effect of emissivity uncertainty on surface temperature retrieval over urban areas: Investigations based on spectral libraries, ISPRS Journal of Photogrammetry and Remote Sensing, 114, 53–65, https://doi.org/10.1016/j.isprsjprs.2016.01.007, 2016.

Chen, K., Boomsma, J., and Holmes, H. A.: A multiscale analysis of heatwaves and urban heat islands in the western U.S. during the summer of 2021, Sci Rep, 13, 9570, https://doi.org/10.1038/s41598-023-35621-7, 2023.

Cheng, Y., Zhao, L., Chakraborty, T., Oleson, K., Demuzere, M., Liu, X., Che, Y., Liao, W., Zhou, Y., and Li, X.: U-Surf: A Global 1km spatially continuous urban surface property dataset for kilometer-scale urban-resolving Earth system modeling (1.0.0), Zenodo [data set], https://doi.org/10.5281/zenodo.11247599, 2024.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Ching, J., Mills, G., Bechtel, B., See, L., Feddema, J., Wang, X., Ren, C., Brousse, O., Martilli, A., Neophytou, M., Mouzourides, P., Stewart, I., Hanna, A., Ng, E., Foley, M., Alexander, P., Aliaga, D., Niyogi, D., Shreevastava, A., Bhalachandran, P., Masson, V., Hidalgo, J., Fung, J., Andrade, M., Baklanov, A., Dai, W., Milcinski, G., Demuzere, M., Brunsell, N., Pesaresi, M., Miao, S., Mu, Q., Chen, F., and Theeuwes, N.: WUDAPT: An Urban Weather, Climate, and Environmental Modeling Infrastructure for the Anthropocene, https://doi.org/10.1175/BAMS-D-16-0236.1, 2018.

Conigliaro, E., Monti, P., Leuzzi, G., and Cantelli, A.: A three-dimensional urban canopy model for mesoscale atmospheric simulations and its comparison with a two-dimensional urban canopy model in an idealized case, Urban Climate, 37, 100831, https://doi.org/10.1016/j.uclim.2021.100831, 2021.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001916, https://doi.org/10.1029/2019MS001916, 2020.

Demuzere, M., De Ridder, K., and Van Lipzig, N. P. M.: Modeling the energy balance in Marseille: Sensitivity to roughness length parameterizations and thermal admittance, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009113, 2008.

Demuzere, M., Oleson, K., Coutts, A. M., Pigeon, G., and van Lipzig, N. P. M.: Simulating the surface energy balance over two contrasting urban environments using the Community Land Model Urban, International Journal of Climatology, 33, 3182–3205, https://doi.org/10.1002/joc.3656, 2013.

Demuzere, M., Coutts, A. M., Göhler, M., Broadbent, A. M., Wouters, H., van Lipzig, N. P. M., and Gebert, L.: The implementation of biofiltration systems, rainwater tanks and urban irrigation in a single-layer urban canopy model, Urban Climate, 10, 148–170, https://doi.org/10.1016/j.uclim.2014.10.012, 2014.

Demuzere, M., Harshan, S., Järvi, L., Roth, M., Grimmond, C. S. B., Masson, V., Oleson, K. W., Velasco, E., and Wouters, H.: Impact of urban canopy models and external parameters on the modelled urban energy balance in a tropical city, Quarterly Journal of the Royal Meteorological Society, 143, 1581–1596, https://doi.org/10.1002/qj.3028, 2017.

Demuzere, M., Kittner, J., and Bechtel, B.: LCZ Generator: A Web Application to Create Local Climate Zone Maps, Front. Environ. Sci., 9, https://doi.org/10.3389/fenvs.2021.637455, 2021.

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

Demuzere, M., Argüeso, D., Zonato, A., and Kittner, J.: W2W: A Python package that injects WUDAPT's Local Climate Zone information in WRF, Journal of Open Source Software, 7, 4432, https://doi.org/10.21105/joss.04432, 2022b.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Esch, T., Brzoska, E., Dech, S., Leutner, B., Palacios-Lopez, D., Metz-Marconcini, A., Marconcini, M., Roth, A., and Zeidler, J.: World Settlement Footprint 3D - A first three-dimensional survey of the global building stock, Remote Sensing of Environment, 270, 112877, https://doi.org/10.1016/j.rse.2021.112877, 2022.

Fang, B., Zhao, L., Oleson, K. W., Zhang, K., Lawrence, P. J., Sacks, B., Cao, C., He, C., Huang, Q., Liu, Z., and Lee, X.: Representing dynamic urban land change in the Community Earth System Model (CESM), https://doi.org/10.22541/essoar.168676909.95382628/v1, 14 June 2023.

Feng, B., Zhang, Y., and Bourke, R.: Urbanization impacts on flood risks based on urban growth data and coupled flood models, Nat Hazards, 106, 613–627, https://doi.org/10.1007/s11069-020-04480-0, 2021.

Fitria, R., Kim, D., Baik, J., and Choi, M.: Impact of Biophysical Mechanisms on Urban Heat Island Associated with Climate Variation and Urban Morphology, Sci Rep, 9, 19503, https://doi.org/10.1038/s41598-019-55847-8, 2019.

Furuya, M. T. G., Furuya, D. E. G., de Oliveira, L. Y. D., da Silva, P. A., Cicerelli, R. E., Gonçalves, W. N., Junior, J. M., Osco, L. P., and Ramos, A. P. M.: A machine learning approach for mapping surface urban heat island using environmental and socioeconomic variables: a case study in a medium-sized Brazilian city, Environ Earth Sci, 82, 325, https://doi.org/10.1007/s12665-023-11017-8, 2023.

Gao, J. and Bukovsky, M. S.: Urban land patterns can moderate population exposures to climate extremes over the 21st century, Nat Commun, 14, 6536, https://doi.org/10.1038/s41467-023-42084-x, 2023.

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

Georgescu, M.: Challenges Associated with Adaptation to Future Urban Expansion, Journal of Climate, 28, 2544–2563, https://doi.org/10.1175/JCLI-D-14-00290.1, 2015.

Golaz, J.-C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., Bradley, A. M., Tang, Q., Maltrud, M. E., Forsyth, R. M., Zhang, C., Zhou, T., Zhang, K., Zender, C. S., Wu, M., Wang, H., Turner, A. K., Singh, B., Richter, J. H., Qin, Y., Petersen, M. R., Mametjanov, A., Ma, P.-L., Larson, V. E., Krishna, J., Keen, N. D., Jeffery, N., Hunke, E. C., Hannah, W. M., Guba, O., Griffin, B. M., Feng, Y., Engwirda, D., Di Vittorio, A. V., Dang, C., Conlon, L. M., Chen, C.-C.-J., Brunke, M. A., Bisht, G., Benedict, J. J., Asay-Davis, X. S., Zhang, Y., Zhang, M., Zeng, X., Xie, S., Wolfram, P. J., Vo, T., Veneziani, M., Tesfa, T. K., Sreepathi, S., Salinger, A. G., Reeves Eyre, J. E. J., Prather, M. J., Mahajan, S., Li, Q., Jones, P. W., Jacob, R. L., Huebler, G. W., Huang, X., Hillman, B. R., Harrop, B. E., Foucar, J. G., Fang, Y., Comeau, D. S., Caldwell, P. M., Bartoletti, T., Balaguru, K., Taylor, M. A., McCoy, R. B., Leung, L. R., and Bader, D. C.: The DOE E3SM Model Version 2: Overview of the Physical Model and Initial Model Evaluation, Journal of Advances in Modeling Earth Systems, 14, e2022MS003156, https://doi.org/10.1029/2022MS003156, 2022.

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

Grimmond, C. S. B., Blackett, M., Best, M. J., Baik, J.-J., Belcher, S. E., Beringer, J., Bohnenstengel, S. I., Calmet, I., Chen, F., Coutts, A., Dandou, A., Fortuniak, K., Gouvea, M. L., Hamdi, R., Hendry, M., Kanda, M., Kawai, T., Kawamoto, Y., Kondo, H., Krayenhoff, E. S., Lee, S.-H., Loridan, T., Martilli, A., Masson, V., Miao, S., Oleson, K.,

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Ooka, R., Pigeon, G., Porson, A., Ryu, Y.-H., Salamanca, F., Steeneveld, G. j., Tombrou, M., Voogt, J. A., Young, D. T., and Zhang, N.: Initial results from Phase 2 of the international urban energy balance model comparison, International Journal of Climatology, 31, 244–272, https://doi.org/10.1002/joc.2227, 2011.

Harman, I. N., Best, M. J., and Belcher, S. E.: Radiative Exchange in an Urban Street Canyon, Boundary-Layer Meteorology, 110, 301–316, https://doi.org/10.1023/A:1026029822517, 2004.

He, W., Li, X., Zhou, Y., Shi, Z., Yu, G., Hu, T., Wang, Y., Huang, J., Bai, T., Sun, Z., Liu, X., and Gong, P.: Global urban fractional changes at a 1 km resolution throughout 2100 under eight scenarios of Shared Socioeconomic Pathways (SSPs) and Representative Concentration Pathways (RCPs), Earth System Science Data, 15, 3623–3639, https://doi.org/10.5194/essd-15-3623-2023, 2023.

Hertwig, D., Ng, M., Grimmond, S., Vidale, P. L., and McGuire, P. C.: High-resolution global climate simulations: Representation of cities, International Journal of Climatology, 41, 3266–3285, https://doi.org/10.1002/joc.7018, 2021.

Hidalgo, J., Dumas, G., Masson, V., Petit, G., Bechtel, B., Bocher, E., Foley, M., Schoetter, R., and Mills, G.: Comparison between local climate zones maps derived from administrative datasets and satellite observations, Urban Climate, 27, 64–89, https://doi.org/10.1016/j.uclim.2018.10.004, 2019.

Huang, X., Rhoades, A. M., Ullrich, P. A., and Zarzycki, C. M.: An evaluation of the variable-resolution CESM for modeling California's climate, Journal of Advances in Modeling Earth Systems, 8, 345–369, https://doi.org/10.1002/2015MS000559, 2016.

Huang, X., Liu, A., and Li, J.: Mapping and analyzing the local climate zones in China's 32 major cities using Landsat imagery based on a novel convolutional neural network, Geo-spatial Information Science, 24, 528–557, https://doi.org/10.1080/10095020.2021.1892459, 2021.

Hulley, G. C., Hook, S. J., Abbott, E., Malakar, N., Islam, T., and Abrams, M.: The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale, Geophysical Research Letters, 42, 7966–7976, https://doi.org/10.1002/2015GL065564, 2015.

Intergovernmental Panel On Climate Change (Ed.): Climate change 2014: mitigation of climate change Working Group III contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge university press, New York, 2014.

Intergovernmental Panel On Climate Change: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 1st ed., Cambridge University Press, https://doi.org/10.1017/9781009157896, 2023.

Islam, S. N. and Winkel, J.: Climate Change and Social Inequality, 2017.

Jackson, T. L., Feddema, J. J., Oleson, K. W., Bonan, G. B., and Bauer, J. T.: Parameterization of Urban Characteristics for Global Climate Modeling, Annals of the Association of American Geographers, 100, 848–865, https://doi.org/10.1080/00045608.2010.497328, 2010.

Jia, S., Weng, Q., Yoo, C., Xiao, H., and Zhong, Q.: Building energy savings by green roofs and cool roofs in current and future climates, npj Urban Sustain, 4, 1–13, https://doi.org/10.1038/s42949-024-00159-8, 2024.

Jongen, H. J., Lipson, M. J., Teuling, A. J., Grimmond, S., Baik, J.-J., Best, M. J., Demuzere, M., Fortuniak, K., Huang, Y., Kauwe, M. G. D., Li, R., McNorton, J. R., Meili, N., Oleson, K. W., Park, S.-B., Sun, T., Tsiringakis, A., Varentsov,

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

M., Wang, C., Wang, Z., and Steeneveld, G.-J.: The water balance representation in Urban-PLUMBER land surface models, https://doi.org/10.22541/essoar.170688849.94066619/v1, 2 February 2024.

Kim, S. K., Bennett, M. M., van Gevelt, T., and Joosse, P.: Urban agglomeration worsens spatial disparities in climate adaptation, Sci Rep, 11, 8446, https://doi.org/10.1038/s41598-021-87739-1, 2021.

Krayenhoff, E. S., Moustaoui, M., Broadbent, A. M., Gupta, V., and Georgescu, M.: Diurnal interaction between urban expansion, climate change and adaptation in US cities, Nature Clim Change, 8, 1097–1103, https://doi.org/10.1038/s41558-018-0320-9, 2018.

Krayenhoff, E. S., Broadbent, A. M., Zhao, L., Georgescu, M., Middel, A., Voogt, J. A., Martilli, A., Sailor, D. J., and Erell, E.: Cooling hot cities: a systematic and critical review of the numerical modelling literature, Environ. Res. Lett., 16, 053007, https://doi.org/10.1088/1748-9326/abdcf1, 2021.

Langendijk, G. S., Halenka, T., Hoffmann, P., Adinolfi, M., Aldama Campino, A., Asselin, O., Bastin, S., Bechtel, B., Belda, M., Bushenkova, A., Campanale, A., Chun, K. P., Constantinidou, K., Coppola, E., Demuzere, M., Doan, Q.-V., Evans, J. P., Feldmann, H., Fernandez, J., Fita, L., Hadjinicolaou, P., Hamdi, R., Hundhausen, M., Grawe, D., Johannsen, F., Milovac, J., Katragkou, E., El Islam Kerroumi, N., Kotlarski, S., Le Roy, B. G., Lemonsu, A., Lennard, C., Lipson, M., Mandal, S., Muñoz Pabón, L. E., Pavlidis, V., Pietikäinen, J.-P., Raffa, M., Raluy-López, E., Rechid, D., Rui, I., Schulz, J.-P., Soares, P. M. M., Takane, Y., Teichmann, C., Thatcher, M., Top, S., Van Schaeybroeck, B., Wang, F., and Yuan, J.: Towards Better Understanding the Urban Environment and its Interactions with Regional Climate Change -The Wcrp Cordex Flagship Pilot Study Urb-Rcc, https://doi.org/10.2139/ssrn.4846089, 2024.

Lawrance, D., Fisher, R., Koven, C., Oleson, K., Swenson, S., Vertenstein, M., Andre, B., Bonan, G., Ghimire, B., Kampenhout, L. van, Kennedy, D., Kluzek, E., Knox, R., Lawrence, P., Li, F., Li, H., Lombardozzi, D., Lu, Y., Perket, J., Riley, W., Sacks, W., Shi, M., Wieder, W., Xu, C., Ali, A., Badger, A., Bisht, G., Broxton, P., Brunke, M., Buzan, J., Clark, M., Craig, T., Dahlin, K., Drewniak, B., Emmons, L., Fisher, J., Flanner, M., Gentine, P., Lenaerts, J., Levis, S., Leung, L. R., Lipscomb, W., Pelletier, J., Ricciuto, D. M., Sanderson, B., Shuman, J., Slater, A., Subin, Z., Tang, J., Tawfik, A., Thomas, Q., Tilmes, S., Vitt, F., and Zeng, X.: Technical Description of version 5.0 of the Community Land Model (CLM), 2018.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Val Martin, M., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, Journal of Advances in Modeling Earth Systems, 11, 4245–4287, https://doi.org/10.1029/2018MS001583, 2019.

Li, D. and Bou-Zeid, E. R.: Synergistic interactions between urban heat islands and heat waves: The impact in cities is larger than the sum of its parts, Journal of Applied Meteorology and Climatology, 52, 2051–2064, https://doi.org/10.1175/JAMC-D-13-02.1, 2013.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

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

Li, X., Zhou, Y., Yu, S., Jia, G., Li, H., and Li, W.: Urban heat island impacts on building energy consumption: A review of approaches and findings, Energy, 174, 407–419, https://doi.org/10.1016/j.energy.2019.02.183, 2019.

Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., Hu, T., Xiao, Y., Xu, B., Yang, J., Liu, X., Cai, W., Huang, H., Wu, T., Wang, X., Lin, P., Li, X., Chen, J., He, C., Li, X., Yu, L., Clinton, N., and Zhu, Z.: Mapping global urban boundaries from the global artificial impervious area (GAIA) data, Environ. Res. Lett., 15, 094044, https://doi.org/10.1088/1748-9326/ab9be3, 2020a.

Li, X., Zhou, Y., Hejazi, M., Wise, M., Vernon, C., Iyer, G., and Chen, W.: Global urban growth between 1870 and 2100 from integrated high resolution mapped data and urban dynamic modeling, Commun Earth Environ, 2, 1–10, https://doi.org/10.1038/s43247-021-00273-w, 2021.

Li, X., Feng, M., Ran, Y., Su, Y., Liu, F., Huang, C., Shen, H., Xiao, Q., Su, J., Yuan, S., and Guo, H.: Big Data in Earth system science and progress towards a digital twin, Nat Rev Earth Environ, 4, 319–332, https://doi.org/10.1038/s43017-023-00409-w, 2023a.

Li, X., Yang, B., Liang, F., Zhang, H., Xu, Y., and Dong, Z.: Modeling urban canopy air temperature at city-block scale based on urban 3D morphology parameters– A study in Tianjin, North China, Building and Environment, 230, 110000, https://doi.org/10.1016/j.buildenv.2023.110000, 2023b.

Li, X. "Cathy," Zhao, L., Oleson, K., Zhou, Y., Qin, Y., Zhang, K., and Fang, B.: Enhancing Urban Climate-Energy Modeling in the Community Earth System Model (CESM) Through Explicit Representation of Urban Air-Conditioning Adoption, Journal of Advances in Modeling Earth Systems, 16, e2023MS004107, https://doi.org/10.1029/2023MS004107, 2024.

Li, Y., Schubert, S., Kropp, J. P., and Rybski, D.: On the influence of density and morphology on the Urban Heat Island intensity, Nat Commun, 11, 2647, https://doi.org/10.1038/s41467-020-16461-9, 2020b.

Liang, S., Strahler, A. H., and Walthall, C.: Retrieval of Land Surface Albedo from Satellite Observations: A Simulation Study, JOURNAL OF APPLIED METEOROLOGY, 38, 1999.

Lin, S., Feng, J., Wang, J., and Hu, Y.: Modeling the contribution of long-term urbanization to temperature increase in three extensive urban agglomerations in China, Journal of Geophysical Research: Atmospheres, 121, 1683–1697, https://doi.org/10.1002/2015JD024227, 2016.

Lin, X., Wu, S., Chen, B., Lin, Z., Yan, Z., Chen, X., Yin, G., You, D., Wen, J., Liu, Q., Xiao, Q., Liu, Q., and Lafortezza, R.: Estimating 10-m land surface albedo from Sentinel-2 satellite observations using a direct estimation approach with Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing, 194, 1–20, https://doi.org/10.1016/j.isprsjprs.2022.09.016, 2022.

Lipson, M., Grimmond, S., Best, M., Chow, W. T. L., Christen, A., Chrysoulakis, N., Coutts, A., Crawford, B., Earl, S., Evans, J., Fortuniak, K., Heusinkveld, B. G., Hong, J.-W., Hong, J., Järvi, L., Jo, S., Kim, Y.-H., Kotthaus, S., Lee, K., Masson, V., McFadden, J. P., Michels, O., Pawlak, W., Roth, M., Sugawara, H., Tapper, N., Velasco, E., and

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Ward, H. C.: Harmonized gap-filled datasets from 20 urban flux tower sites, Earth System Science Data, 14, 5157–5178, https://doi.org/10.5194/essd-14-5157-2022, 2022.

Lipson, M. J., Grimmond, S., Best, M., Abramowitz, G., Coutts, A., Tapper, N., Baik, J.-J., Beyers, M., Blunn, L., Boussetta, S., Bou-Zeid, E., De Kauwe, M. G., de Munck, C., Demuzere, M., Fatichi, S., Fortuniak, K., Han, B.-S., Hendry, M. A., Kikegawa, Y., Kondo, H., Lee, D.-I., Lee, S.-H., Lemonsu, A., Machado, T., Manoli, G., Martilli, A., Masson, V., McNorton, J., Meili, N., Meyer, D., Nice, K. A., Oleson, K. W., Park, S.-B., Roth, M., Schoetter, R., Simón-Moral, A., Steeneveld, G.-J., Sun, T., Takane, Y., Thatcher, M., Tsiringakis, A., Varentsov, M., Wang, C., Wang, Z.-H., and Pitman, A. J.: Evaluation of 30 urban land surface models in the Urban-PLUMBER project: Phase 1 results, Quarterly Journal of the Royal Meteorological Society, 150, 126–169, https://doi.org/10.1002/qj.4589, 2024.

Lobo, J., Aggarwal, R. M., Alberti, M., Allen-Dumas, M., Bettencourt, L. M. A., Boone, C., Brelsford, C., Broto, V. C., Eakin, H., Bagchi-Sen, S., Meerow, S., D'Cruz, C., Revi, A., Roberts, D. C., Smith, M. E., York, A., Lin, T., Bai, X., Solecki, W., Pataki, D., Tapia, L. B., Rockman, M., Wolfram, M., Schlosser, P., and Gauthier, N.: Integration of urban science and urban climate adaptation research: opportunities to advance climate action, npj Urban Sustain, 3, 1–9, https://doi.org/10.1038/s42949-023-00113-0, 2023.

Ma, J. and Mostafavi, A.: Urban form and structure explain variability in spatial inequality of property flood risk among US counties, Commun Earth Environ, 5, 1–12, https://doi.org/10.1038/s43247-024-01337-3, 2024.

Mackey, C. W., Lee, X., and Smith, R. B.: Remotely sensing the cooling effects of city scale efforts to reduce urban heat island, Building and Environment, 49, 348–358, https://doi.org/10.1016/j.buildenv.2011.08.004, 2012.

Malakar, N. K., Hulley, G. C., Hook, S. J., Laraby, K., Cook, M., and Schott, J. R.: An Operational Land Surface Temperature Product for Landsat Thermal Data: Methodology and Validation, IEEE Transactions on Geoscience and Remote Sensing, 56, 5717–5735, https://doi.org/10.1109/TGRS.2018.2824828, 2018.

Masson, V., Heldens, W., Bocher, E., Bonhomme, M., Bucher, B., Burmeister, C., De Munck, C., Esch, T., Hidalgo, J., Kanani-Sühring, F., Kwok, Y.-T., Lemonsu, A., Lévy, J.-P., Maronga, B., Pavlik, D., Petit, G., See, L., Schoetter, R., Tornay, N., Votsis, A., and Zeidler, J.: City-descriptive input data for urban climate models: Model requirements, data sources and challenges, Urban Climate, 31, 100536, https://doi.org/10.1016/j.uclim.2019.100536, 2020.

Microsoft: microsoft/GlobalMLBuildingFootprints, 2022.

Oke, T. R., Mills, G., Christen, A., and Voogt, J. A.: Urban Climates, 1st ed., Cambridge University Press, https://doi.org/10.1017/9781139016476, 2017.

Oleson, K. W. and Feddema, J.: Parameterization and Surface Data Improvements and New Capabilities for the Community Land Model Urban (CLMU), Journal of Advances in Modeling Earth Systems, 12, e2018MS001586, https://doi.org/10.1029/2018MS001586, 2020.

Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., and Grimmond, C. S. B.: An Urban Parameterization for a Global Climate Model. Part I: Formulation and Evaluation for Two Cities, Journal of Applied Meteorology and Climatology, 47, 1038–1060, https://doi.org/10.1175/2007JAMC1597.1, 2008a.

Oleson, K. W., Bonan, G. B., Feddema, J., and Vertenstein, M.: An Urban Parameterization for a Global Climate Model. Part II: Sensitivity to Input Parameters and the Simulated Urban Heat Island in Offline Simulations, Journal of Applied Meteorology and Climatology, 47, 1061–1076, https://doi.org/10.1175/2007JAMC1598.1, 2008b.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Oleson, K. W., Bonan, G. B., Feddema, J. J., Vertenstein, M., and Kluzek, E.: Technical Description of an Urban Parameterization for the Community Land Model (CLMU), 2010.

Pasquarella, V. J., Brown, C. F., Czerwinski, W., and Rucklidge, W. J.: Comprehensive quality assessment of optical satellite imagery using weakly supervised video learning, in: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2125–2135, https://doi.org/10.1109/CVPRW59228.2023.00206, 2023.

Pfister, G. G., Eastham, S. D., Arellano, A. F., Aumont, B., Barsanti, K. C., Barth, M. C., Conley, A., Davis, N. A., Emmons, L. K., Fast, J. D., Fiore, A. M., Gaubert, B., Goldhaber, S., Granier, C., Grell, G. A., Guevara, M., Henze, D. K., Hodzic, A., Liu, X., Marsh, D. R., Orlando, J. J., Plane, J. M. C., Polvani, L. M., Rosenlof, K. H., Steiner, A. L., Jacob, D. J., and Brasseur, G. P.: The Multi-Scale Infrastructure for Chemistry and Aerosols (MUSICA), https://doi.org/10.1175/BAMS-D-19-0331.1, 2020.

Qi, M., Xu, C., Zhang, W., Demuzere, M., Hystad, P., Lu, T., James, P., Bechtel, B., and Hankey, S.: Mapping urban form into local climate zones for the continental US from 1986–2020, Sci Data, 11, 195, https://doi.org/10.1038/s41597-024-03042-4, 2024.

Reinhart, C. F. and Cerezo Davila, C.: Urban building energy modeling – A review of a nascent field, Building and Environment, 97, 196–202, https://doi.org/10.1016/j.buildenv.2015.12.001, 2016.

Robinson, A., Lehmann, J., Barriopedro, D., Rahmstorf, S., and Coumou, D.: Increasing heat and rainfall extremes now far outside the historical climate, npj Clim Atmos Sci, 4, 1–4, https://doi.org/10.1038/s41612-021-00202-w, 2021.

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Girolamo, S. D., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T. C., Sprenger, M., Ubbiali, S., and Wernli, H.: Kilometer-Scale Climate Models: Prospects and Challenges, Bulletin of the American Meteorological Society, 101, E567–E587, https://doi.org/10.1175/BAMS-D-18-0167.1, 2020.

Schär, C., Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Di Girolamo, S., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T., Sprenger, M., Ubbiali, S., and Wernli, H.: Prospects for Kilometer-Scale Climate Models, Bulletin of the American Meteorological Society, 102, 47–52, 2021.

Scheuer, S., Haase, D., and Volk, M.: Integrative assessment of climate change for fast-growing urban areas: Measurement and recommendations for future research, PLOS ONE, 12, e0189451, https://doi.org/10.1371/journal.pone.0189451, 2017.

Sezer, N., Yoonus, H., Zhan, D., Wang, L. (Leon), Hassan, I. G., and Rahman, M. A.: Urban microclimate and building energy models: A review of the latest progress in coupling strategies, Renewable and Sustainable Energy Reviews, 184, 113577, https://doi.org/10.1016/j.rser.2023.113577, 2023.

Sharma, A., Fernando, H. J. S., Hamlet, A. F., Hellmann, J. J., Barlage, M., and Chen, F.: Urban meteorological modeling using WRF: a sensitivity study, International Journal of Climatology, 37, 1885–1900, https://doi.org/10.1002/joc.4819, 2017.

Sharma, A., Wuebbles, D. J., and Kotamarthi, R.: The Need for Urban-Resolving Climate Modeling Across Scales, AGU Advances, 2, e2020AV000271, https://doi.org/10.1029/2020AV000271, 2021.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Shi, Q., Zhu, J., Liu, Z., Guo, H., Gao, S., Liu, M., Liu, Z., and Liu, X.: The Last Puzzle of Global Building Footprints—Mapping 280 Million Buildings in East Asia Based on VHR Images, Journal of Remote Sensing, 4, 0138, https://doi.org/10.34133/remotesensing.0138, 2024.

Shu, E. G., Porter, J. R., Hauer, M. E., Sandoval Olascoaga, S., Gourevitch, J., Wilson, B., Pope, M., Melecio-Vazquez, D., and Kearns, E.: Integrating climate change induced flood risk into future population projections, Nat Commun, 14, 7870, https://doi.org/10.1038/s41467-023-43493-8, 2023.

Sjöstrand, K.: Urbanization impacts on floods, Nat Rev Earth Environ, 3, 738–738, https://doi.org/10.1038/s43017-022-00367-9, 2022.

Stewart, I. D. and Oke, T. R.: Local Climate Zones for Urban Temperature Studies, Bulletin of the American Meteorological Society, 93, 1879–1900, https://doi.org/10.1175/BAMS-D-11-00019.1, 2012.

Sun, Y., Zhang, N., Miao, S., Kong, F., Zhang, Y., and Li, N.: Urban Morphological Parameters of the Main Cities in China and Their Application in the WRF Model, Journal of Advances in Modeling Earth Systems, 13, e2020MS002382, https://doi.org/10.1029/2020MS002382, 2021.

Tabari, H.: Climate change impact on flood and extreme precipitation increases with water availability, Sci Rep, 10, 13768, https://doi.org/10.1038/s41598-020-70816-2, 2020.

Tang, Q., Golaz, J.-C., Van Roekel, L. P., Taylor, M. A., Lin, W., Hillman, B. R., Ullrich, P. A., Bradley, A. M., Guba, O., Wolfe, J. D., Zhou, T., Zhang, K., Zheng, X., Zhang, Y., Zhang, M., Wu, M., Wang, H., Tao, C., Singh, B., Rhoades, A. M., Qin, Y., Li, H.-Y., Feng, Y., Zhang, Y., Zhang, C., Zender, C. S., Xie, S., Roesler, E. L., Roberts, A. F., Mametjanov, A., Maltrud, M. E., Keen, N. D., Jacob, R. L., Jablonowski, C., Hughes, O. K., Forsyth, R. M., Di Vittorio, A. V., Caldwell, P. M., Bisht, G., McCoy, R. B., Leung, L. R., and Bader, D. C.: The fully coupled regionally refined model of E3SM version 2: overview of the atmosphere, land, and river results, Geoscientific Model Development, 16, 3953–3995, https://doi.org/10.5194/gmd-16-3953-2023, 2023a.

Tang, W., Pfister, G. G., Kumar, R., Barth, M., Edwards, D. P., Emmons, L. K., and Tilmes, S.: Capturing High-Resolution Air Pollution Features Using the Multi-Scale Infrastructure for Chemistry and Aerosols Version 0 (MUSICAv0) Global Modeling System, Journal of Geophysical Research: Atmospheres, 128, e2022JD038345, https://doi.org/10.1029/2022JD038345, 2023b.

United Nations: World Urbanization Prospects The 2018 Revision, 2018.

Vardoulakis, S., Fisher, B. E. A., Pericleous, K., and Gonzalez-Flesca, N.: Modelling air quality in street canyons: a review, Atmospheric Environment, 37, 155–182, https://doi.org/10.1016/S1352-2310(02)00857-9, 2003.

Wang, D., Schwartz, P., Yuan, F., Thornton, P., and Zheng, W.: Toward Ultrahigh-Resolution E3SM Land Modeling on Exascale Computers, Computing in Science & Engineering, 24, 44–53, https://doi.org/10.1109/MCSE.2022.3218990, 2022.

Wang, Y., Sun, G., Wu, Y., and Rosenberg, M. W.: Urban 3D building morphology and energy consumption: empirical evidence from 53 cities in China, Sci Rep, 14, 12887, https://doi.org/10.1038/s41598-024-63698-1, 2024.

van der Wiel, K. and Bintanja, R.: Contribution of climatic changes in mean and variability to monthly temperature and precipitation extremes, Commun Earth Environ, 2, 1–11, https://doi.org/10.1038/s43247-020-00077-4, 2021.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Wu, S., Lin, X., Bian, Z., Lipson, M., Lafortezza, R., Liu, Q., Grimmond, S., Velasco, E., Christen, A., Masson, V., Crawford, B., Ward, H. C., Chrysoulakis, N., Fortuniak, K., Parlow, E., Pawlak, W., Tapper, N., Hong, J., Hong, J.-W., Roth, M., An, J., Lin, C., and Chen, B.: Satellite observations reveal a decreasing albedo trend of global cities over the past 35 years, Remote Sensing of Environment, 303, 114003, https://doi.org/10.1016/j.rse.2024.114003, 2024.

Yang, J., Zhao, L., and Oleson, K.: Large humidity effects on urban heat exposure and cooling challenges under climate change, Environ. Res. Lett., 18, 044024, https://doi.org/10.1088/1748-9326/acc475, 2023.

Yuan, F., Wang, D., Kao, S.-C., Thornton, M., Ricciuto, D., Salmon, V., Iversen, C., Schwartz, P., and Thornton, P.: An ultrahigh-resolution E3SM land model simulation framework and its first application to the Seward Peninsula in Alaska, Journal of Computational Science, 73, 102145, https://doi.org/10.1016/j.jocs.2023.102145, 2023.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.-E., Xu, P., Ramoino, F., and Arino, O.: ESA WorldCover 10 m 2021 v200 (v200), https://doi.org/10.5281/zenodo.7254221, 2022.

Zhan, Y., Yao, Z., Groffman, P. M., Xie, J., Wang, Y., Li, G., Zheng, X., and Butterbach-Bahl, K.: Urbanization can accelerate climate change by increasing soil N2O emission while reducing CH4 uptake, Global Change Biology, 29, 3489–3502, https://doi.org/10.1111/gcb.16652, 2023.

Zhang, K., Cao, C., Chu, H., Zhao, L., Zhao, J., and Lee, X.: Increased heat risk in wet climate induced by urban humid heat, Nature, 617, 738–742, https://doi.org/10.1038/s41586-023-05911-1, 2023a.

Zhang, W., Cui, R., Li, C., Ge, H., Zhang, Z., and Tang, X.: Impact of urban agglomeration construction on urban air quality–empirical test based on PSM–DID model, Sci Rep, 13, 15099, https://doi.org/10.1038/s41598-023-42314-8, 2023b.

Zhang, Y. and Gu, Z.: Air quality by urban design, Nature Geosci, 6, 506–506, https://doi.org/10.1038/ngeo1869, 2013.

Zhang, Z., Qian, Z., Zhong, T., Chen, M., Zhang, K., Yang, Y., Zhu, R., Zhang, F., Zhang, H., Zhou, F., Yu, J., Zhang, B., Lü, G., and Yan, J.: Vectorized rooftop area data for 90 cities in China, Sci Data, 9, 66, https://doi.org/10.1038/s41597-022-01168-x, 2022.

Zhao, L.: Urban growth and climate adaptation, Nature Clim Change, 8, 1034–1034, https://doi.org/10.1038/s41558-018-0348-x, 2018.

Zhao, L., Lee, X., Smith, R. B., and Oleson, K.: Strong contributions of local background climate to urban heat islands, Nature, 511, 216–219, https://doi.org/10.1038/nature13462, 2014.

Zhao, L., Lee, X., and Schultz, N. M.: A wedge strategy for mitigation of urban warming in future climate scenarios, Atmospheric Chemistry and Physics, 17, 9067–9080, https://doi.org/10.5194/acp-17-9067-2017, 2017.

Zhao, L., Oppenheimer, M., Zhu, Q., Baldwin, J. W., Ebi, K. L., Bou-Zeid, E., Guan, K., and Liu, X.: Interactions between urban heat islands and heat waves, Environ. Res. Lett., 13, 034003, https://doi.org/10.1088/1748-9326/aa9f73, 2018.

Zhao, L., Oleson, K., Bou-Zeid, E., Krayenhoff, E. S., Bray, A., Zhu, Q., Zheng, Z., Chen, C., and Oppenheimer, M.: Global multi-model projections of local urban climates, Nat. Clim. Chang., 11, 152–157, https://doi.org/10.1038/s41558-020-00958-8, 2021.

https://doi.org/10.5194/essd-2024-416
Preprint. Discussion started: 1 October 2024

Zhao, M., Cheng, C., Zhou, Y., Li, X., Shen, S., and Song, C.: A global dataset of annual urban extents (1992–2020) from harmonized nighttime lights, Earth System Science Data, 14, 517–534, https://doi.org/10.5194/essd-14-517-2022, 2022.

Zheng, Z., Zhao, L., and Oleson, K. W.: Large model structural uncertainty in global projections of urban heat waves, Nat Commun, 12, 3736, https://doi.org/10.1038/s41467-021-24113-9, 2021.

Zhou, Y., Smith, S. J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., Asrar, G. R., Zhang, X., He, C., and Elvidge, C. D.: A global map of urban extent from nightlights, Environ. Res. Lett., 10, 054011, https://doi.org/10.1088/1748-9326/10/5/054011, 2015.