Earth System
Science
Data

Open Access

Supplement of

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

Yifan Cheng et al.

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

Section S1: Aggregation of 1km raw U-Surf data to coarser resolutions
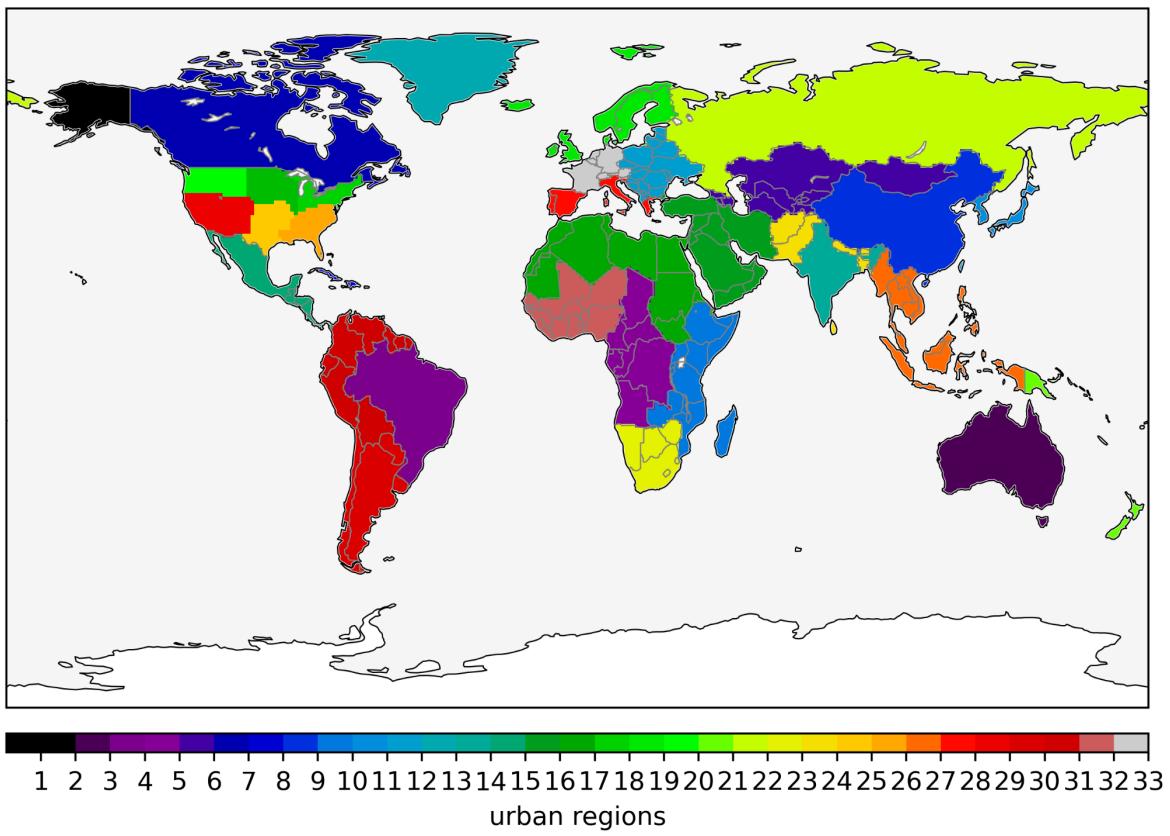
The aggregation of urban canopy parameters (UCPs) from 1km to coarser resolutions requires careful consideration of the physical properties and conservation principles of different parameters. As shown in Table S1, we have classified UCPs into two categories, area-based conservative and non-conservative. We employed direct spatial averaging for urban percentage. For conservative parameters – roof and pervious fraction – which are inherently weighted by area, we used their urban percentages as weights to aggregate. We implemented a facet-fraction weighted averaging method for all the non-conservative parameters to ensure physically meaningful aggregation. For example, when aggregating roof or impervious canyon floor emissivity, we used the respective facet areal fractions (roof or impervious canyon floor fraction) with respect to the 1 km grid as their weights. This way ensures that the contribution of each parameter to the coarser resolution is proportional to its actual surface coverage.

The aggregation of canyon height-to-width ratio (H/W) is slightly more complex as it is derived from multiple primary parameters. We evaluated two potential aggregation methods: 'aggregating first' and 'aggregating after' (Li et al., 2024a), both using urban density (urban percentage × roof fraction) as weights. The former is to aggregate the primary input parameters (e.g., building height, roof fraction) to the target resolution before calculating H/W. The latter calculates H/W at the original 1km resolution before spatial aggregation. Our analysis revealed that the 'aggregating after' method generally produces slightly higher values and preserves more spatial variation compared to the 'aggregating first' method (Figure S27). In addition, the 'aggregating after' method better maintains the non-linear relationships between input and output parameters and hence preserves local canyon characteristics during the upscaling process. This choice aligns with the recommendations from previous studies (e.g., (Dai et al., 2019; Shangguan et al., 2014)) and helps prevent the smoothing of local variations in the 'aggregating first' method. Therefore, in the published dataset with this study we choose the 'aggregating after' method to aggregate H/W to coarser resolutions (0.125° and 1°).
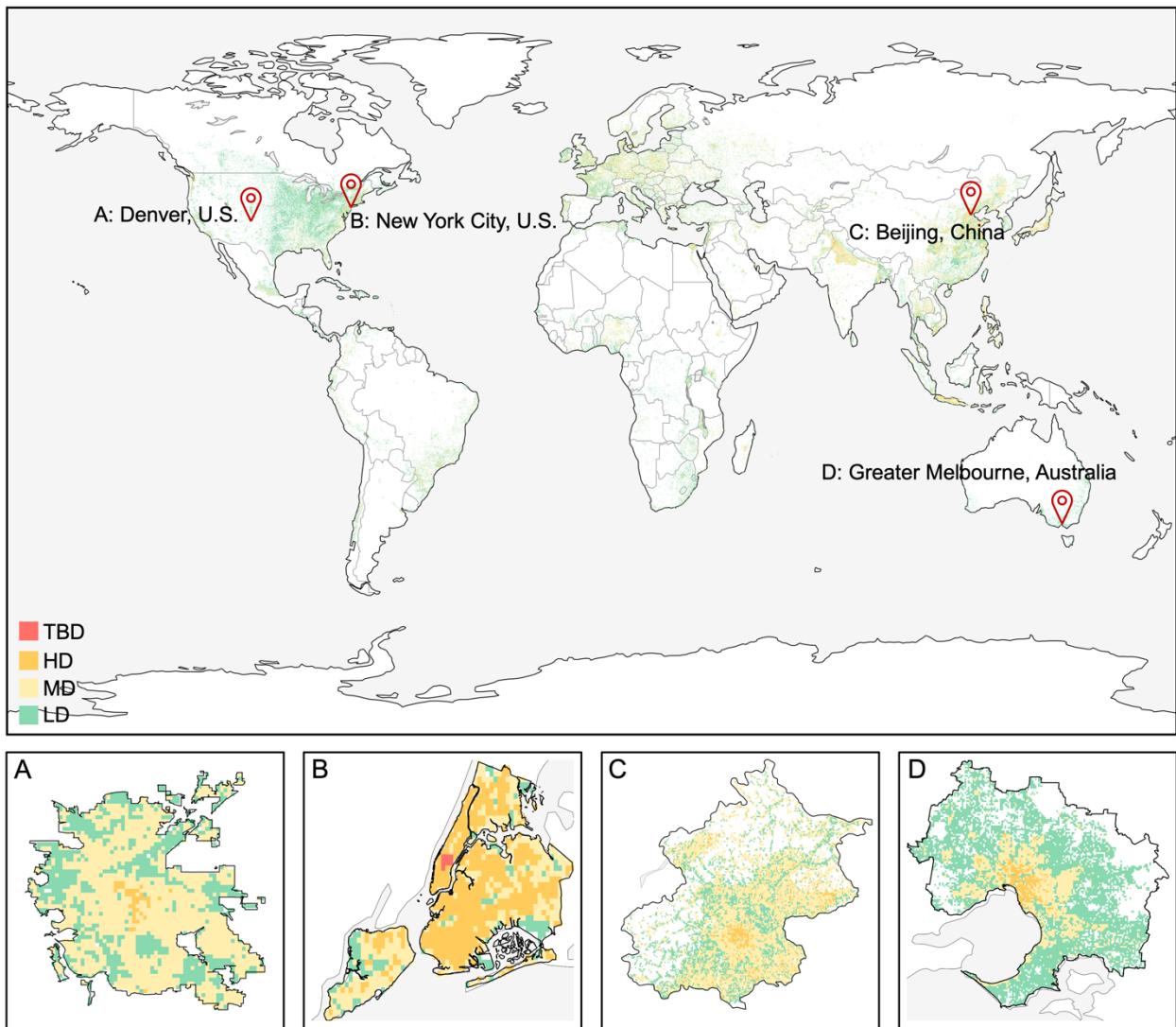
Section S2: Uncertainty propagation in data synthesis and processing

In our uncertainty assessment, we employed Monte Carlo simulation approach that assumes the uncertainties from different data products are independent. For each simulation, we introduced normally distributed perturbations based on the documented uncertainties of individual data sources (Table 3) to evaluate how these variations affect our 1km output parameters. Most input datasets provided only RMSE values as their uncertainty metric, thus we adopted a conservative approach by approximating the standard deviation with RMSE ($RMSE^2 = bias^2 + variance$, thus $variance \leq RMSE$) in these cases, thereby ensuring our uncertainty estimates remain conservative.
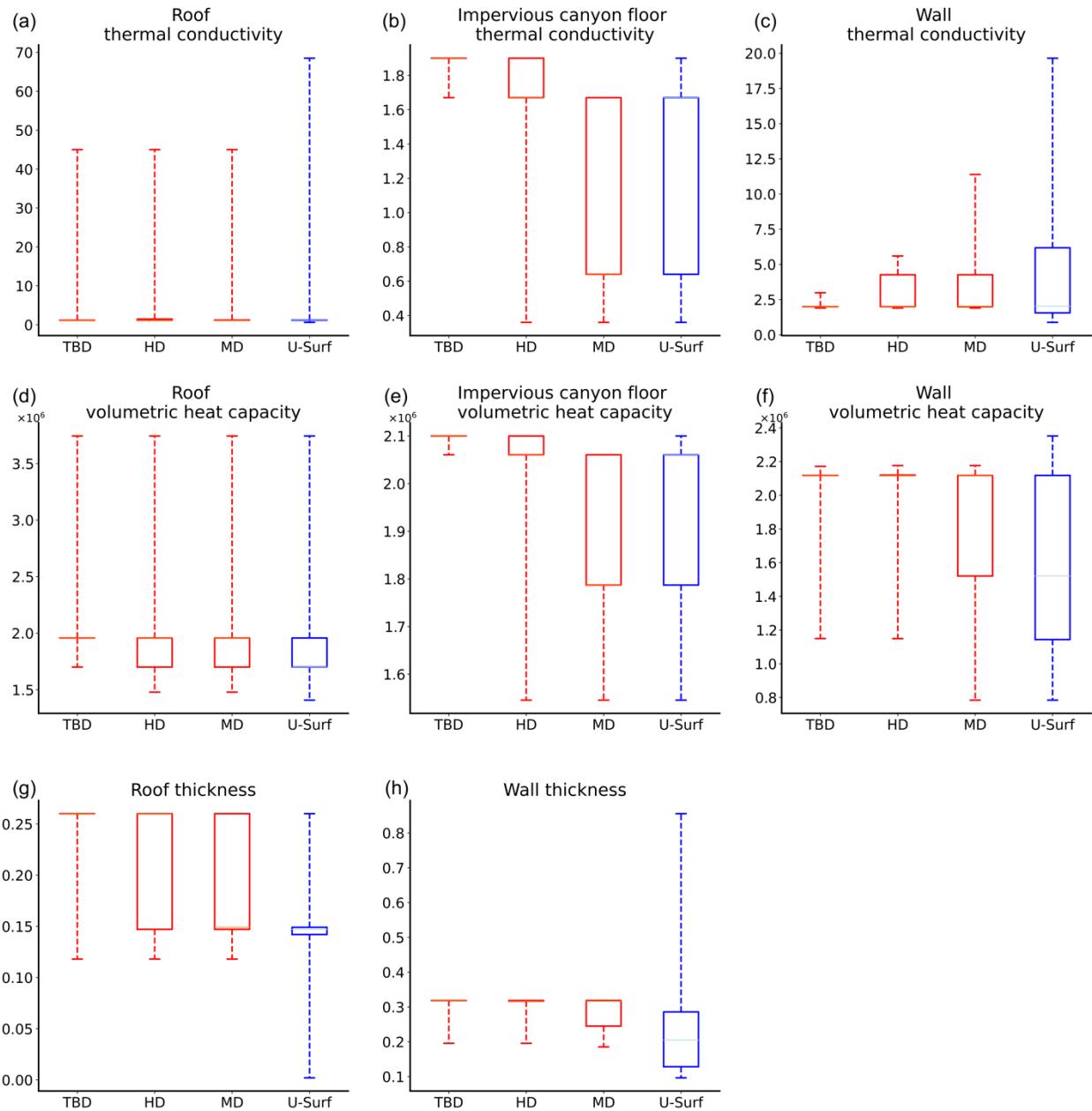
For most input datasets, we could directly obtain uncertainty values from original literature or calculate them through simple averaging. However, the uncertainty estimation for the Sentinel-2 blue-sky albedo product required additional steps. We estimate the uncertainty of the Sentinel-2 blue-sky albedo by examining the RMSE of black-sky albedo α_{black} and white-sky albedo α_{white}. According to Lin et al. (2022), the average uncertainty of the unevenly and uniformly distributed urban areas gives $\sigma_{black} = 0.0185$ and $\sigma_{white} = 0.0205$, respectively. We then calculated the blue-sky albedo α_{blue} as $(1 - D)\alpha_{black} + D\alpha_{white}$, where D is the diffuse skylight ratio and is assigned the commonly used value of 0.3 here based on the BaRAD2019 dataset from (Chakraborty and Lee, 2021). Thus, $\sigma_{black} = \sqrt{(1 - D)^2 {\alpha_{black}}^2 + D^2 {\alpha_{white}}^2} = 0.0154$.
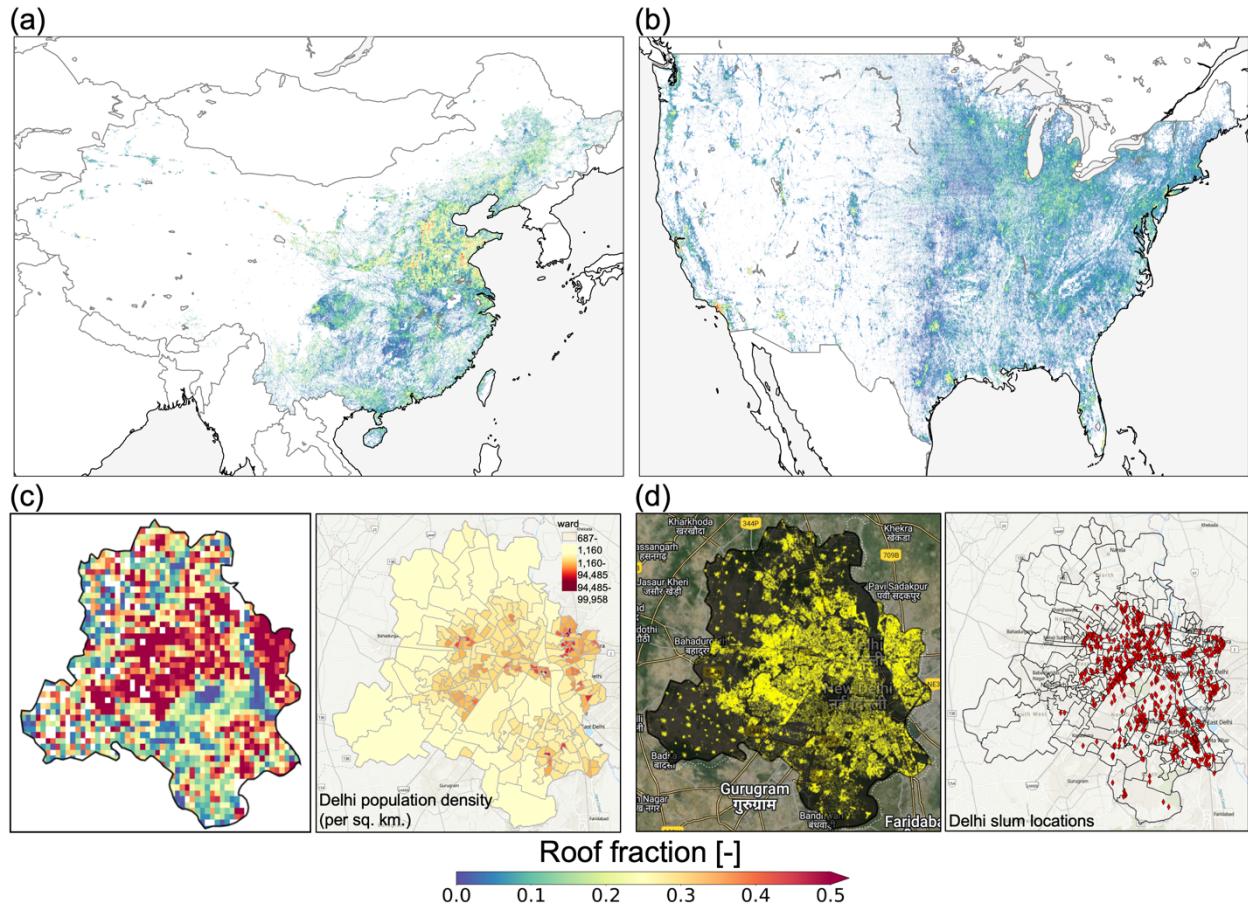
Figure S1. The thirty-three regions used to constrain the global variability of urban surface properties (Jackson et al., 2010).
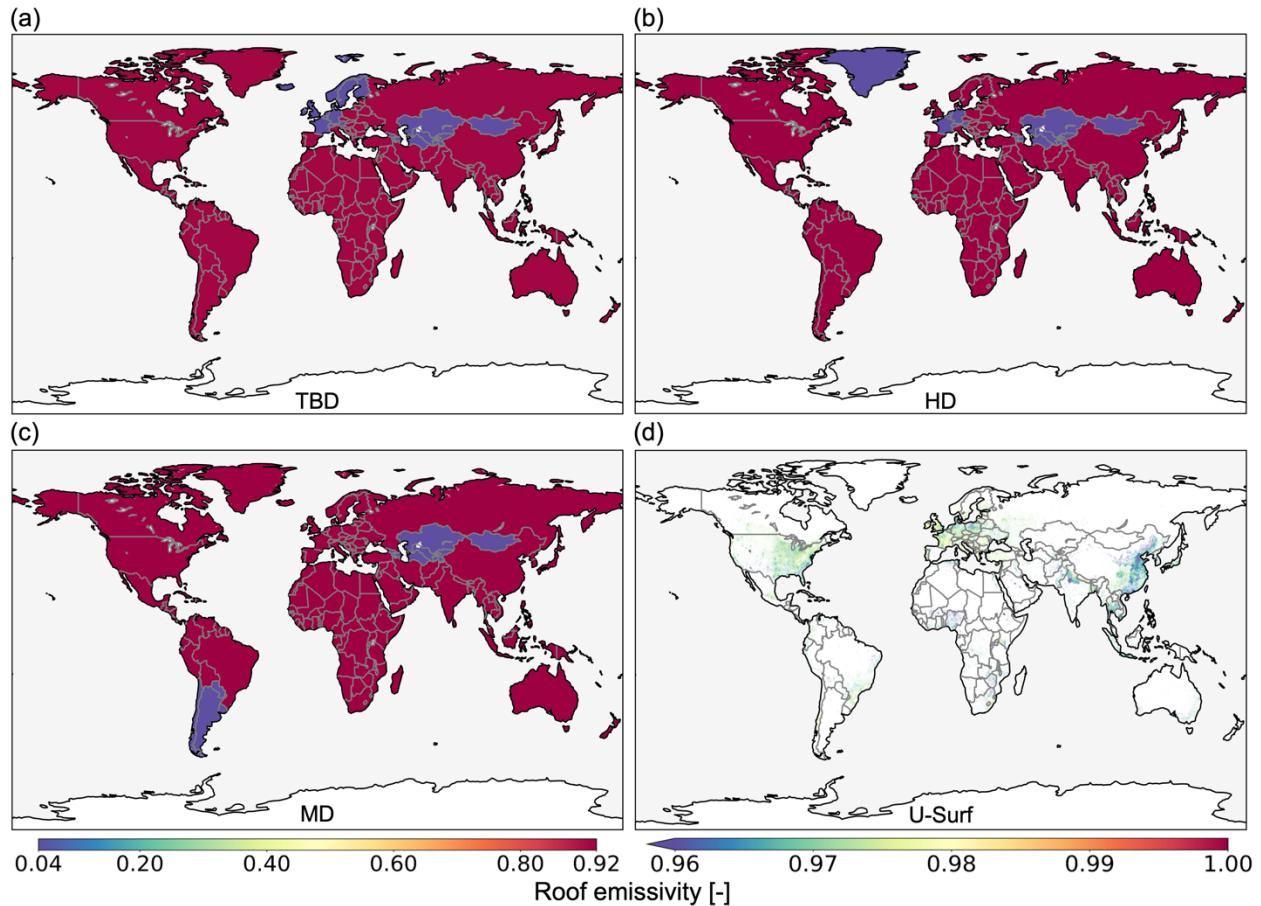
Figure S2. The four urban density classes used to attribute thermal parameters in U-Surf: TBD (red), HD (orange), MD (yellow), LD (green), based on the canyon height-to-width ratio percentiles from J2010 for the purpose of assigning thermal parameters as described in section 2.2.3 of the main text. The lower plots illustrate this classification in four selected cities: (A) Denver, U.S., (B) New York City, U.S., (C) Beijing, China, (D) Greater Melbourne, Australia.
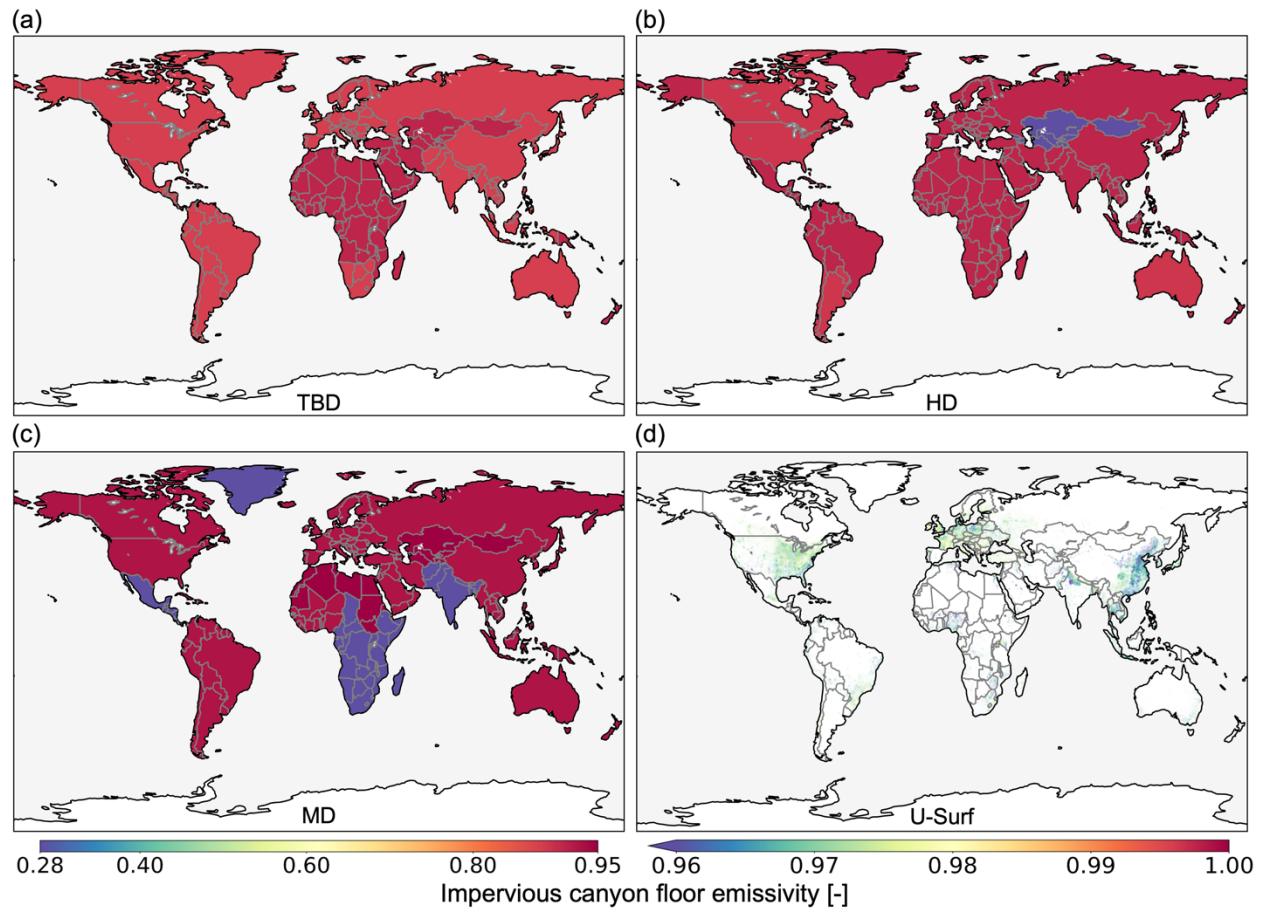
Figure S3. Distributions of thermal properties. Red bars represent default CLMU values (discrete, 33 regions, 3 density classes), and blue bars show the raw new U-Surf values (continuous, 1km). Values of thermal properties are derived based on Oleson and Feddema, (2020).
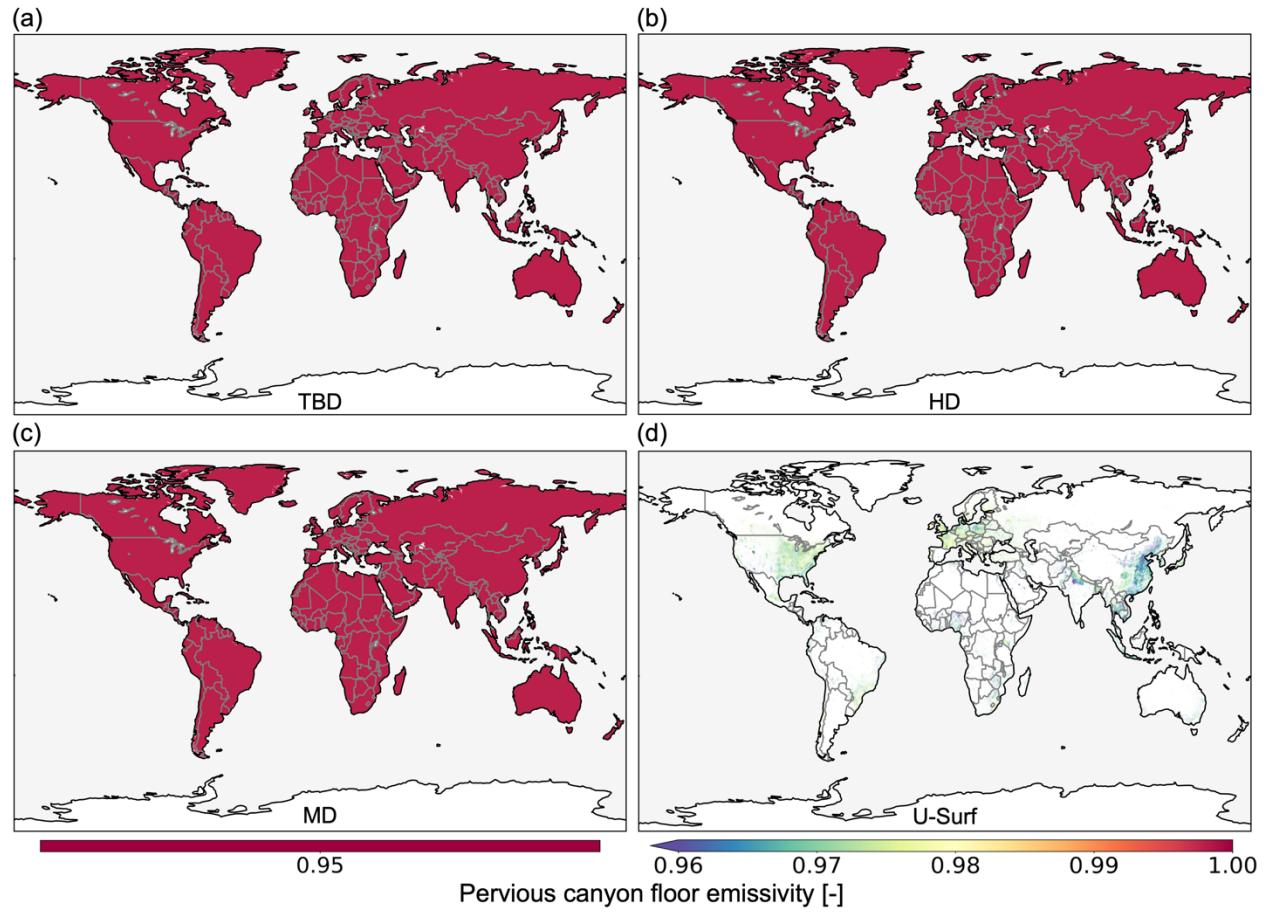
Figure S4. Roof fraction at 1 km-resolution in: (a) China, (b) CONUS, (c) Delhi, India – roof fraction (left) and population density (2022) (right; data source: Esri India, 2024), (d) Delhi, India—building footprints used to derive roof fraction (left) and slum locations (2011) (right; data source: OpenCity - Urban Data Portal, 2024; Delhi Public Geoportal, 2024). All roof fraction maps share the same color scale. The satellite images are from Google Earth Engine (Gorelick et al., 2017) and © Esri.
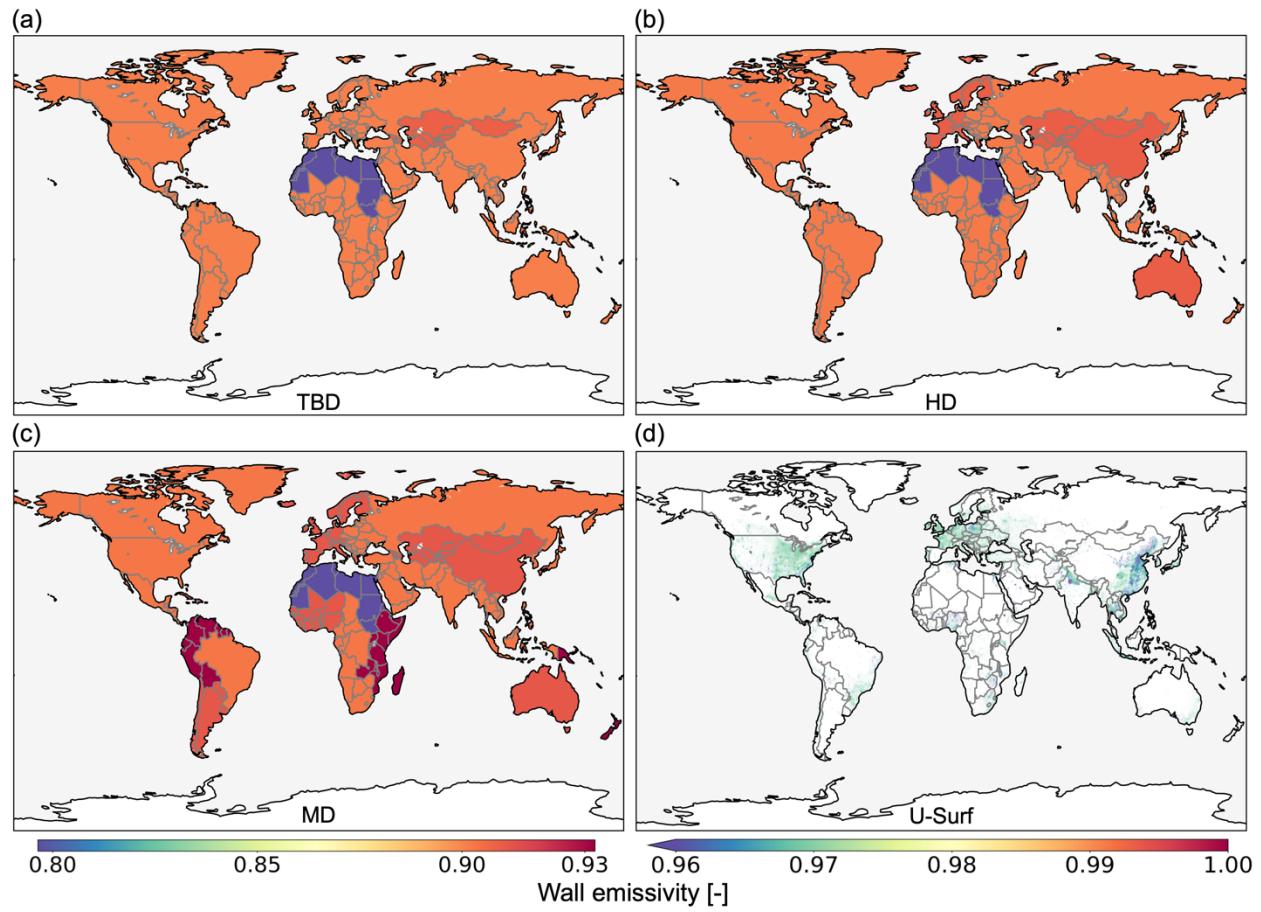
Figure S5. Comparison between CLMU, including (a) Tall Building District (TBD), (b) High Density (HD), (c) Medium Density (MD) (discrete values), and (d) U-Surf roof emissivity (1km resolution) at global scale. Note that (a-c) CLMU parameters share a common colorbar at the bottom left and (d) U-Surf parameter uses a separate one (bottom right) for visualization purposes.
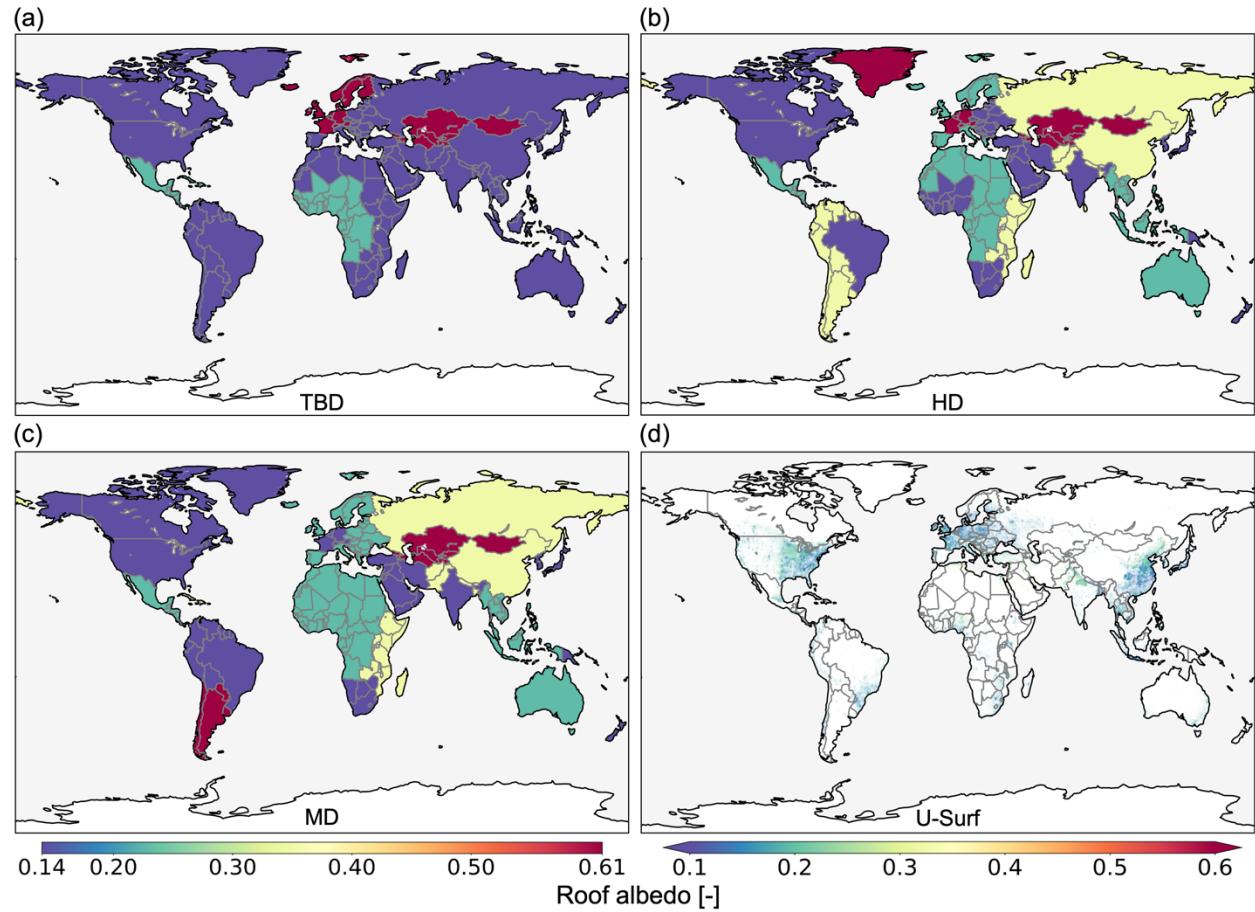
**Figure S6. Same as S5, but for impervious canyon floor emissivity.**
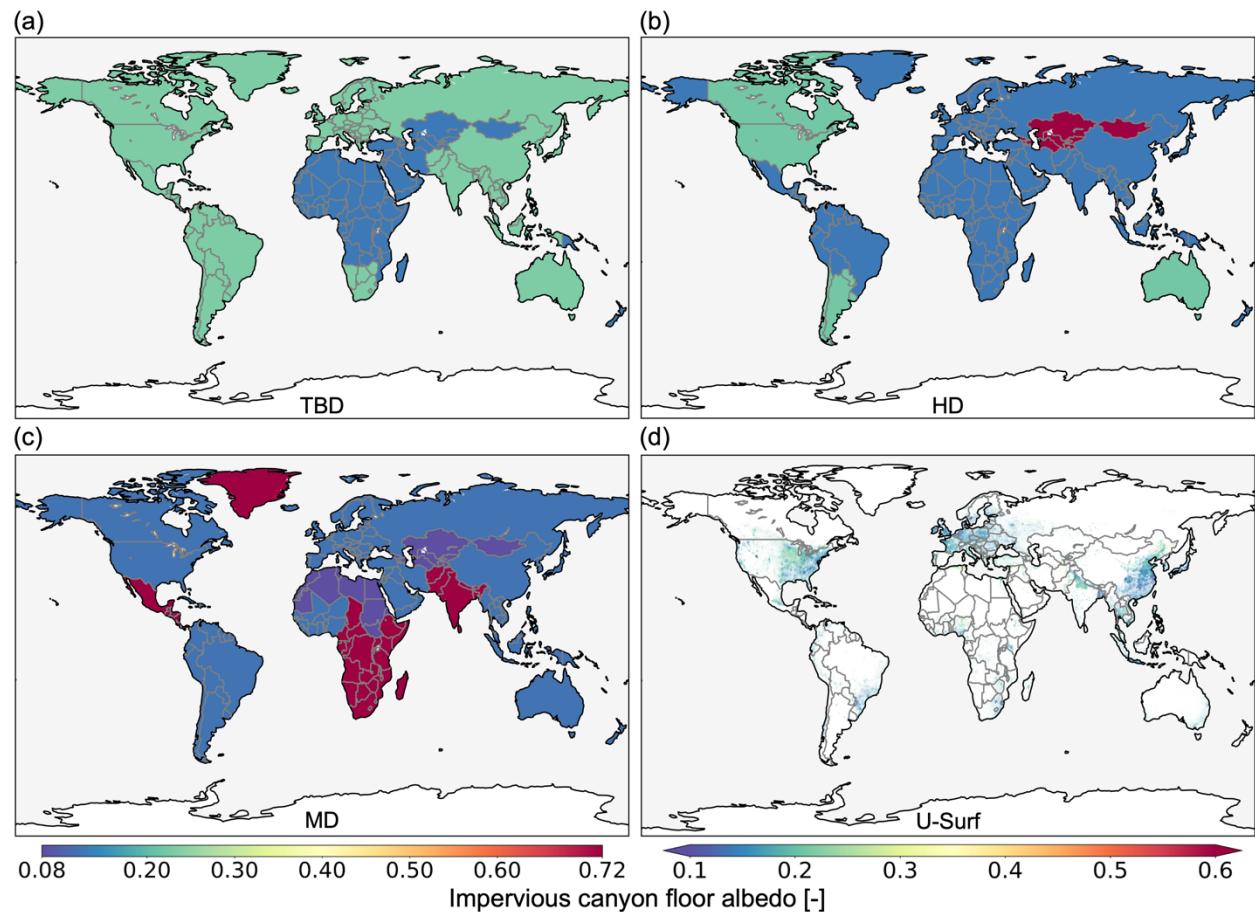
Figure S7. Same as S5, but for pervious canyon floor emissivity. Note that the emissivity is set to be 0.95 everywhere to represent a typical value for vegetation using a bulk parameterization scheme (Oleson et al., 2010).
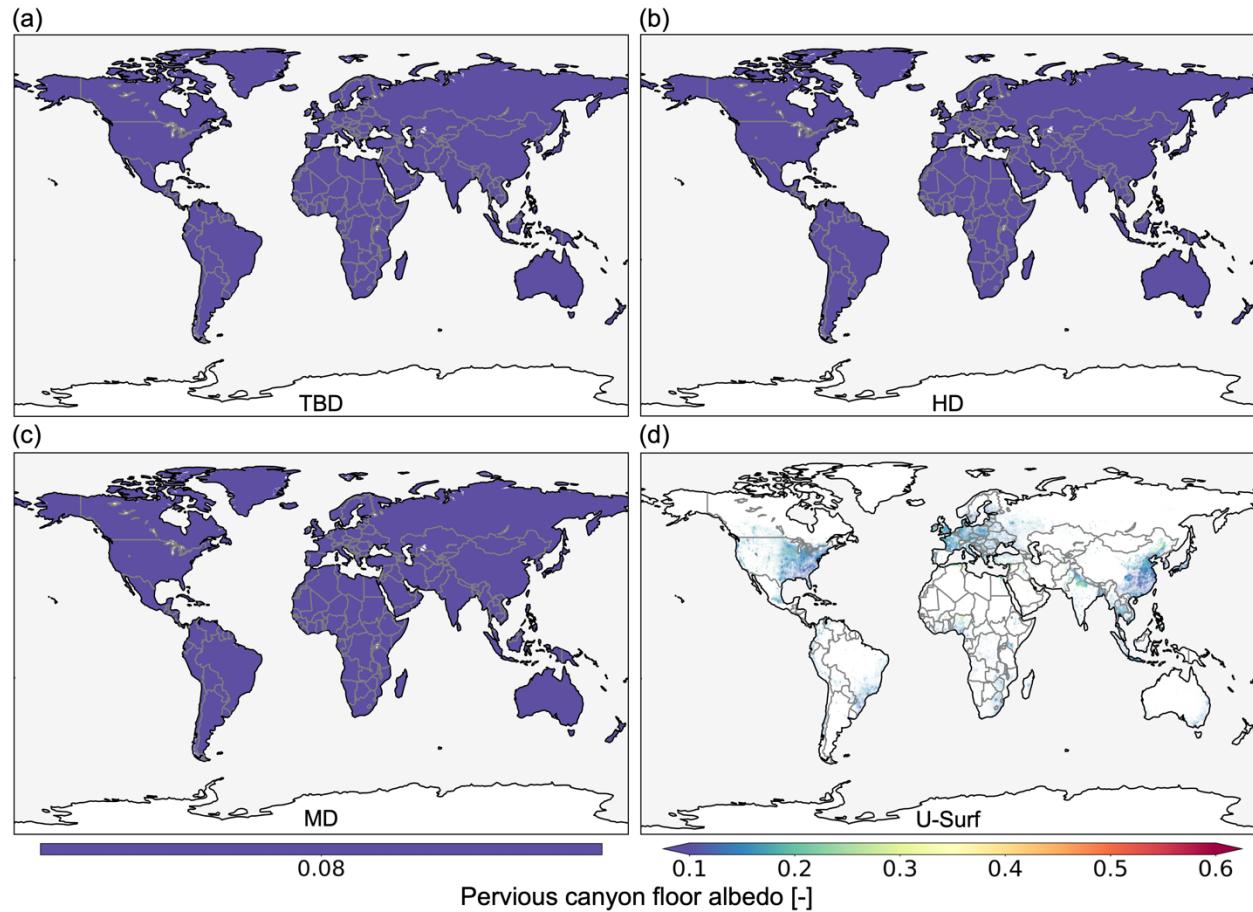
**Figure S8. Same as S5, but for wall emissivity.**
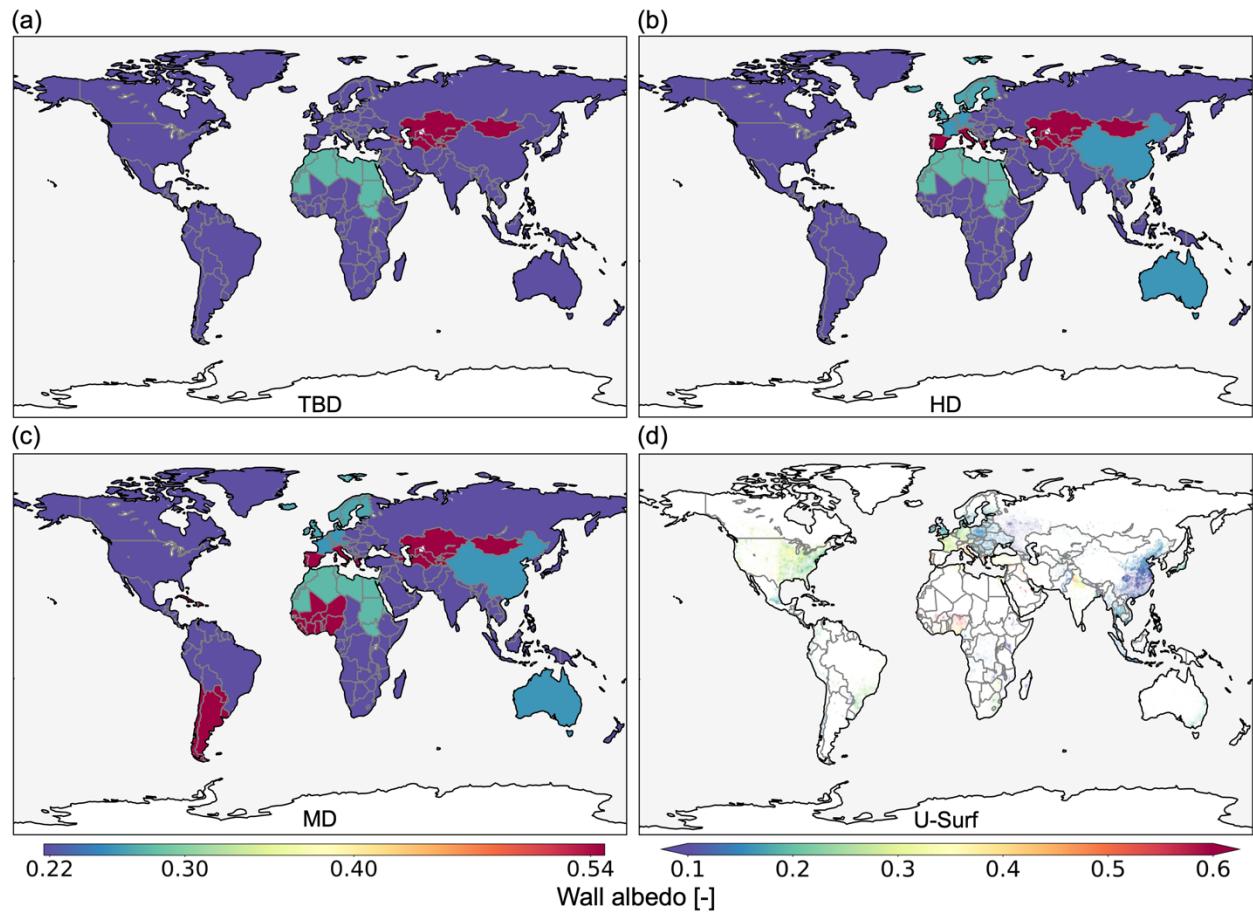
**Figure S9. Same as S5, but for roof albedo.**
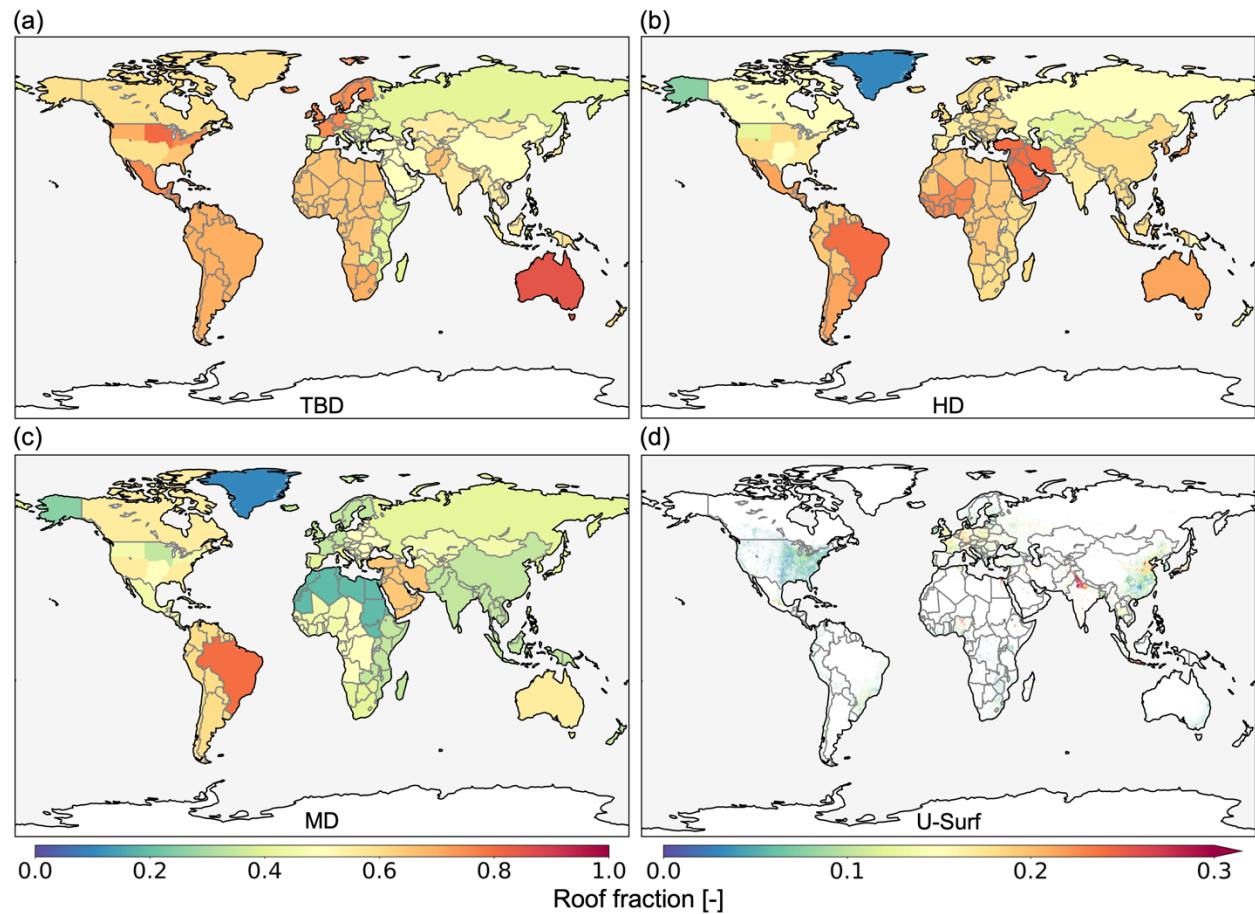
**Figure S10. Same as S5, but for impervious canyon floor albedo.**
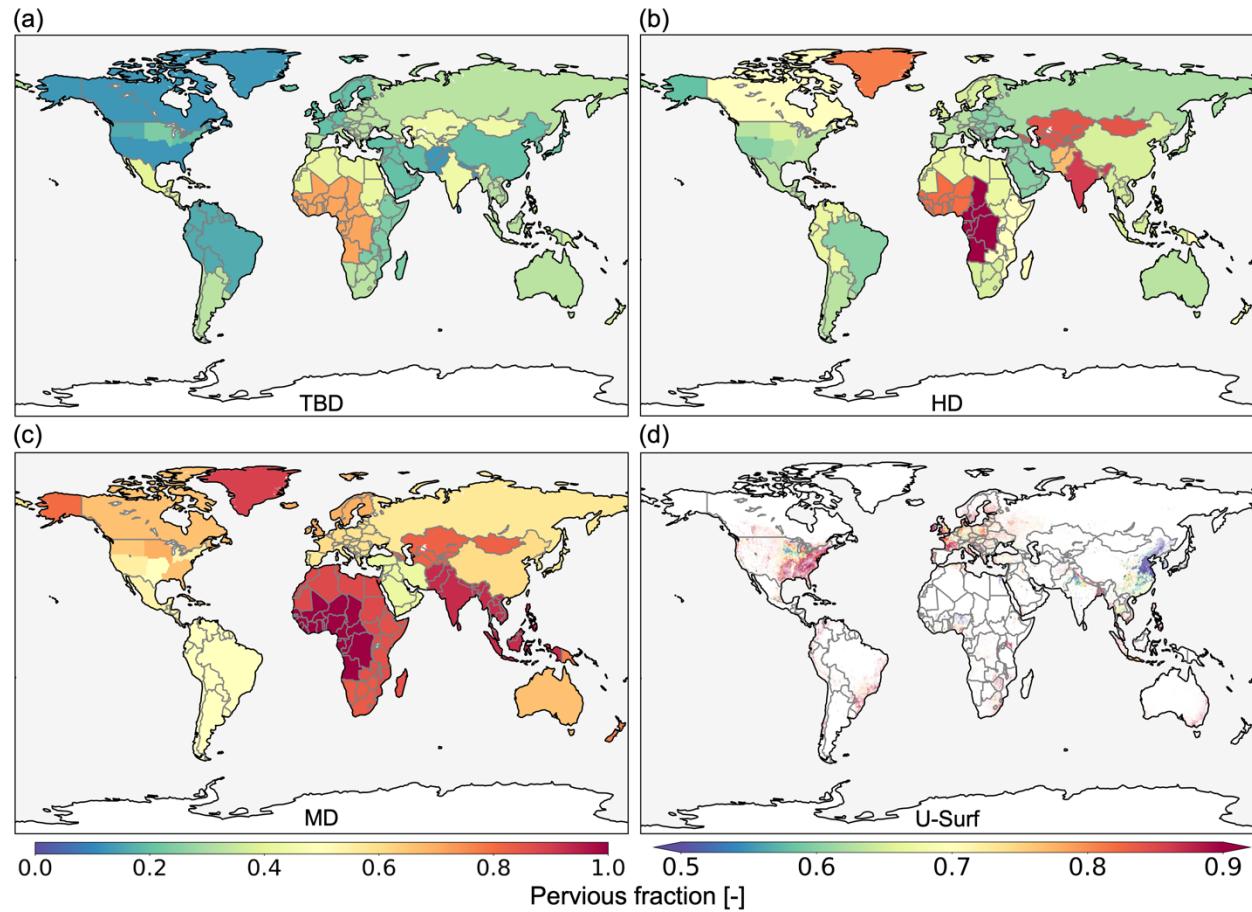
Figure S11. Same as S5, but for pervious canyon floor albedo. Note that the albedo is set to be 0.08 everywhere to represent a typical value for vegetation using a bulk parameterization scheme (Oleson et al., 2010).
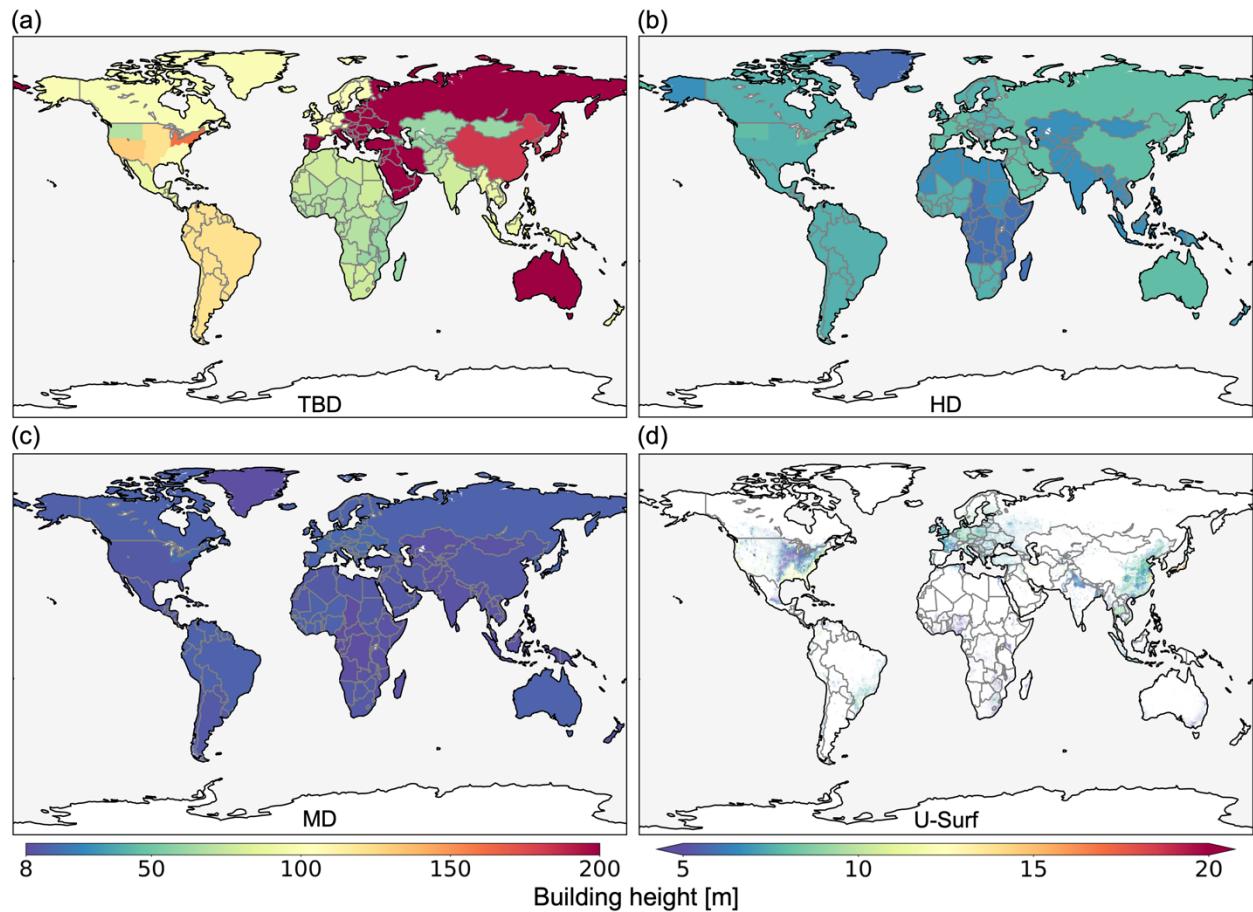
**Figure S12. Same as S5, but for wall albedo.**
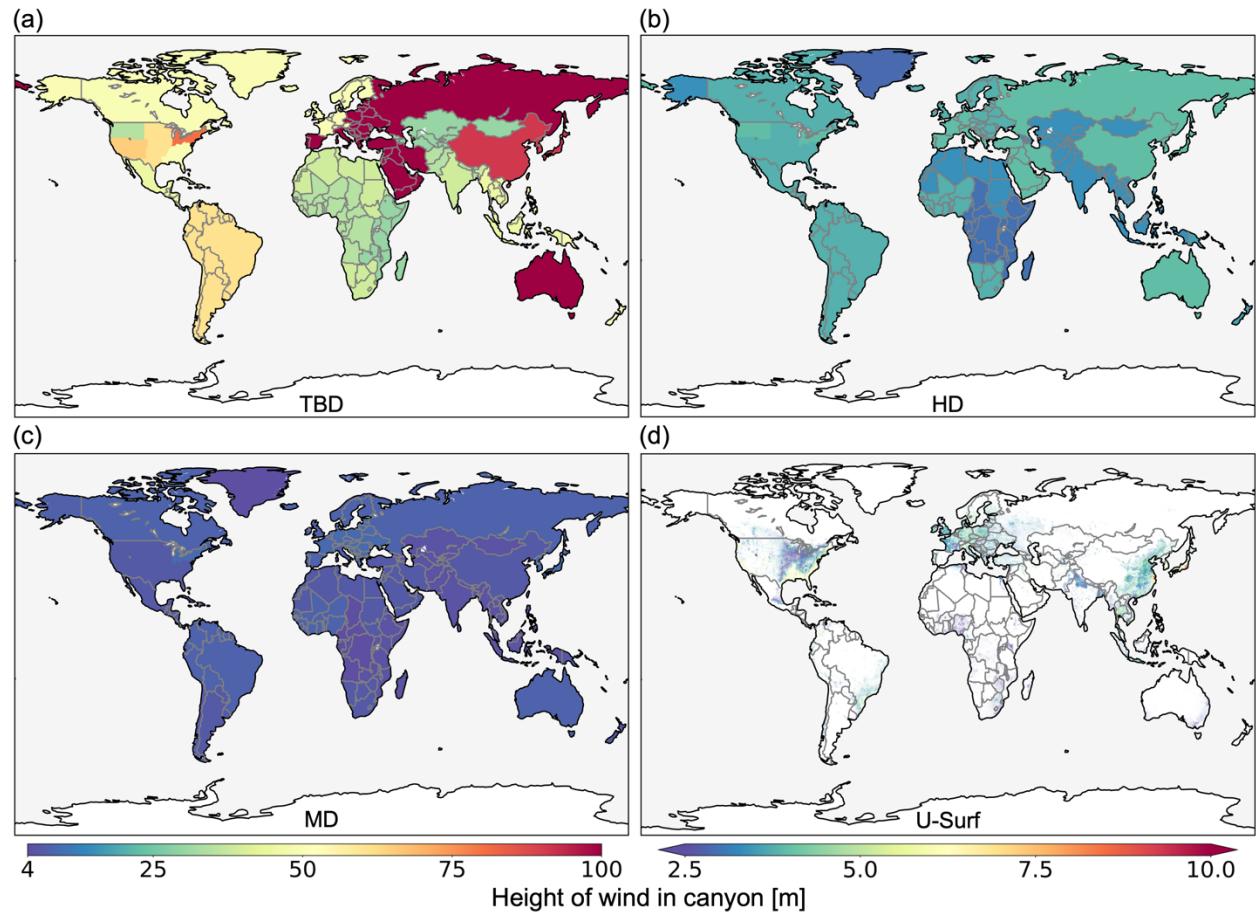
Figure S13. Same as S5, but for roof fraction.
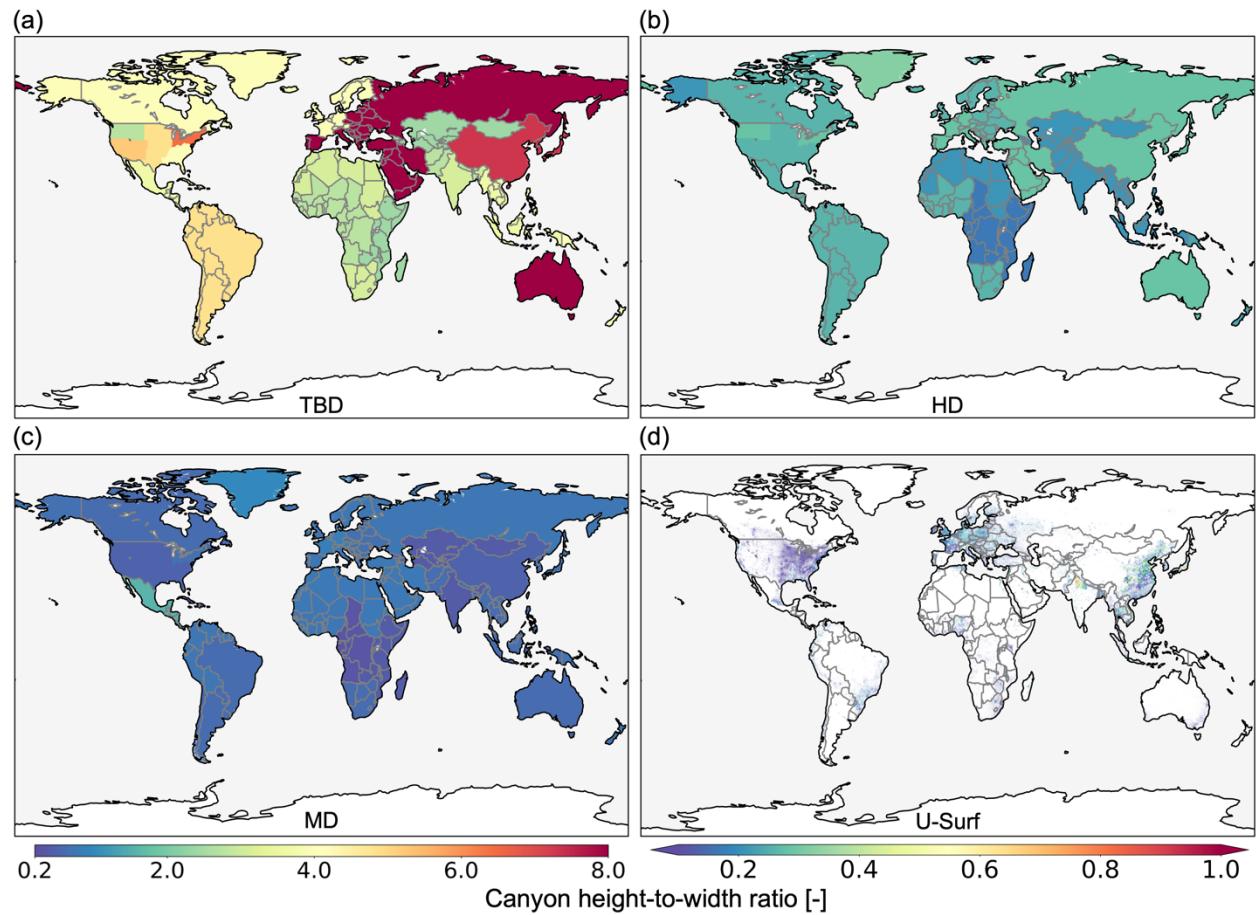
**Figure S14. Same as S5, but for pervious fraction of the canyon floor.**
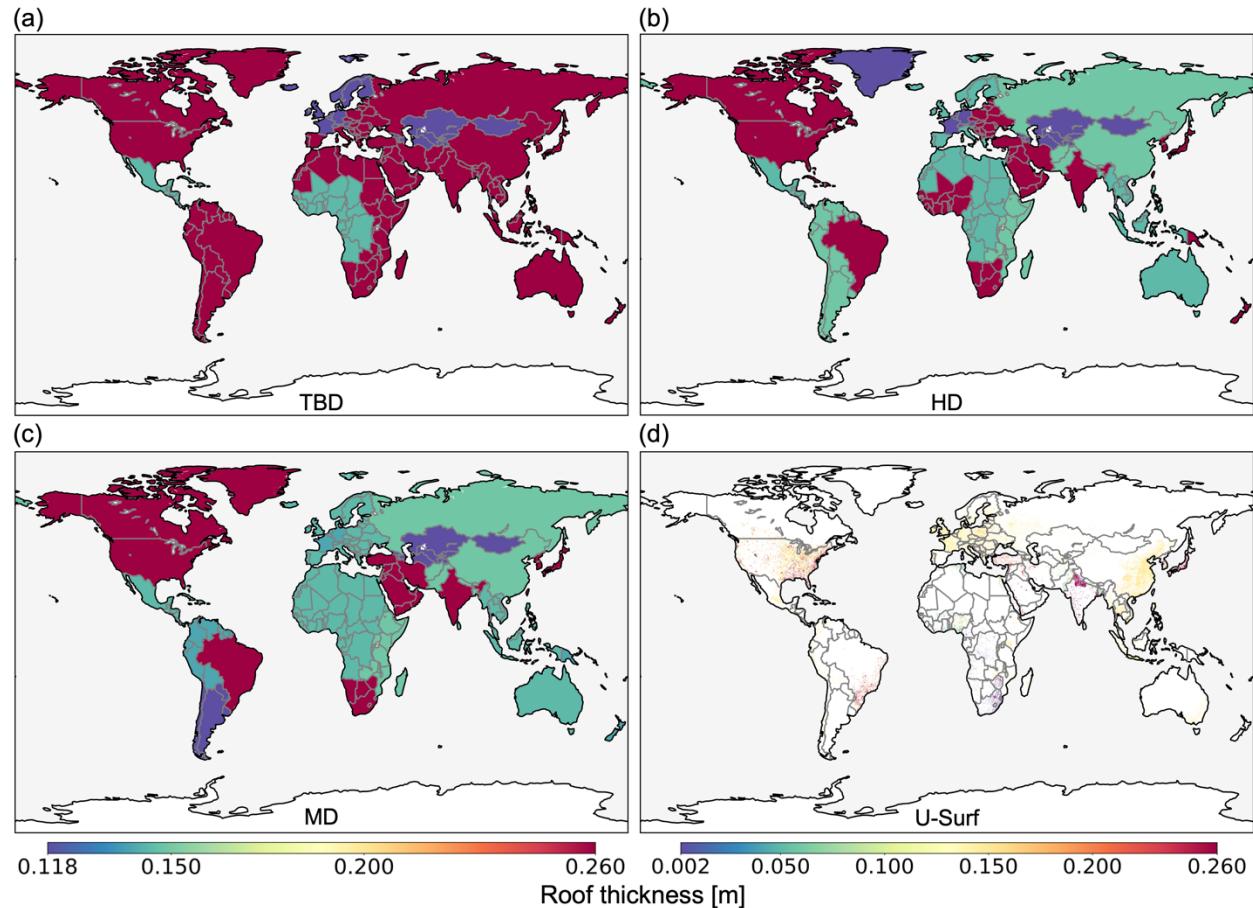
**Figure S15. Same as S5, but for roof height [meter].**
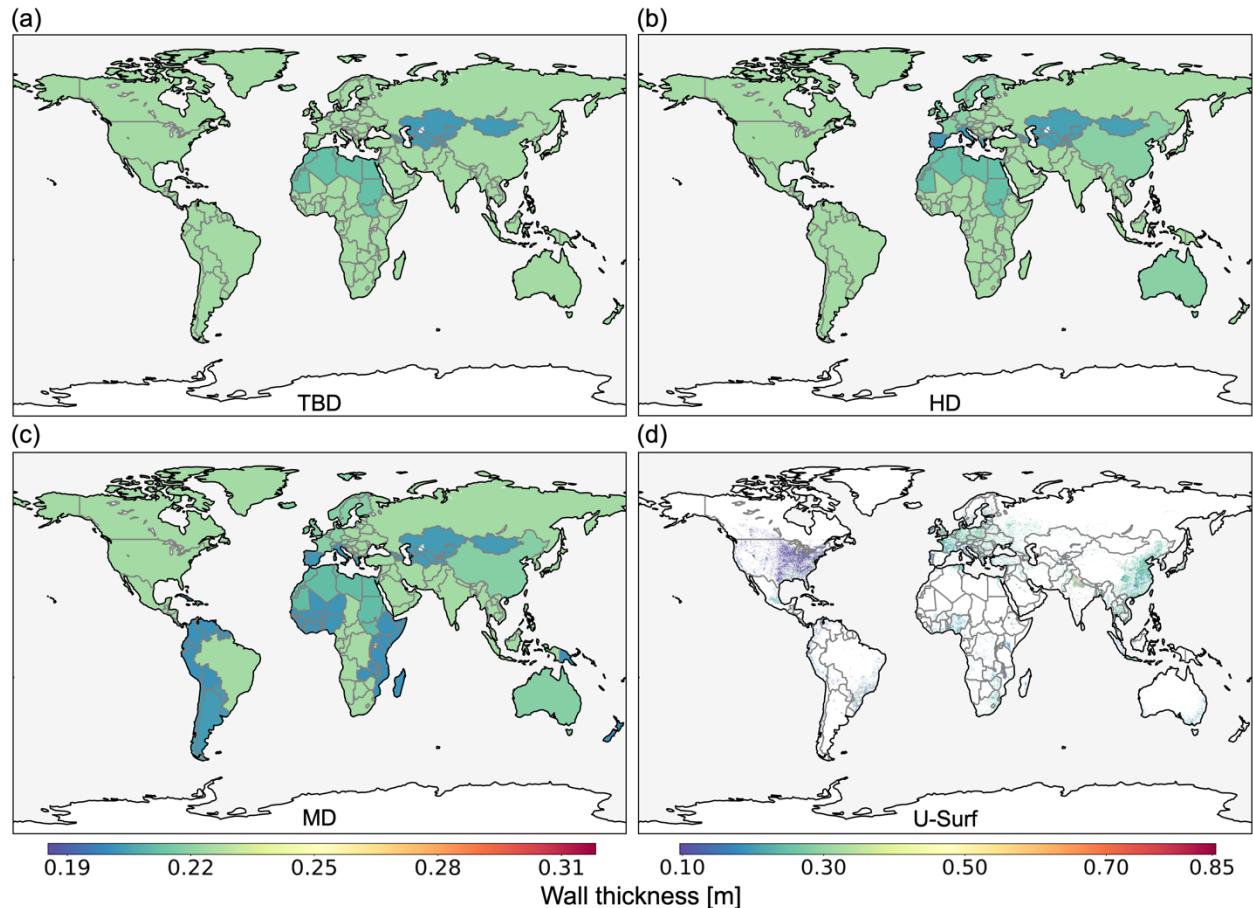
Figure S16. Same as S5, but for height of wind in canyon [meter].
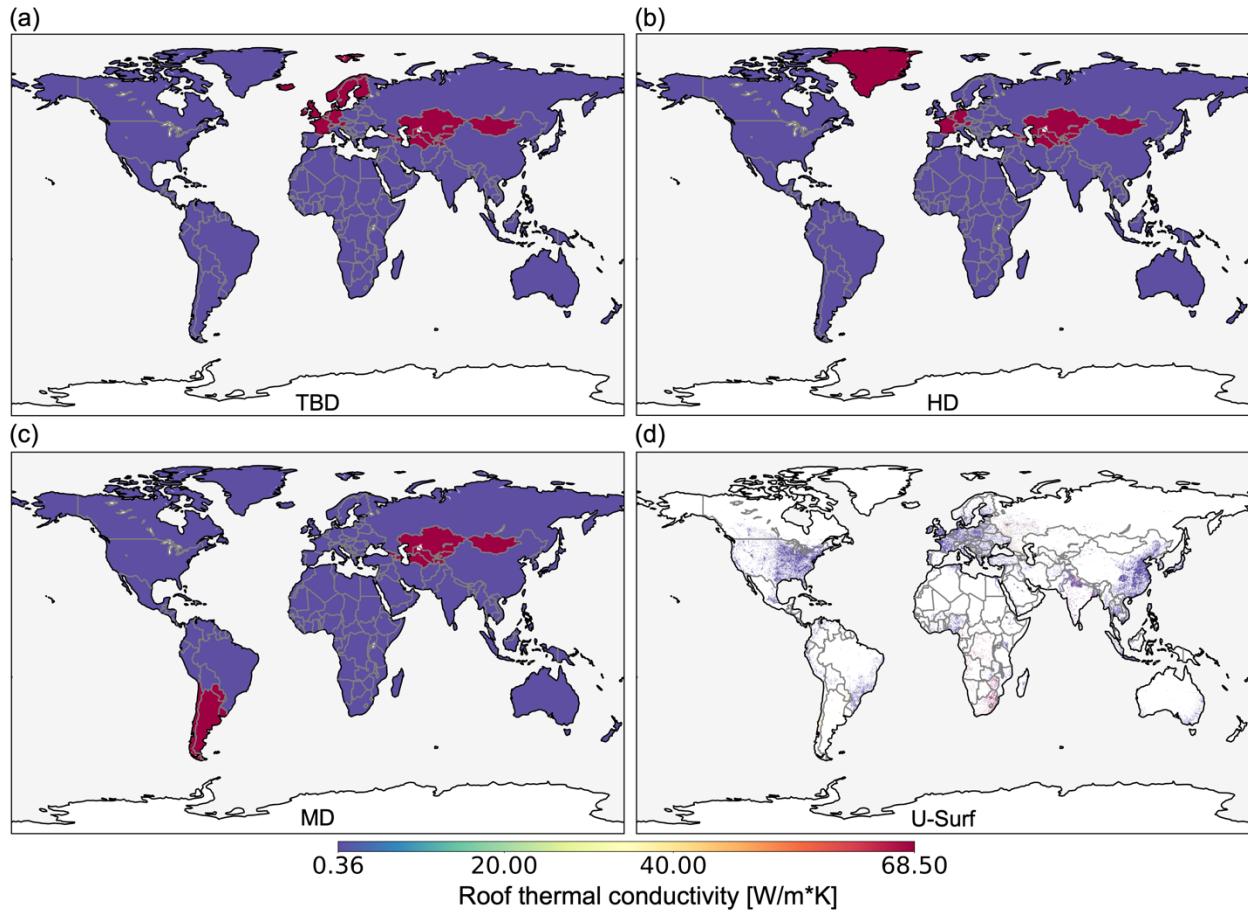
Figure S17. Same as S5, but for canyon height-to-width ratio.
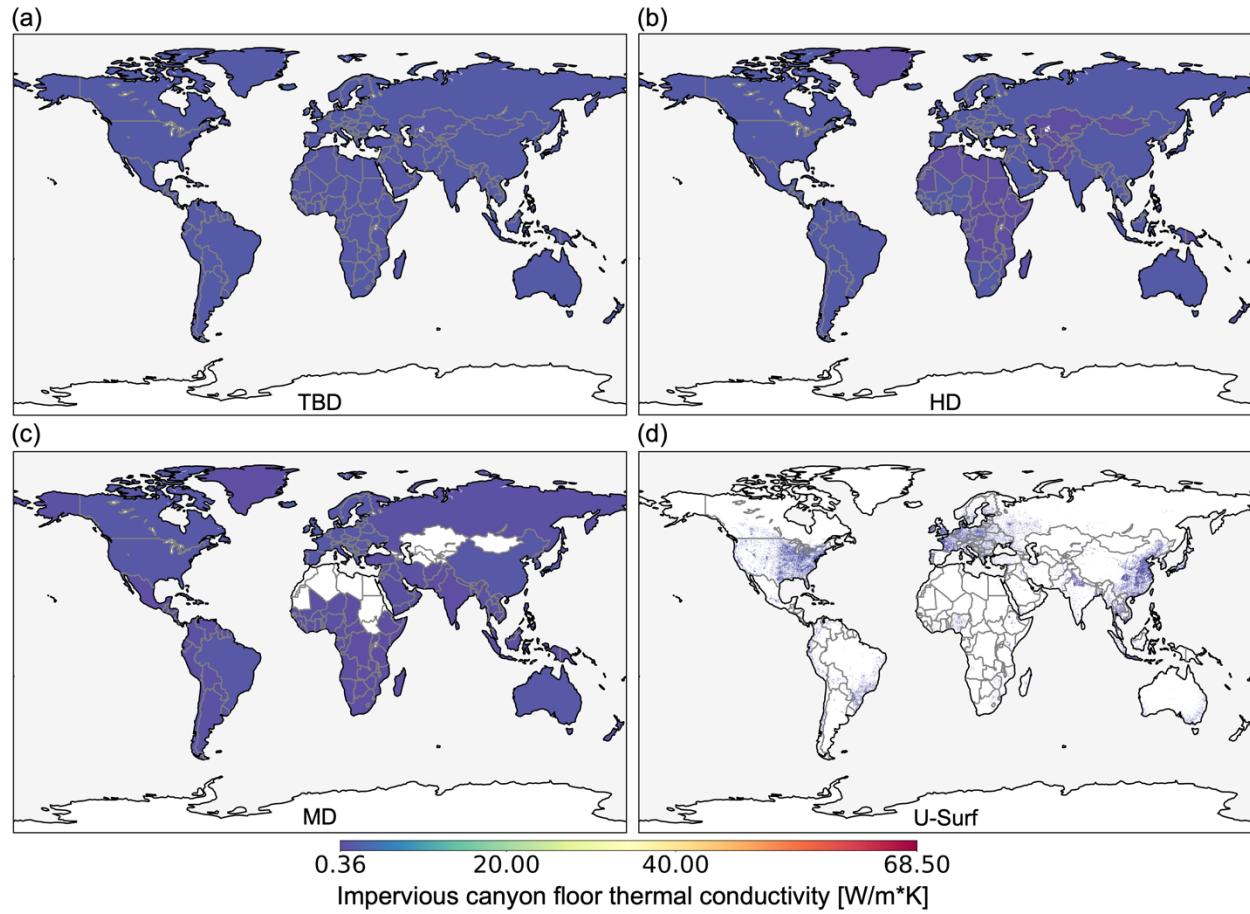
Figure S18. Same as S5, but for roof thickness [meter].
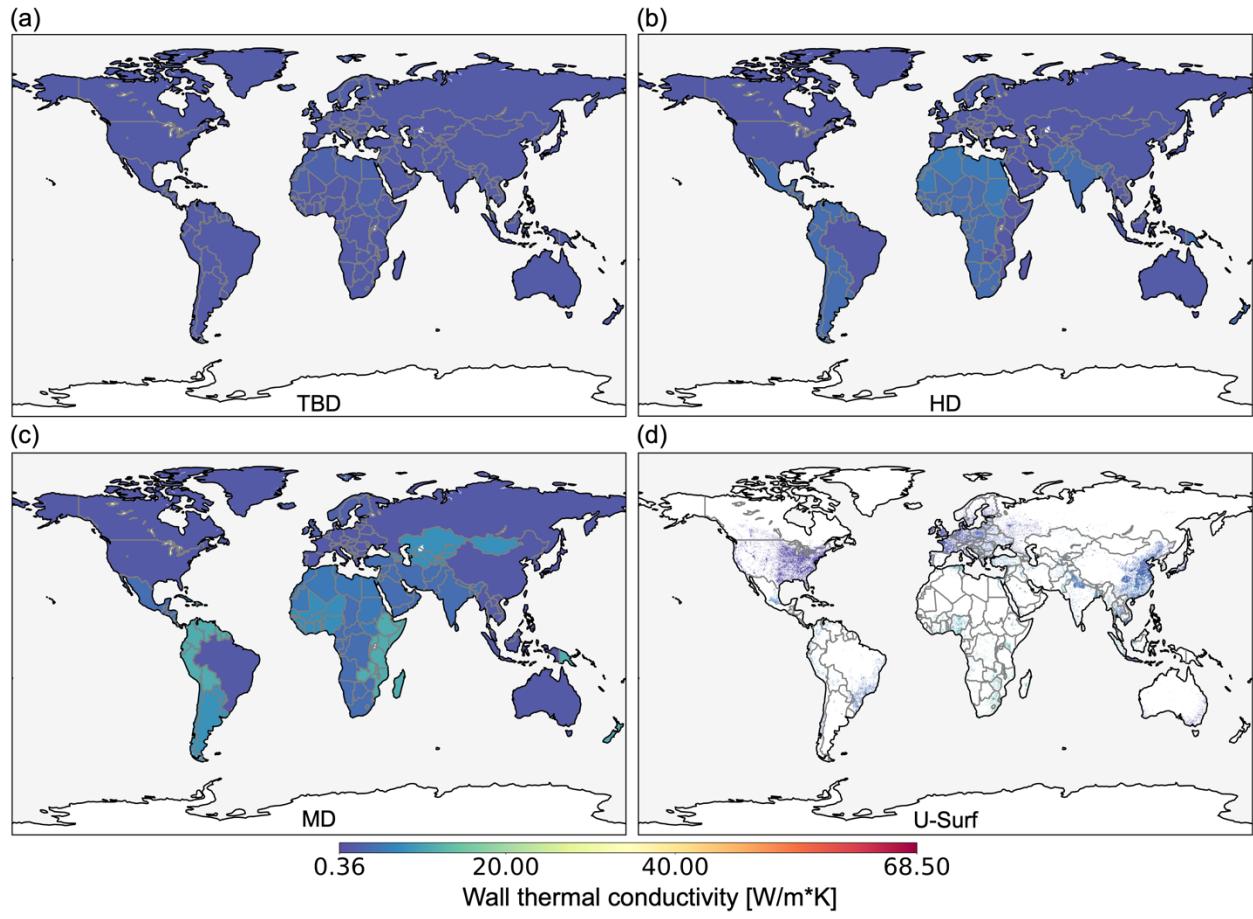
Figure S19. Same as S18, but for wall thickness [meter].
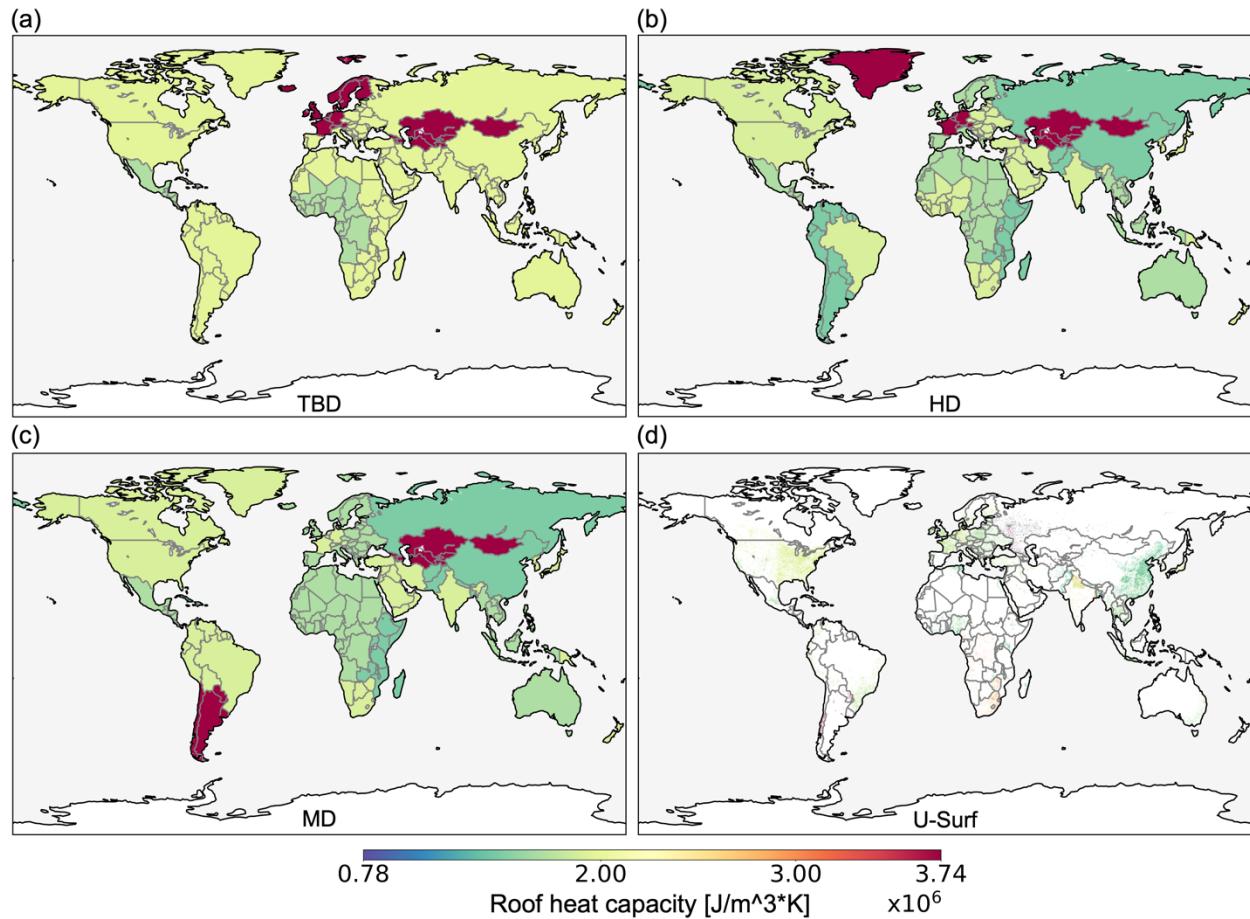
Figure S20. Same as S18, but for roof thermal conductivity [$W/m \cdot K$]. Note that there are 10 distinct layers for roofs and walls, each comes with a different value for thermal conductivity and volumetric heat capacity that depends on the material of each layer. The maps only show the value of the first layer for visualization purposes.

(a) TBD

(b) HD

(c) MD

(d) U-Surf

0.36 20.00 40.00 68.50

Impervious canyon floor thermal conductivity [W/m*K]
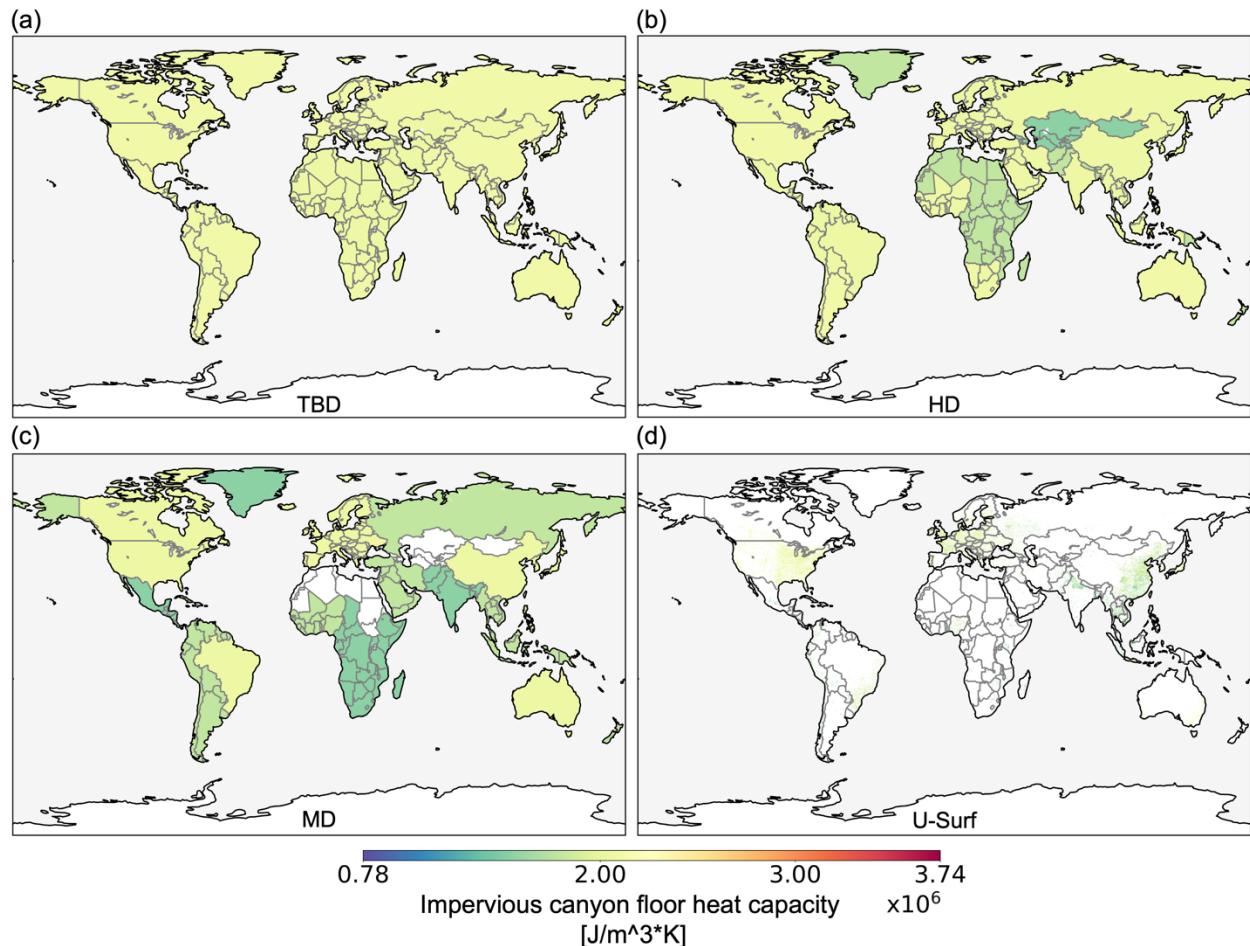
Figure S21. Same as S20, but for impervious canyon floor thermal conductivity $[W/m \cdot K]$.
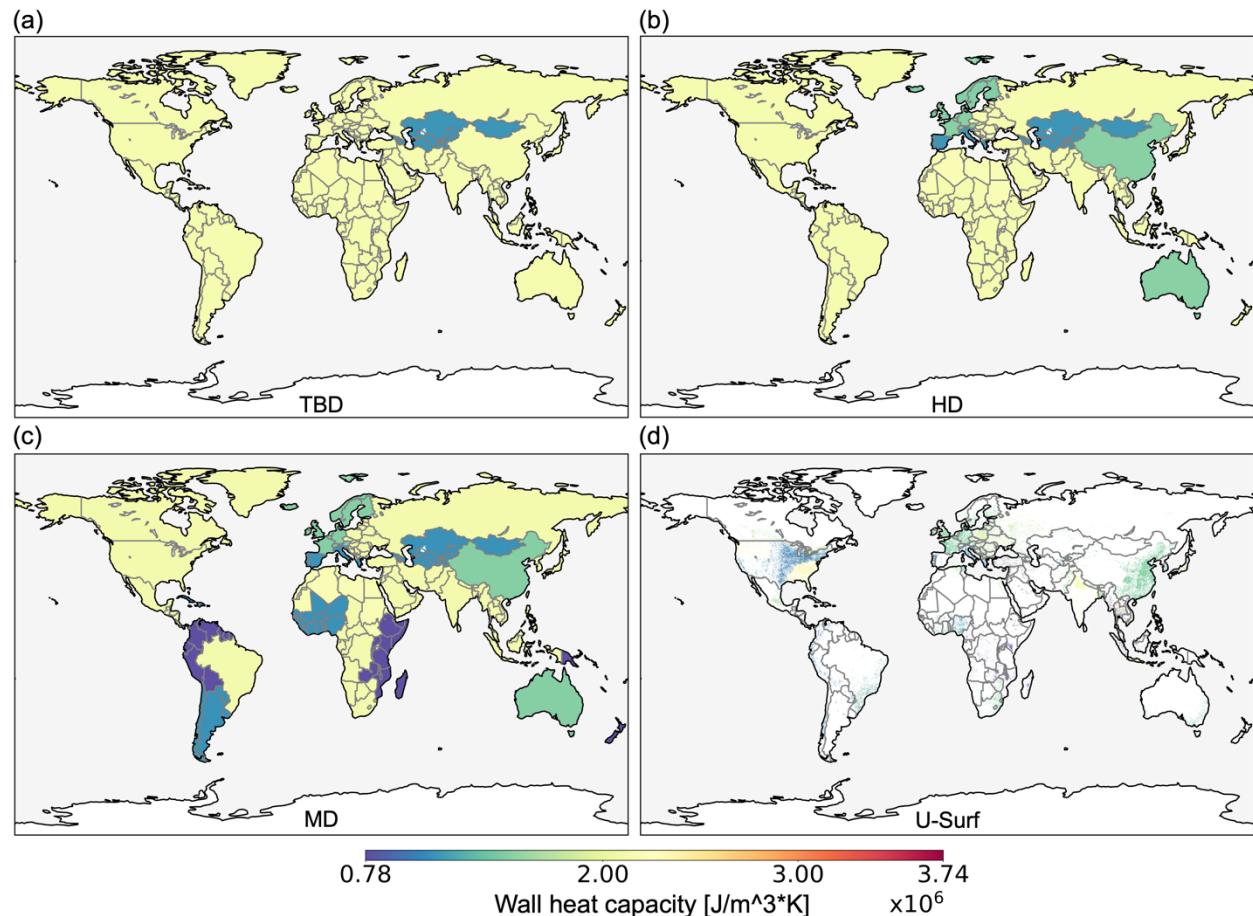
Figure S22. Same as S20, but for wall thermal conductivity $[W/m \cdot K]$.
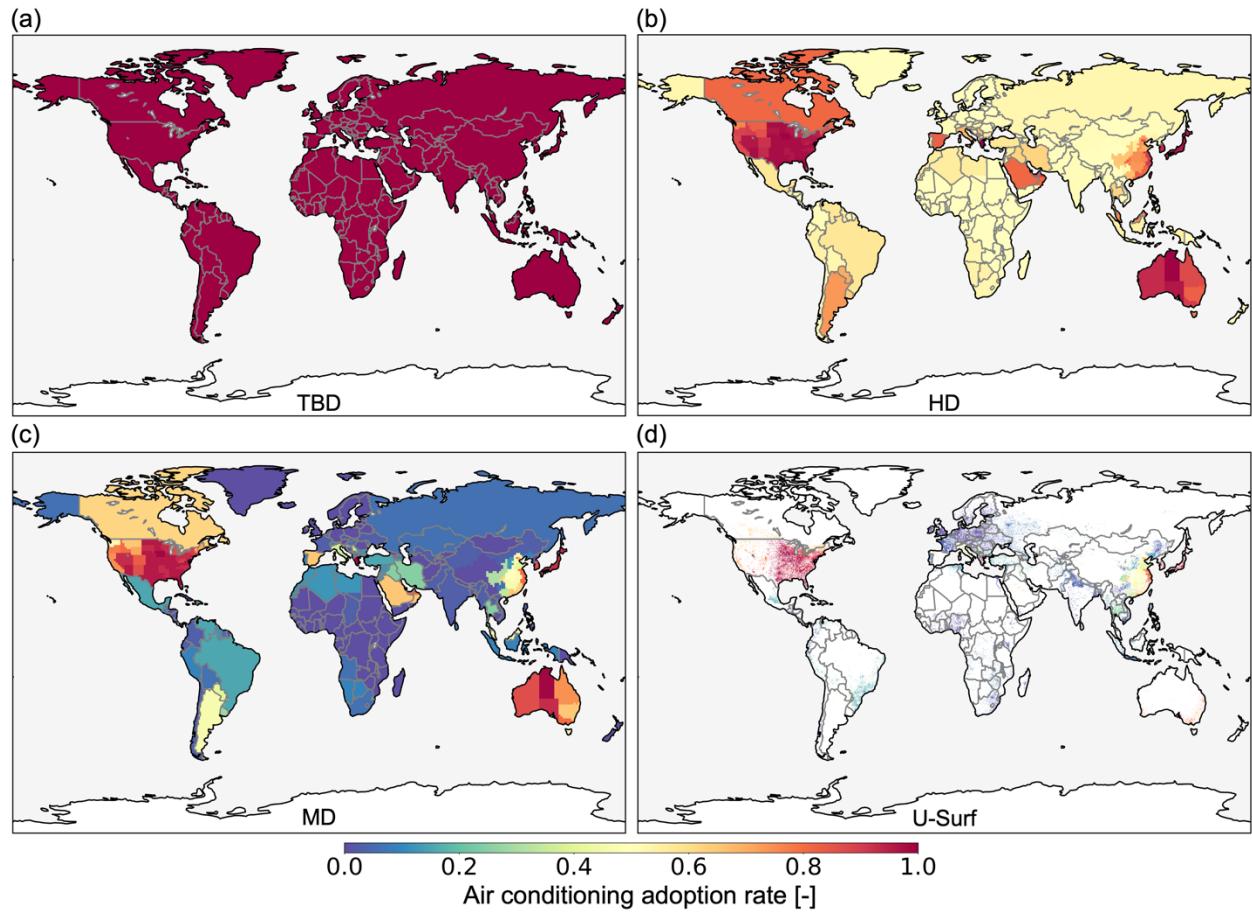
Figure S23. Same as S20, but for roof volumetric heat capacity $[J/m^3 \cdot K]$.
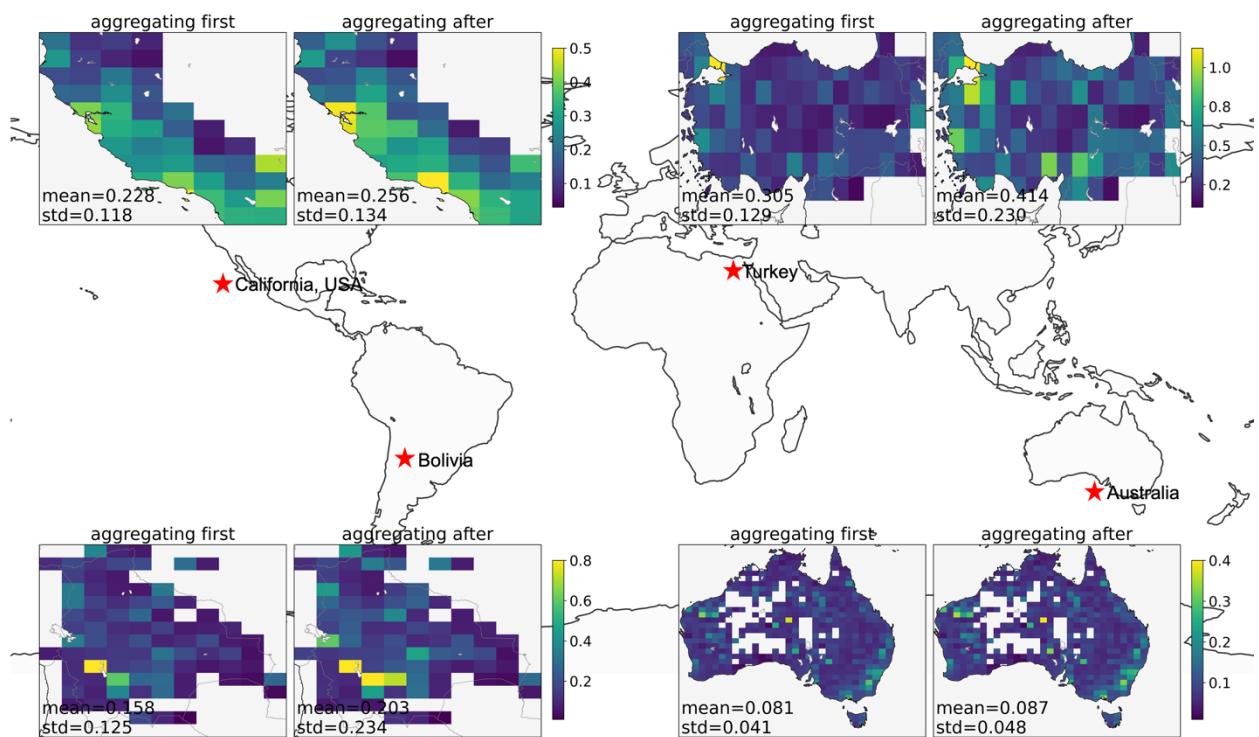
Figure S24. Same as S20, but for impervious volumetric heat capacity $[J/m^3 \cdot K]$.
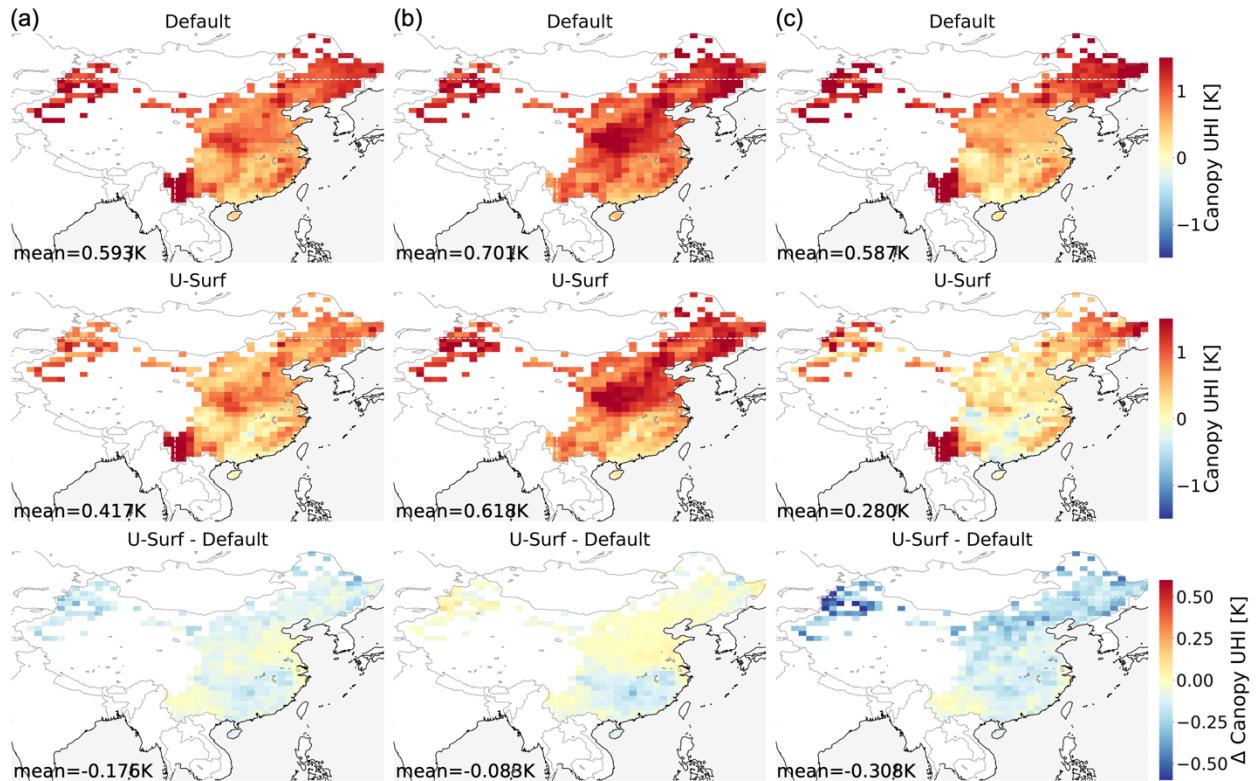
Figure S25. Same as S20, but for wall volumetric heat capacity $[J/m^3 \cdot K]$.

Figure S26. AC adoption rate adapted from Li et al. (2024b). Note that the AC adoption rate was collected at national and subnational level, the gridded dataset used to generate the 1km-resolution map is area-weighted averages of the three density types at $0.9375° × 1.25°$ (latitude × longitude).

Figure S27. Comparison between 'aggregating fist' and 'aggregating after' method for aggregating H/W from 1km to 1º in selected regions/countries.

Figure S28. Spatial distribution of 5-year average canopy urban heat island (UHI) intensity [K] in China. The panels show (a) annual, (b) summer (JJA), and (c) winter (DJF) averages simulated by CESM2 using default urban canopy parameters (UCPs) (top row), U-Surf parameters (middle row), and their difference (U-Surf minus default, bottom row). Negative values in the difference plots indicate weaker UHI intensity with U-Surf parameters.

Table S1. Conservativeness of urban surface property parameters under spatial aggregation.

Category	Parameter	Facet type			
		Roof	Impervious canyon floor	Pervious canyon floor	Wall
Radiative	Emissivity	N	N	N	N
	Albedo	N	N	N	N
Morphological	Fraction*	Y	-	Y	-
	Building height	N	-	-	-
	Canyon height-to-width ratio	N	-	-	-
Thermal	Thickness	N	-	-	N
	Volumetric heat capacity	N	N	-	N
	Thermal conductivity	N	N	-	N

Y: conservative parameters; N: non-conservative parameters; -: not applicable. * Urban percentage is another fractional (conservative) parameter.

Table S2. Selected countries/regions to demonstrate the uncertainty propagation.

Continent	North America		South America		Europe		Asia		Oceania	Africa
Country	United States	Mexico	Argentina	Bolivia	France	Poland	China	Malaysia	Australia	Nigeria
Koppen climate zone*	-	BSh	Cfa	Aw	Cfb	Dfb	-	Af	Bwh	Aw
	Varying	Hot semi-arid	Humid subtropical	Tropical savanna	Temperate oceanic	Warm summer continental	Varying	Tropical rainforest	Hot desert	Topical savanna

* The table only shows the predominant Koppen climate zone if any.

Table S3. Thematic validation results at 21 Urban-PLUMBER sites.

Site	City	Country	Roof fraction MAE	Pervious fraction MAE	Building height MAE (m)	Canyon height-to-width ratio MAE
AU-Preston	Melbourne	Australia	0.110	0.370	0.929	0.124
AU-Surreyhills	Melbourne	Australia	0.061	0.015	4.407	0.204
CA-Sunset	Vancouver	Canada	0.030	0.111	1.664	0.159
FI-Kumpula	Helsinki	Finland	0.005	0.005	0.887	0.013
FI-Torni	Helsinki	Finland	0.165	0.112	4.220	0.582
FR-Capitole	Toulouse	France	0.014	0.046	2.224	0.446
GR-HECKOR	Crete	Greece	0.025	0.023	1.484	0.855
JP-Yoyogi	Tokyo	Japan	0.065	0.180	16.354	0.816
KR-Jungnang	Seoul	South Korea	0.397	0.014	2.900	0.441
KR-Ochang	Ochang	South Korea	0.022	0.271	12.600	0.056
MX-Escandon	Mexico City	Mexico	0.042	0.180	2.562	1.019
NL-Amsterdam	Amsterdam	Netherlands	0.111	0.194	3.780	0.240
PL-Lipowa	Lodz	Poland	0.052	0.101	21.590	0.316
PL-Narutowicza	Lodz	Poland	0.059	0.078	1.008	0.358
SG-TelokKurau06	-	Singapore	0.160	0.118	15.737	1.493
UK-KingsCollege	London	United Kingdom	0.120	0.016	1.652	0.696
UK-Swindon	Swindon	United Kingdom	0.049	0.054	2.446	0.010
US-Baltimore	Baltimore	United States	0.044	0.002	7.239	0.082
US-Minneapolis1	Minnesota	United States	0.050	0.047	6.436	0.057
US-Minneapolis2	Minnesota	United States	0.060	0.114	6.436	0.057
US-WestPhoenix	Arizona	United States	0.050	0.561	7.716	0.113
Average			0.081	0.124	5.918	0.387

Table S4. Thematic validation results at 17 WSF-3D sites.

Site	Country	Roof fraction MAE	Roof fraction RMSE	Building height MAE (m)	Building height RMSE (m)
Almaty	Kazakhstan	0.054	0.070	8.150	10.586
Amsterdam	Netherlands	0.070	0.095	5.098	7.078
Bavaria	Germany	0.065	0.084	4.217	5.955
Cartagena	Colombia	0.088	0.116	8.551	11.996
Dongying	China	0.111	0.130	7.627	9.805
Gyeonggi	South Korea	0.067	0.083	6.029	8.378
Indianapolis	United States	0.045	0.067	5.000	7.264
Kigali	Rwanda	0.062	0.084	6.475	9.823
Lipa	Philippines	0.094	0.104	5.274	6.542
Munich	Germany	0.053	0.069	5.587	7.626
Nairobi	Kenya	0.058	0.082	12.445	17.714
NewYork	United States	0.076	0.105	11.653	14.889
Niamey	Niger	0.112	0.133	10.636	14.473
Seoul	South Korea	0.113	0.124	11.210	13.386
Tanauan	Philippines	0.092	0.104	4.881	5.675
Vienna	Austria	0.063	0.082	6.614	8.872
Washington	United States	0.072	0.094	7.132	8.276
Average		0.076	0.096	7.446	9.902

References

Chakraborty, T. C. and Lee, X.: Using supervised learning to develop BaRAD, a 40-year monthly bias-adjusted global gridded radiation dataset, Sci Data, 8, 238, https://doi.org/10.1038/s41597-021-01016-4, 2021.

Dai, Y., Shangguan, W., Wei, N., Xin, Q., Yuan, H., Zhang, S., Liu, S., Lu, X., Wang, D., and Yan, F.: A review of the global soil property maps for Earth system models, SOIL, 5, 137–158, https://doi.org/10.5194/soil-5-137-2019, 2019.

Delhi Public Geoportal: https://gsdl.org.in/DelhiPublic.html, last access: 16 September 2024.

Esri India: https://www.esri.in/en-in/home, last access: 16 September 2024.

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

Li, L., Bisht, G., Hao, D., and Leung, L. R.: Global 1km land surface parameters for kilometer-scale Earth system modeling, Earth System Science Data, 16, 2007–2032, https://doi.org/10.5194/essd-16-2007-2024, 2024a.

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, 2024b.

Oleson, K., Bonan, B., Feddema, J., Vertenstein, M., and Kluzek, E.: Technical Description of an Urban Parameterization for the Community Land Model (CLMU), https://doi.org/10.5065/D6K35RM9, 2010.

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.

OpenCity - Urban Data Portal: https://opencity.in/, last access: 16 September 2024.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H.: A global soil data set for earth system modeling, Journal of Advances in Modeling Earth Systems, 6, 249–263, https://doi.org/10.1002/2013MS000293, 2014.