JAMES Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2025MS005573

Key Points:

- High-resolution urban-resolving modeling capacity was achieved in Community Terrestrial System Model by integrating U-Surf, a global 1-km urban surface property data set
- More accurate facet-level urban properties improve simulated urban surface meteorology and energy fluxes
- Impacts on surface energy fluxes due to improved urban property constraints have major implications for land-atmosphere interaction

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Zhao,
leizhao@illinois.edu

Citation:

Cheng, Y., Zhao, L., Oleson, K., Chakraborty, T., Zhang, K., He, C., & Yang, J. (2026). Importance of spatially continuous urban surface properties in urban-resolving Earth system modeling. *Journal of Advances in Modeling Earth Systems*, *18*, e2025MS005573. https://doi.org/10.1029/2025MS005573

Received 14 OCT 2025
Accepted 18 MAY 2026

Author Contributions:

Conceptualization: Yifan Cheng, Lei Zhao, Keith Oleson
Data curation: Yifan Cheng
Formal analysis: Yifan Cheng
Funding acquisition: Lei Zhao, Keith Oleson, TC Chakraborty, Keer Zhang
Investigation: Yifan Cheng, Lei Zhao, Keith Oleson, TC Chakraborty, Keer Zhang, Cenlin He, Joyce Yang

Importance of Spatially Continuous Urban Surface Properties in Urban-Resolving Earth System Modeling

Yifan Cheng[1], Lei Zhao[1,2,3], Keith Oleson[4], TC Chakraborty[5], Keer Zhang[6], Cenlin He[7], and Joyce Yang[1]

[1]Department of Civil and Environmental Engineering, The Grainger College of 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]Department of Climate, Meteorology & Atmospheric Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA, [4]Climate and Global Dynamics Laboratory, NSF National Center for Atmospheric Research, Boulder, CO, USA, [5]Atmospheric, Climate, and Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA, [6]High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA, [7]Research Applications Laboratory, NSF National Center for Atmospheric Research, Boulder, CO, USA

Abstract Accurate representation of urban properties and processes at higher resolutions in global modeling systems is essential for advancing our ability to capture the complexities of urban systems and informing effective resilience strategies. However, the prescription of coarse global-scale urban properties in most state-of-the-art Earth system models (ESMs) is limiting their potential for capturing urban signals as they advance toward kilometer-scale simulation capabilities. To bridge this gap in inadequate urban property representation and to advance urban-resolving Earth system modeling, this work integrates the newly-developed global 1 km-resolution facet-level urban surface property data set, U-Surf, into the land component of Community Earth System Model (CESM)—Community Terrestrial System Model (CTSM). The land-only CTSM simulations are validated against satellite measurements, ground-based urban weather stations, flux tower observations, and reanalysis data. Results demonstrate that the enhanced urban properties allow improved simulations of urban meteorology and surface energy fluxes compared to the default coarse-resolution categorical urban canopy parameters. Spatial scaling analysis reveals regime-dependent information loss during resolution aggregation, as well as substantial scale-dependent variations in urban surface energy flux representation. These findings have critical implications for coupled Earth system modeling when including the effect of land-atmosphere interaction. This work establishes a foundation for future urban-resolving kilometer-scale ESM development, which will enable systematic intra- and inter-city comparisons that inform urban adaptation strategies across diverse global urban environments.

Plain Language Summary The lack of fine-scale urban properties is a major limitation of Earth system models. We integrate U-Surf, a global 1 km spatially-continuous data set providing facet-level urban surface properties, into the Community Terrestrial System Model. Validation against satellite measurements, reanalysis data, weather station data and flux tower observations show the improved model performance in capturing urban surface meteorology and energy fluxes brought by more accurate, high-resolution urban representation. Spatial scaling analysis also reveals the climate-regime-dependent information loss when aggregating to coarser grids. Our results demonstrate the importance of urban representation in kilometer-scale Earth system modeling.

1. Introduction

Urban areas, where more than 50% of the world's population reside, exhibit distinct radiative, morphological and thermal properties, which can have profound impacts on local to regional weather (Chakraborty & Qian, 2024; Qian et al., 2022) and climate (Liao et al., 2018; United Nations, 2019; Zhao et al., 2021). As cities worldwide experience and modulate climate change across scales, a global perspective becomes critical to systematically quantify the two-way urban-climate interactions. Offering a universally consistent framework to simulate urban effects globally, Earth System Models (ESMs) or global climate models have therefore become essential tools for urban climate risks analysis (X. C. Li, Zhao, Qin, et al., 2024; Yang et al., 2023; K. Zhang et al., 2023), urban adaptation strategies assessment (Demuzere et al., 2014; J. Zhang et al., 2016; K. Zhang et al., 2025), and cross-sectoral urban resilience planning (X. C. Li, Zhao, Oleson, et al., 2024; Sun et al., 2024; J. Zhang et al., 2016).

Methodology: Yifan Cheng, Lei Zhao, Keith Oleson, TC Chakraborty, Keer Zhang
Project administration: Lei Zhao, Keith Oleson
Resources: Lei Zhao, Keith Oleson
Supervision: Lei Zhao, Keith Oleson
Validation: Yifan Cheng, Lei Zhao, TC Chakraborty, Keer Zhang, Cenlin He, Joyce Yang
Visualization: Yifan Cheng, Joyce Yang
Writing – original draft: Yifan Cheng
Writing – review & editing: Yifan Cheng, Lei Zhao, Keith Oleson, TC Chakraborty, Keer Zhang, Cenlin He, Joyce Yang

Nevertheless, despite growing interests and a pressing need of ESMs for urban climate research, the urban-resolving capabilities in these models are fundamentally limited. A longstanding deficiency in these global-scale models is their near-universal lack of urban representation. This is mainly because early versions of these models were designed for large-scale purposes and were commonly applied at coarse resolutions (usually 1° or coarser), with urban areas assumed to be either too small to resolve or to cause discernible changes in broader-scale dynamics. However, neither of the assumptions are necessarily true. First, ESMs are increasingly being used for regional climate assessments in addition to their original purpose of global climate modeling. Given the importance of spatial heterogeneity at these regional scales, various ESMs have been run at higher and higher spatial resolutions in recent years (Haarsma et al., 2016; Roberts et al., 2025). Many modeling groups now have specific development plans for running ESMs at kilometer scales (km-scale) in the near- or medium-term, such as NCAR's Community Earth System Model (CESM), DOE's Simple Cloud Resolving E3SM Atmosphere Model (SCREAM) and ECMWF's Integrated Forecasting System (Caldwell et al., 2021; Rackow et al., 2025; Skamarock et al., 2012). At these fine spatial scales, resolving urban areas and their interactions with background climate is critical to capture both the heterogeneity in urban microclimate and potential feedback between urban areas and the regional climate system in ESMs. For instance, resolving urban impacts at the km-scale can provide insights about hotspots of heat stress within cities, with implications for urban planning (Jiang et al., 2025), environmental justice (Hsu et al., 2021), and optimizing energy grids (Battini et al., 2025). Similarly, urbanization-induced cloud modification and its consequences on extreme precipitation and urban flooding, often examined using limited observations (Torelló-Sentelles et al., 2024; Vo et al., 2023) or regional weather and climate models (Moraglia et al., 2024; Young et al., 2025), can be studied at hemispherical or global scales using kilometer-resolution ESM simulations, revealing broader patterns of urban impacts on the hydrological cycle.

Second, beyond the well-established urban modulation effect on local surface energy balance (Demuzere et al., 2013; Kotthaus & Grimmond, 2014), meteorological signals (Yang et al., 2023; K. Zhang et al., 2023), and energy systems (X. C. Li, Zhao, Oleson, et al., 2024; X. C. Li, Zhao, Qin, et al., 2024), a growing body of research from both observational studies and regional modeling has shown that urban surface characteristics also fundamentally alter regional climate processes. Cities could reshape regional temperature distributions (Chakraborty & Lee, 2019b; Chakraborty & Qian, 2024; Scott et al., 2018), modify precipitation patterns and shallow cumulus cloud formation (Singh et al., 2020; Tian et al., 2022, 2024; Wu et al., 2021; Yan et al., 2024), and affect boundary layer dynamics (Hou et al., 2013; Zhu et al., 2016) and extreme weather events (D. Li & Bou-Zeid, 2013; W. Zhang et al., 2018; Zhao et al., 2018). Moreover, the assumption of negligible urban impacts on regional climate will be increasingly unjustified going forward as urban areas continue to rapidly expand in the future (G. Chen et al., 2020; X. Li et al., 2019).

Despite considerable progress in the development of km-scale ESMs in recent years (Moon et al., 2025; Moreno-Chamarro et al., 2025; S. Zhang et al., 2020), the urban components of these models, both in terms of properties and processes, are substantially lagging behind compared to their atmospheric counterparts. This creates a fundamental mismatch where fine-grain atmospheric modeling capabilities are limited by simplistic urban surface representation, thus underestimating the surface spatial heterogeneity and failing to resolve urban processes at kilometer scales. This mismatch is seen in the Community Earth System Model (CESM) and Energy Exascale Earth System Model (E3SM), which share similar version of urban component and are already among the very few exceptions in state-of-the-art ESMs that have an explicit, physically-based urban land scheme (Lawrence et al., 2018). Their urban representations, however, are not sufficient for km-scale modeling capabilities. The urban canopy parameters in these models are coarse-grained, low-resolution, and spatially discontinuous, thus providing poor constraints for the biophysical properties of built surfaces and failing to realistically represent the spatial heterogeneity across and within cities. In CESM and E3SM, the globe is divided into 33 regions with up to three urban density classes per region. Large countries like China and Russia are labeled as one region per country, and even data-abundant areas such as the contiguous United States (CONUS) consist of only six sub-regions (Figure S1 in Supporting Information S1). These categorical parameters cannot resolve urban properties at the kilometer scale. Moreover, some of these existing materials-based urban estimates diverge greatly from those derived from direct satellite observations (Chakraborty et al., 2021).

Currently, there are two major challenges toward km-scale urban-resolving ESMs: (a) missing global high-resolution (km-level) urban surface constraints in the land model of ESMs, and consequently, (b) inaccurate representation of urban surface processes at km scales. This study represents the first step in advancing global high-resolution urban-resolving Earth system modeling by leveraging a newly developed global 1-km facet-level urban surface property data set, namely U-Surf (Cheng, Zhao, Chakraborty, et al., 2025). Here we investigate a key scientific question: how enhanced urban surface representation and spatial resolution advance the urban-resolving capability of ESMs. To address this question, we incorporate the U-Surf data into CESM and conduct land-only simulations across various spatial scales. We evaluate the simulation results against multiple reference data sets including satellite measurements, observation-based reanalysis data, station and flux tower measurements. We further conduct spatial scaling analysis to quantify the impact of using coarser resolution data sets on information loss to provide insights on land-atmosphere interaction. This work establishes a parametric baseline for future km-scale urban process simulations within Earth system modeling frameworks.

This paper is structured as follows. Section 2 outlines the technical details in the integrated data set, the modeling system, and simulation setup. Section 3 describes the workflow of model performance evaluation and spatial scaling analysis. Section 4 documents the assessment of simulation results, discussions and broad implications. Conclusions follow in Section 5.

2. Methods and Data

2.1. Overview of CLMU Parameterization and Modeling Scheme

The land component of CESM, namely the Community Terrestrial System Model (CTSM), models different land surfaces as sub-grid landunits. Its urban module (Community Land Model Urban, hereafter referred to as CLMU) employs a single-layer urban canopy structure to capture the interactions between urban surfaces and surrounding environments (Oleson et al., 2010; Oleson & Feddema, 2020). This structure has been widely adopted by other ESMs including the U.S. Department of Energy's E3SM, the Norwegian Earth System Model (NorESM), and the Climate Change coupled climate model (CMCC-CM2) (Bentsen et al., 2013; Caldwell et al., 2019; Cherchi et al., 2019). This single-layer model conceptualizes the urban canopy as an infinitely long street bordered by two building walls with identical height. The urban land unit is further divided into five columns (or facets): roof, impervious and pervious canyon floors, sunlit and shaded walls (Figure 1). Each facet is associated with a set of parameters with three major categories: radiative, morphological and thermal properties (Table 1). The default urban surface data set in CLMU is based on the global urban canopy parameters (UCPs) developed by Jackson et al. (2010), and has been updated by Oleson and Feddema (2020).

2.2. The Default and U-Surf-Based Urban Canopy Parameters

In the default CLMU configuration, urban canopy parameters are developed over 33 distinct regions (Jackson et al., 2010). Urban areas with similar climates, socio-economic characteristics, and architectural practices are grouped into the same region. To further differentiate the urban form within each region, there are 4 density classes: tall building district (TBD), high density (HD), medium density (MD) and low density (LD). Note that the LD class is not used in the current implementation of surface data set as it is considered to be better modeled as a vegetated/soil surface due to the sparsely distributed built-up areas and the high pervious fraction (Lawrence et al., 2018). The associated urban surface parameters were developed for an older period (circa-2000) when high-resolution satellite images were scarce, thus, these parameters are limited by the spatial resolution and overall quality of the data available at the time of development. The coarse categorization has overlooked the realistic heterogeneity of cities, thus impeding CESM's application in studying urban interactions with background climate, particularly relevant for high-resolution urban climate modeling.

The newly developed U-Surf data set, for the first time, provides the required UCPs for Earth system modeling in a spatially-continuous manner at 1-km resolution (Cheng, Zhao, Chakraborty, et al., 2025). It demonstrates significant improvement in representing urban land heterogeneity both within and across cities and provides globally-consistent, high-fidelity surface biophysical data to urban-resolving ESMs. Figure 2 provides an illustrative example comparing the default urban surface parameters with those derived in U-Surf. The comparison clearly shows the enhanced fine-scale spatial texture in the U-Surf product (Figure 2d) in contrast to the large-scale homogeneous representation in CESM's default urban data set (Figures 2a–2c). More details about the U-Surf data set and its parameter derivations can be found in Cheng, Zhao, Chakraborty, et al. (2025). In this study, we incorporated the U-Surf parameters into CESM/CLMU and examined how the enhanced urban surface representation and spatial resolution affect its urban modeling capability across spatial scales.

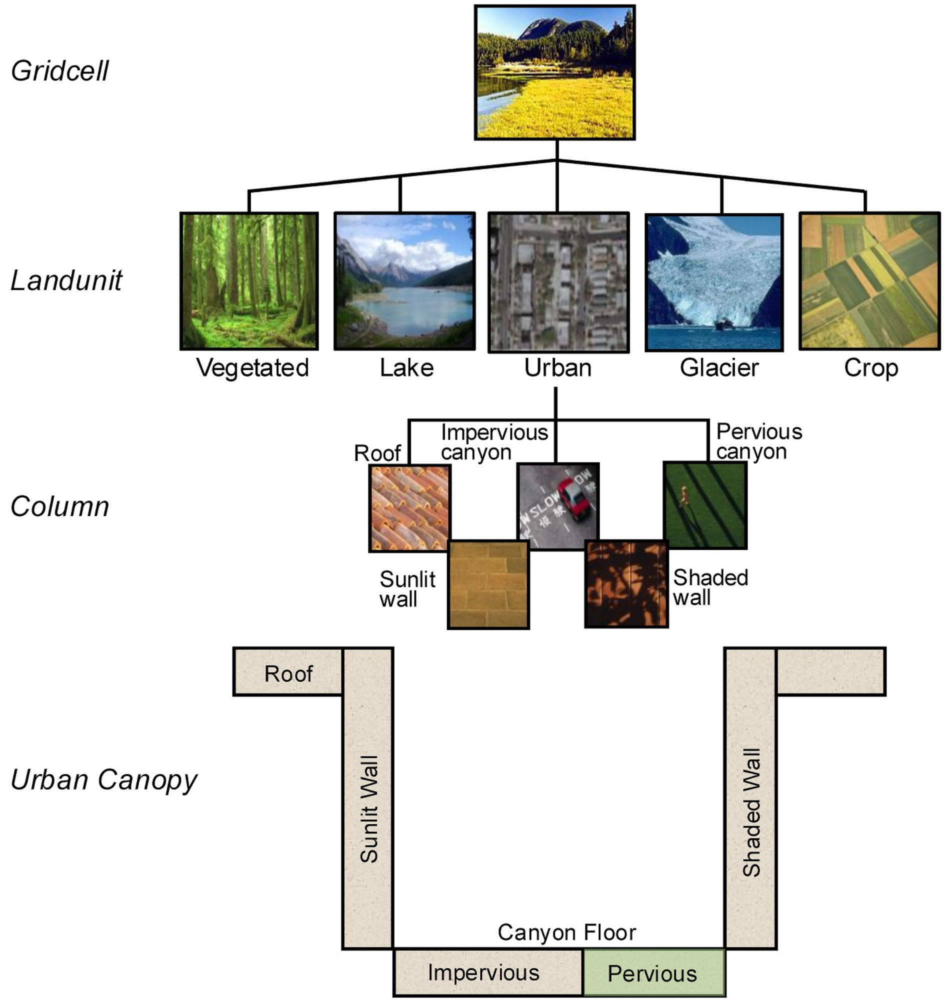

Figure 1. Community Terrestrial System Model representation hierarchy and embedded single-layer urban canopy model (adapted from Figure 2 in Lawrence et al. [2019] and Figure 2 in Oleson et al. [2008]).

2.3. Simulation Configurations

2.3.1. Land-Only Simulations

To evaluate how the improved urban representation influences urban meteorology across scales, we conducted two sets of land-only simulations using CESM (Table 2), where only the land component is active and decoupled from other prescribed components (e.g., atmosphere, ocean). These two sets of simulations, CTRL and USURF, differ only in the urban surface data sets. For CTRL simulations, the surface data set is generated using default urban canopy parameters from Oleson and Feddema (2020), whereas in USURF cases, we used the surface data set derived from 1-km U-Surf parameters. Each simulation set was performed at four different spatial resolutions, labeled as LR (low-resolution), MR (medium-resolution), HR (high-resolution) and UHR (ultra-high-resolution), each with corresponding spatial coverage specified in Table 2. Owing to the intensive computational demands, the HR and UHR simulations were performed only over the CONUS. To support these simulations, the original 1-km U-Surf data was aggregated to corresponding resolutions using the *mksurfdata_esmf* tool. The details on the surface data set creation are documented in Text S1 in Supporting Information S1. The aggregation of 1-km raw data to coarser resolutions involved different treatments to area-based conservative and non-conservative parameters; more details are documented in Cheng, Zhao, Chakraborty, et al. (2025). In addition, to integrate the air conditioning (AC) adoption rate and the latest developed explicit AC scheme in CLMU, the code base used is consistent with X. C. Li, Zhao, Oleson, et al. (2024).

Table 1
Urban Canopy Parameters Used in Community Land Model Urban

Category	Parameter	Long name	Unit
Radiative	EM_ROOF	Emissivity of roof	Unitless [–]
	EM_IMPROAD	Emissivity of impervious canyon floor	
	EM_PERROAD	Emissivity of pervious canyon floor	
	EM_WALL[a]	Emissivity of wall	
	ALB_ROOF[b]	Albedo of roof	
	ALB_IMPROAD	Albedo of impervious canyon floor	
	ALB_PERROAD	Albedo of pervious canyon floor	
	ALB_WALL	Albedo of wall	
Morphological	HT_ROOF	Roof height	m
	WTLUNIT_ROOF	Roof fraction (w.r.t. urban surface)	Unitless [–]
	WTROAD_PERV	Pervious canyon floor fraction (w.r.t. canyon floor)	
	CANYON_HWR	Canyon height to width ratio	
	PCT_URBAN	Urban percentage (w.r.t. gridcell)	
Thermal	TK_ROOF	Thermal conductivity of roof	$Wm^{-1}K^{-1}$
	TK_IMPROAD[c]	Thermal conductivity of impervious canyon floor	
	TK_WALL	Thermal conductivity of wall	
	CV_ROOF	Volumetric heat capacity of roof	$Jm^{-3}K^{-1}$
	CV_IMPROAD	Volumetric heat capacity of impervious canyon floor	
	CV_WALL	Volumetric heat capacity of wall	
	THICK_ROOF	Roof thickness	m
	THICK_WALL	Wall thickness	
	NLEV_IMPROAD	Number of impervious canyon floor layers	Unitless [–]
	T_BUILDING_MIN	Minimum interior building temperature	K
	T_BUIILDING_MAX[d]	Maximum interior building temperature	
	P_AC[d]	Air conditioning adoption rate	Unitless [–]

[a]The properties of sunlit and shaded walls are identical, however, the radiation partitioning and transfer at these two facets are modeled differently (Oleson et al., 2010). [b]There are two albedo parameters, ALB_{}_DIF, ALB_{}_DIR for each facet, representing diffuse and direct albedo. They are set to be the same in the current version of CLMU. [c]The thermal conductivity, volumetric heat capacity are defined with each layer of roofs, impervious canyon floors and walls based on the construction material. [d]T_BUILDING_MAX and P_AC are stored in the urbantv stream file that includes parameters with additional time dimension while the rest of parameters are stored in the surface data set file.

All simulations were driven by the same atmospheric forcings, which come from hourly bias-corrected data of the European Centre for Medium-Range Weather Forecasts Reanalysis version 5 (ERA5) reanalysis data (Cucchi et al., 2020), including precipitation, incident solar and longwave radiation, surface temperature, surface pressure, specific humidity and wind at the lowest atmospheric level. To reduce computational costs caused by spatial interpolation on-the-fly by CESM, the forcing data for UHR cases went through spatial interpolation using the same bilinear regridding method in advance to match the model resolutions. All simulations were fully spun up through cyclic forcing, where the 2010–2014 bias-corrected ERA5 climatology was iteratively applied. The spin-up period varied by spatial resolution but typically required between 15 and 25 years. The production run spanned from 2010 to 2019, producing outputs at monthly and daily frequency.

2.3.2. Site-Level Simulations

Traditional flux tower observations, which provide all the components of the surface energy budget, generally represent natural surfaces, making it difficult to evaluate the urban surface energy balance. The recent Urban-PLUMBER project primarily aims to enhance the understanding of the accuracy of current urban climate models and has also produced a harmonized data set of quality-controlled and gap-filled observations from 21

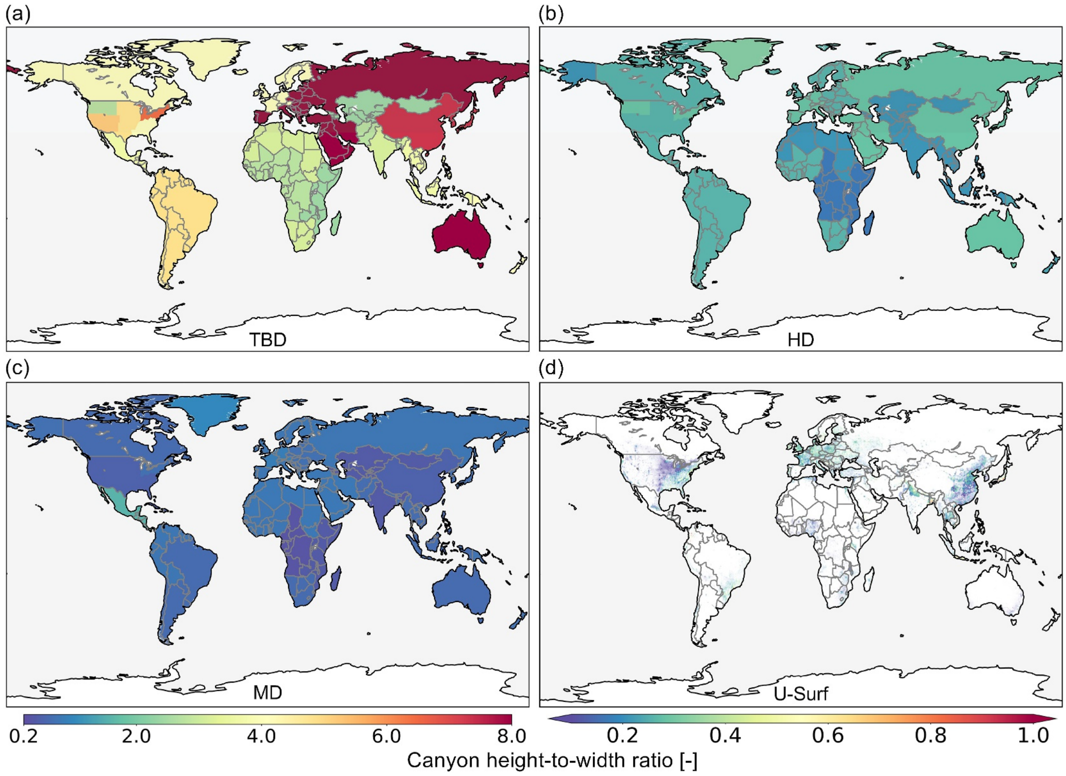

Figure 2. Comparison between (a–c) default Community Land Model Urban and (d) U-Surf urban canopy parameter. The parameter shown here is canyon height-to-width ratio (unitless), which represents the structural layout and compactness of the built area. Panels (a–c) share the left colorbar and panel (d) uses the right one. In the default case, each density class (TBD, Tall building district; HD, high density; MD, medium density) within an urban region is assigned a corresponding value. In the U-Surf data set, parameters are spatially continuous without classification into density classes (adapted from Figure S17 in Cheng, Zhao, Chakraborty, et al. [2025]).

urban flux tower sites across different climate zones and urban built environments (Lipson et al., 2022; Figure S3 in Supporting Information S1). We further performed point-scale (site-level) simulations customized for the 21 Urban-PLUMBER flux tower stations to compare with observations on surface energy flux components. We ran two sets of point-scale simulations labeled SP_CTRL and SP_USURF. Each set consisted of four different spatial resolutions (LR, MR, UHR and 1 km; Table 3). Note that here the resolution specifically refers to the scale to which the urban surface parameters were aggregated, whereas all simulations were conducted as single points (conceptually as 1 × 1 grid at each site) using the same atmospheric forcing. The continuous Urban-PLUMBER forcing was constructed using both observational and ERA5-derived meteorological data for each site (Lipson et al., 2022). Most sites feature data at a 30-min frequency, except for JP-Yoyogi, PL-Lipowa, PL-Narutowicza,

Table 2
Land-Only Simulation Configurations

Atmosphere forcing	Spatial extent	Spatial resolution	Simulation ID	UCPs
Bias corrected ERA5 hourly data (WFDE5)	Global	0.9375° × 1.25°	CTRL_LR	Default
			USURF_LR	U-Surf
		0.23° × 0.31°	CTRL_MR	Default
			USURF_MR	U-Surf
	CONUS	0.125° × 0.125°	CTRL_HR	Default
			USURF_HR	U-Surf
		0.03125° × 0.03125°	CTRL_UHR	Default
			USURF_UHR	U-Surf

Table 3
Site-Level Single Point Simulation Configurations

Atmosphere forcing	Spatial resolution[a]	Simulation ID	UCPs
Urban-PLUMBER forcing	0.9375° × 1.25°	SP_CTRL_LR	Default
		SP_USURF_LR	U-Surf
	0.23° × 0.31°	SP_CTRL_MR	Default
		SP_USURF_MR	U-Surf
	0.03125° × 0.03125°	SP_CTRL_UHR	Default
		SP_USURF_UHR	U-Surf
	0.00833° × 0.00833°[b]	SP_CTRL_1 km	Default
		SP_USURF_1 km	U-Surf

[a]The resolution here only indicates the scale to which the urban surface parameters are aggregated; all simulations are conducted at single points (1 × 1 grid at each site). [b]0.00833° refers to the 1 km raw data values from U-Surf data set.

UK-KingsCollege, UK-Swindon, and US-Baltimore, which have hourly forcing. Note that the calculation of urban surface aerodynamic constants in point-scale simulations employed a modified approach compared to global and regional land-only simulations. Detailed methodology is described in Text S2 in Supporting Information S1.

3. Model Validation and Scale Analysis

3.1. Data for Model Evaluation

Validation of urban modeling results remains challenging due to the scarcity of globally available and consistent measurements. For the purpose of model evaluation, we combined multiple observed data sources, including satellite, weather station, reanalysis data and flux tower measurements, to evaluate both large-scale patterns and localized urban signals as comprehensively as possible.

The 1-km quality controlled land surface temperature (LST) observational data were calculated following Chakraborty and Qian (2024). We used the annual summertime and wintertime composites of observations from the AQUA satellite carrying the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor with local passing time at approximately 1:30 p.m. and 1:30 a.m. Pixels with uncertainty greater than 3 K based on the product quality control flag have been removed from the collection (Chakraborty & Qian, 2024). Before regridding to the desired grids to validate against the urban skin temperature, which is derived by surface upward longwave radiation, from the 3.5-km (0.03125°) simulations over CONUS, we only included pixels with positive urban percentage as defined in the U-Surf data set. Additionally, we also utilized the 4-km gridded observation-based daily average, minimum and maximum near-surface (2-m) air temperature from the Parameter-elevation Regressions on Independent Slopes (PRISM) data set in 10 CONUS metropolitan areas across diverse climate zones. Error metrics were calculated exclusively using urban land pixels, with all water-containing (ocean/lake) pixels excluded following the practice in Krayenhoff et al. (2018).

The near-surface air temperature and relative humidity (RH) measured by urban stations were obtained from K. Zhang et al. (2023), which compiled a global data set of 133 urban sites. The stations were selected to ensure their locations are within urban landscapes rather than transitional settings such as municipal airports. Each site meets specific criteria, including a built-up fraction greater than 0.45 within the 1-km radius, coordinates' accuracy better than 200 m, and sensor placement at a height between 1.3 and 3 m above the ground (K. Zhang et al., 2023). To evaluate model performance at global scale, we compared near-surface air temperature and RH from global 0.23° × 0.31° simulations (CTRL_MR and USURF_MR) against measurements from these urban stations. Accounting for the varying temporal coverage across stations, here we used 119 out of 133 stations with valid data available during our analysis period of 2010–2019 after data cleaning and quality control.

We also evaluated six modeled surface energy flux variables at 21 Urban-PLUMBER sites: upwelling shortwave radiation (SW_{up}), upwelling longwave radiation (LW_{up}), net radiation (Q^*), sensible heat flux (Q_h), latent heat flux (Q_{le}), and residual term (Q_s) representing the sum of storage heat flux (Q_{stor}) and anthropogenic heat flux (Q_{anth}). Among these, Q^* and Q_s were calculated using Equations 1 and 2:

$$Q^* = \mathrm{LW}_{\mathrm{down}} - \mathrm{LW}_{\mathrm{up}} + \mathrm{SW}_{\mathrm{down}} - \mathrm{SW}_{\mathrm{up}} \tag{1}$$

$$Q_s = Q_{\mathrm{stor}} + Q_{\mathrm{anth}} = Q^* - Q_h - Q_{\mathrm{le}} \tag{2}$$

Note that the length of available observations and analysis period varies across 21 Urban-PLUMBER sites, as summarized in Table S1 in Supporting Information S1. Model performance was evaluated using only the time stamps for which quality controlled observational data were available. These measurements were used to assess four groups of point-scale simulations (cases in Table 3).

In addition to the calculation of changes in absolute bias for long-term averages of surface fluxes, we used the Taylor diagram to summarize the overall model performance at short temporal intervals. Taylor diagrams demonstrate three error statistics simultaneously: normalized standard deviation (σ), Pearson correlation coefficient (R) and normalized centered-root-mean-square error (cRMSE) (Taylor, 2001). A "perfect" model is defined as a model with $\sigma = 1, R = 1, \mathrm{cRMSE} = 0$. These error metrics were calculated using Equations 3–5:

$$\sigma = \frac{\sigma_M}{\sigma_O} = \frac{\sqrt{\frac{1}{n-1}\Sigma\left(M_i - \overline{M}\right)^2}}{\sqrt{\frac{1}{n-1}\Sigma\left(O_i - \overline{O}\right)^2}} = \sqrt{\frac{\Sigma\left(M_i - \overline{M}\right)^2}{\Sigma\left(O_i - \overline{O}\right)^2}} \tag{3}$$

$$R = \frac{\Sigma\left(M_i - \overline{M}\right)\left(O_i - \overline{O}\right)}{\sqrt{\Sigma\left(M_i - \overline{M}\right)^2}\sqrt{\Sigma\left(O_i - \overline{O}\right)^2}} \tag{4}$$

$$\mathrm{cRMSE} = \sqrt{1 + \sigma^2 - 2\sigma \cdot R} \tag{5}$$

where M stands for model output and O represents observations.

3.2. Spatial Scaling Analysis

We conducted the spatial scaling analysis using the five-year (2010–2014) annual average outputs from land-only simulations documented in Table 2. The main purpose of this analysis is to investigate the impact of km-scale spatial heterogeneity and upscale aggregation on modeled urban meteorology. We chose two representative variables, Bowen ratio (BR) and net radiation (NR). The Fourier power spectrum analysis, originally developed to understand energy distribution across atmospheric scales in boundary layer and turbulence studies (Charrondière & Stiperski, 2024; Nastrom et al., 1984), has also been applied to analyze the effective resolution of numerical weather prediction models such as the Weather Research and Forecast (WRF) model (Skamarock, 2004). We performed a two-dimensional Fast Fourier Transform over CONUS to compute the power spectral density as a function of the wavenumber, which under this context, can be interpreted as the relationship between variance and spatial frequency. Then we calculated the local coefficient of variation to demonstrate how spatial resolution affects the capture of small-scale features. The coefficient of variation (CV) is expressed as the ratio between standard deviation and mean values. We applied the sliding window algorithm with resolution-independent window sizing (constant at ~$3° \times 3°$) to calculate local CV. Higher CV values indicate more spatial heterogeneity and better resolution of small-scale features.

In addition, we selected four $5° \times 5°$ regions representing different background climate regimes based on the Koppen climate classification (Beck et al., 2018). For each region, we quantified the information loss introduced by spatial aggregation following the method described in L. Li et al. (2024) and Vergopolan et al. (2022). According to the model configurations in Table 2, the coarser-scale simulations have the scale ratio ($\lambda_{\mathrm{scale}}/\lambda_{\mathrm{finest}}$, where λ represents spatial scale) of 4 (USURF_HR), 8 (USURF_MR) and 32 (USURF_LR) compared to the finest resolution (USURF_UHR). The ratio of spatial standard deviation ($\sigma_{\mathrm{scale}}/\sigma_{\mathrm{finest}}$, where σ is the standard deviation) was calculated within each region at different spatial scales. We then fitted a logarithmic relationship between scale ratio and standard deviation ratio. A steeper negative slope β indicates stronger scale-dependency of data spatial variability, which corresponds to higher information loss γ, defined as $1 - \sigma_{\mathrm{scale}}/\sigma_{\mathrm{finest}}$, through the upscaling process. We further selected one city from each region to showcase the impact of resolutions on the distributions of BR and NR in detail. Building upon the quantification of information loss, we performed a supervised Principal Component Analysis (sPCA) across all urban pixels over CONUS using a 2° × 2° fixed window to identify the major physical drivers underlying the information loss patterns. sPCA uses the target to find latent components that maximize the covariance between predictors X and target Y (Abdi, 2010). In sPCA, X and Y are decomposed as Equations 6 and 7:

$$X = TP^T \tag{6}$$

$$Y = UQ^T \tag{7}$$

where T stands for the latent-variable score matrix, that is, reprojected X in the latent-variable space, P is the loading matrix, which represents the coefficients (contributions) of each predictor to latent variables. Similarly, U and Q are the score and loading matrix on the Y side, respectively. In the sPCA biplot, the two axes show the first and second X score, T_1, T_2. The arrow length represent the loadings while the arrow direction indicates the positive/negative association between predictors and latent variables (Abdi, 2010; Oyedele & Lubbe, 2015). Since loadings mainly describe the alignment between predictors and latent axes, to better identify the most influential drivers of information loss, we used X weight-matrix W and T, Q to calculate the Variable Importance in Projection (VIP) as shown in Text S4 in Supporting Information S1 to select the top 10 important drivers to be displayed in the biplot.

4. Results and Discussions

4.1. Improvement of Modeled Surface Temperature and Humidity

In this section, we focus our analysis on model performance during summertime, defined as the multi-year June-July-August (JJA) average in the Northern Hemisphere and December-January-February (DJF) average in the Southern Hemisphere, for several reasons. First, the remotely-sensed radiative parameters (emissivity and albedo) in the U-Surf data set represent mean states under cloud-free and snow-free scenarios, making them more representative of summer conditions. During winter, snowpack can alter the surface albedo, and snowmelt affects hydrological conditions, further contributing to temperature variability (Simpson et al., 2022). The reduced influence of snow-related processes during summer allows for clearer attribution of model improvements due to enhanced urban surface representation instead of confounding seasonal factors. Second, summertime temperature extremes are more directly associated with heatwaves, with substantial impacts on public health (K. Chen et al., 2023; Su et al., 2025; Wang et al., 2024), making accurate simulations of summertime meteorology particularly important for climate adaptation. The wintertime results corresponding to Figures 3–6, defined as multi-year DJF average in the Northern Hemisphere combined with JJA average in the Southern Hemisphere, can be found in Figures S4–S7 in Supporting Information S1, however, the interpretation of results might require additional considerations that are not explicitly captured in the current modeling framework.

The urban representation from the U-Surf data set significantly improves the performance of urban LST in terms of both magnitude and spatial distribution. Comparing LST patterns in USURF_UHR against CTRL_UHR simulations, the correlation and alignment with satellite measurements are improved across CONUS cities, where the Pearson correlation coefficient (R) increases from 0.85 to 0.87 and the mean absolute error (MAE) decreases from 4.51 to 2.62 K (Figure 3a). For temperature distribution within cities, the USURF_UHR run (middle row in Figures 3b–3d) reproduces LST patterns more accurately and yields lower MAE than CTRL_UHR compared to MODIS observations (bottom row in Figures 3b–3d) across large metropolitan areas such as Chicago, Boston and Miami.

The degree of improvement varies by regional background climate (Figure 4). We conducted a more comprehensive assessment against MODIS LST across 464 CONUS cities within the U.S. Census Bureau boundaries (2024), where only cities with more than 3 × 3 valid pixels (approximately covering 100 km^2) were included to allow for a statistically meaningful correlation coefficient. In general, humid (195 cities) and temperate (208 cities) regions show enhanced R values and reduced MAEs. Semi-arid regions (40 cities) demonstrate modest improvement with correlation distributions shifting toward higher values. However, arid regions (21 cities) exhibit degraded model performance when incorporating the new surface data set. In arid cities, U-Surf shows larger MAE than CTRL runs. Because CLMU does not explicitly represent irrigation in urban areas (Demuzere et al., 2014; K. Zhang et al., 2023), the mismatches between the satellite-detected urban green space and

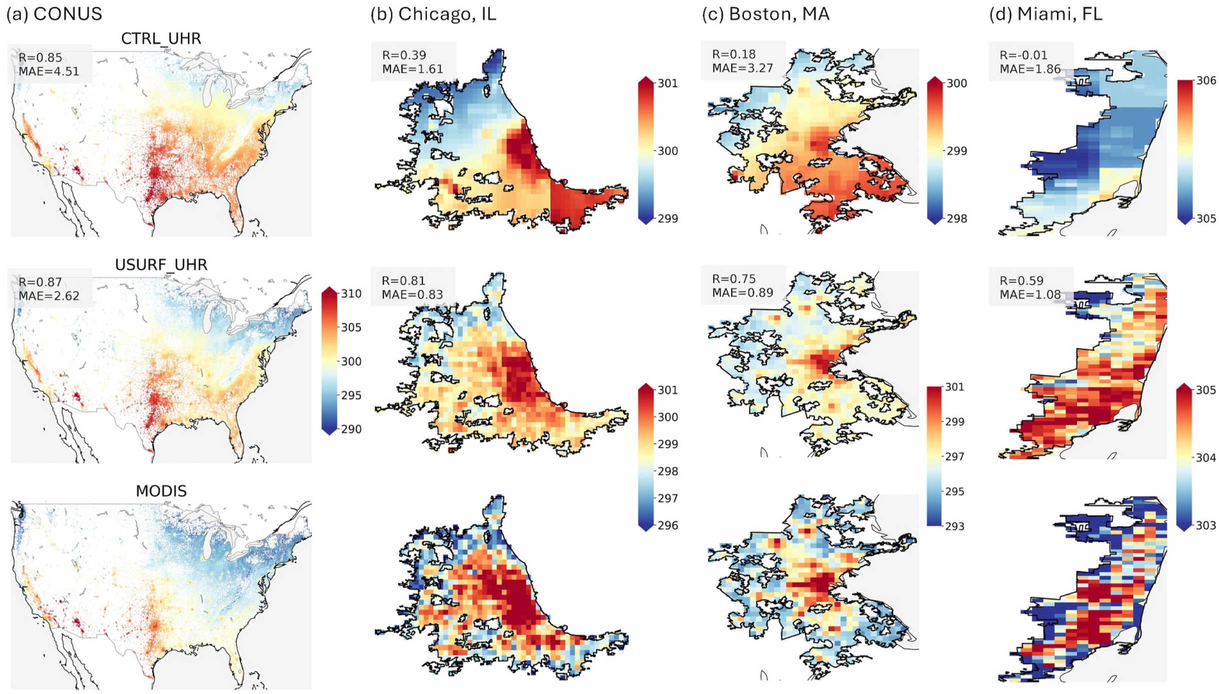

Figure 3. Five-year (2010–2014) summertime average urban land surface temperature comparison between CTRL (top row), USURF (middle row) land-only CESM simulations and MODIS observations (bottom row) at 0.03125° × 0.03125° resolution. The performance metrics correlation coefficient (R) and mean absolute error (MAE) [K] are demonstrated in (a) CONUS and selected U.S. cities: (b) Chicago, IL, (c) Boston, MA, (d) Miami, FL, only using common urban pixels. The city boundary shapefiles are from the U.S. Census Bureau (2024). Panel (a) shares a common colorbar. In panels (b–d), CTRL uses separate colorbars due to its narrower spatial distribution; while USURF and MODIS share the same colorbars for spatial pattern comparison.

inadequate modeling of evapotranspiration are likely to be more pronounced in arid regions, which will impact the surface energy balance and summertime LST (Lo & Famiglietti, 2013; Vahmani & Hogue, 2015).

In addition to LST, we examined daily average (Figure 5a), minimum (Figure 5b) and maximum (Figure 5c) near-surface (2-m) air temperature across 10 CONUS cities. In general, the USURF run produces lower median values than the CTRL run for average (T_a) and maximum air temperatures ($T_{\max}$) in all cities and for minimum air temperatures ($T_{\min}$) in 7 out of 10 cities. Mean differences range from −0.151 K for $T_{\min}$ to −0.447 K for $T_{\max}$. This pattern matches the overall cooling effect observed in both urban LST (Figure 3a, Figure S8a in Supporting Information S1) and near-surface air temperature (Figure S8b in Supporting Information S1). Additionally, we note that the updated urban representation leads to a stronger reduction in $T_{\max}$ than in $T_{\min}$ in all cities except for Chicago, resulting in a decreased diurnal temperature range with average value of −0.295 K (Figure S9 in

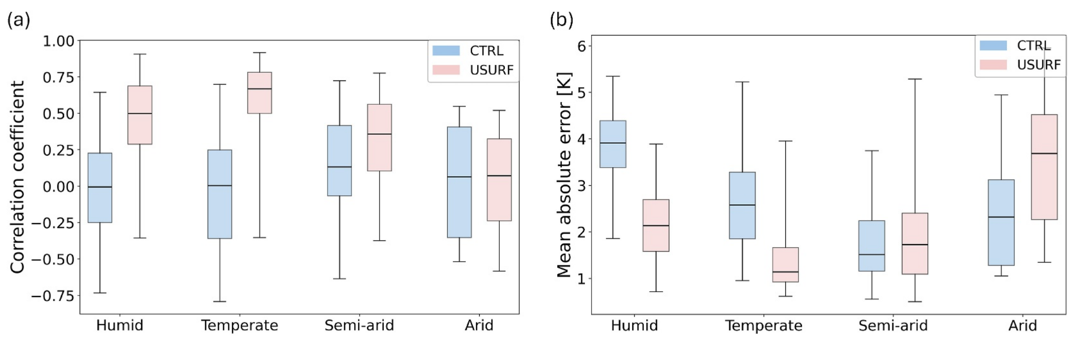

Figure 4. Five-year (2010–2014) summertime average land surface temperature (LST) correlation coefficient and mean absolute error (unit: K) validated against MODIS LST over 464 CONUS cities as defined in the U.S. Census Bureau (2024). Box and whisker plots show the 25th percentile, the median, and the 75th percentile (the bottom, middle, and top horizontal bars) and extend to the 1st and 99th quantiles.

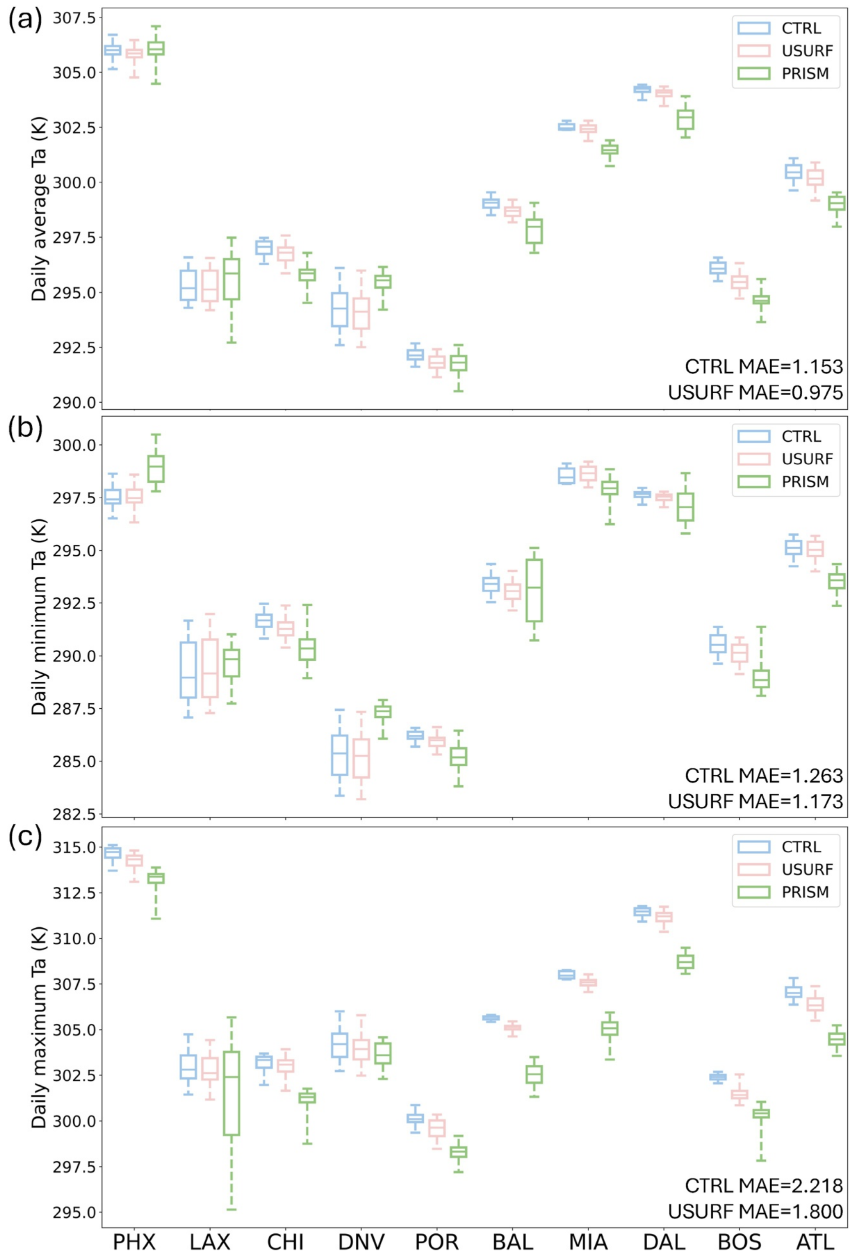

Figure 5. Distribution of 5-year (2010–2014) summertime average daily (a) average, (b) minimum and (c) maximum near-surface air temperature (T_a) for 10 major CONUS cities: Phoenix, AZ (PHX); Los Angeles, CA (LAX); Chicago, IL (CHI); Denver, CO (DNV); Portland, OR (POR); Baltimore, MD (BAL); Miami, FL (MIA); Dallas, TX (DAL); Boston, MA (BOS); Atlanta, GA (ATL). Numbers on the bottom right corner stand for the city average mean absolute error (MAE) comparing USURF and CTRL simulations with PRISM data set as a reference, respectively. Box and whisker plots show the 25th percentile, the median, and the 75th percentile (the bottom, middle, and top horizontal bars) and extend to the 1st and 99th quantiles.

Supporting Information S1). This behavior indicates that changes in urban surface properties have a stronger impact on near-surface air temperature through radiation and convection processes during daytime, suppressing the peak heating. At night, this effect is weaker, likely due to the reduced role of solar radiation-driven processes.

When comparing the modeled results with the PRISM data set (PRISM Group, 2014) as a reference, the USURF run reduces the daily air temperature MAEs within cities by approximately 0.18 K relative to the CTRL run (Figure 5a). The reductions of 0.18 K in daily average air temperature MAE and 1.89 K in LST MAE are

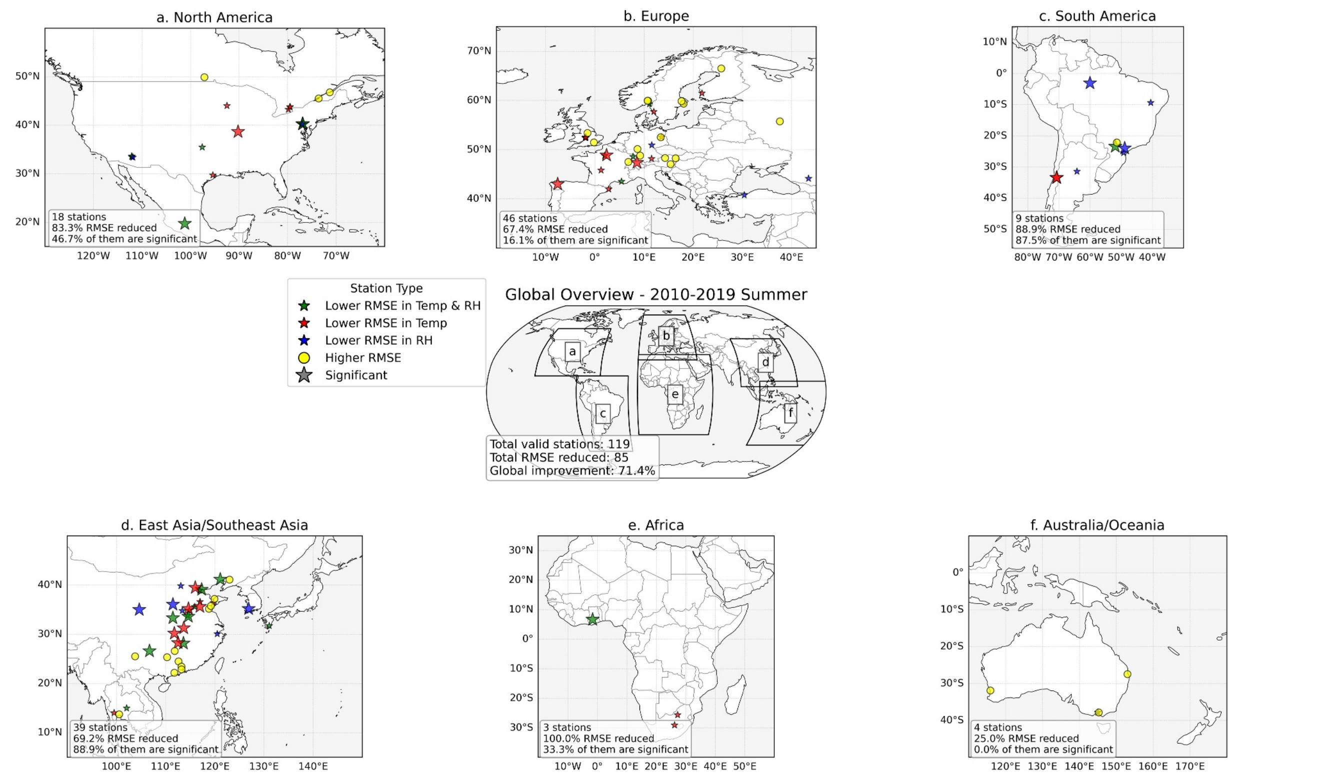

Figure 6. Improved summertime model performance for daily average near-surface air temperature and relative humidity at 119 urban stations during 2010–2019, shown across six continents: (a) North America, (b) Europe, (c) South America, (d) East/Southeast Asia, (e) Africa, and (f) Australia/Oceania. The modeled results are from global 0.23° × 0.31° land-only CESM simulations as described in Table 2. The improvement is demonstrated by the reduction in the root mean squared error (RMSE) in this plot.

consistent with findings from previous studies, which demonstrated that LST changes typically exhibit greater spatial variability (Venter et al., 2021) and can sometimes be higher than changes in near-surface air temperature in urban areas by an order of magnitude (M. Du et al., 2024). Nevertheless, we note that both the PRISM data set and model simulations have limitations that contribute to the model-observation mismatches in the diurnal cycle amplitude and phase. The PRISM data is a spatially interpolated product based on weather station measurements and therefore is largely biased toward "non-urban" areas, underestimating urban signals (Krayenhoff et al., 2018). This explains why the PRISM air temperatures are lower than modeled results across almost all cities. The model overestimation of air temperatures is systematically discussed together with other limitations in Section 4.5. Note that recent km-scale coupled simulations using U-Surf have also shown improvements in estimated urban air temperature and urban canopy heat island intensity when compared with an urban-resolving benchmark (Chakraborty et al., 2025).

At global scale, the model performance improvements are also demonstrated when validated against observations from the aforementioned urban weather stations (Section 3.1). We compare the daily average near-surface air temperature and RH from the CTRL_MR and USURF_MR simulations over the period 2010–2019, focusing on the summertime (Figure 6). In general, USURF runs better capture the daily distributions of both urban air temperature and RH based on the Kolmogorov–Smirnov (KS) test (Figure S10 in Supporting Information S1). We further evaluated the simulation results based on the coefficient of determination (R^2) and the root mean squared error (RMSE) on the daily scale. Compared to KS test, R^2 and RMSE are generally stricter measures, as they require the model to reproduce observed day-to-day variability. Current ESMs often struggle in this regard due to their reliance on a simplistically "parameterized," rather than "resolved," urban boundary layer, especially those surface- and canopy-layer processes. However, even under these stricter metrics, our results still demonstrate marked improvements in the USURF runs relative to the CTRL runs. Here we present the RMSE results in the main text (Figure 6), and the R^2 results are provided in Figure S11 in Supporting Information S1. We find that 85 out of 119 stations (71.4%) show improvement in at least one variable during summertime. This finding

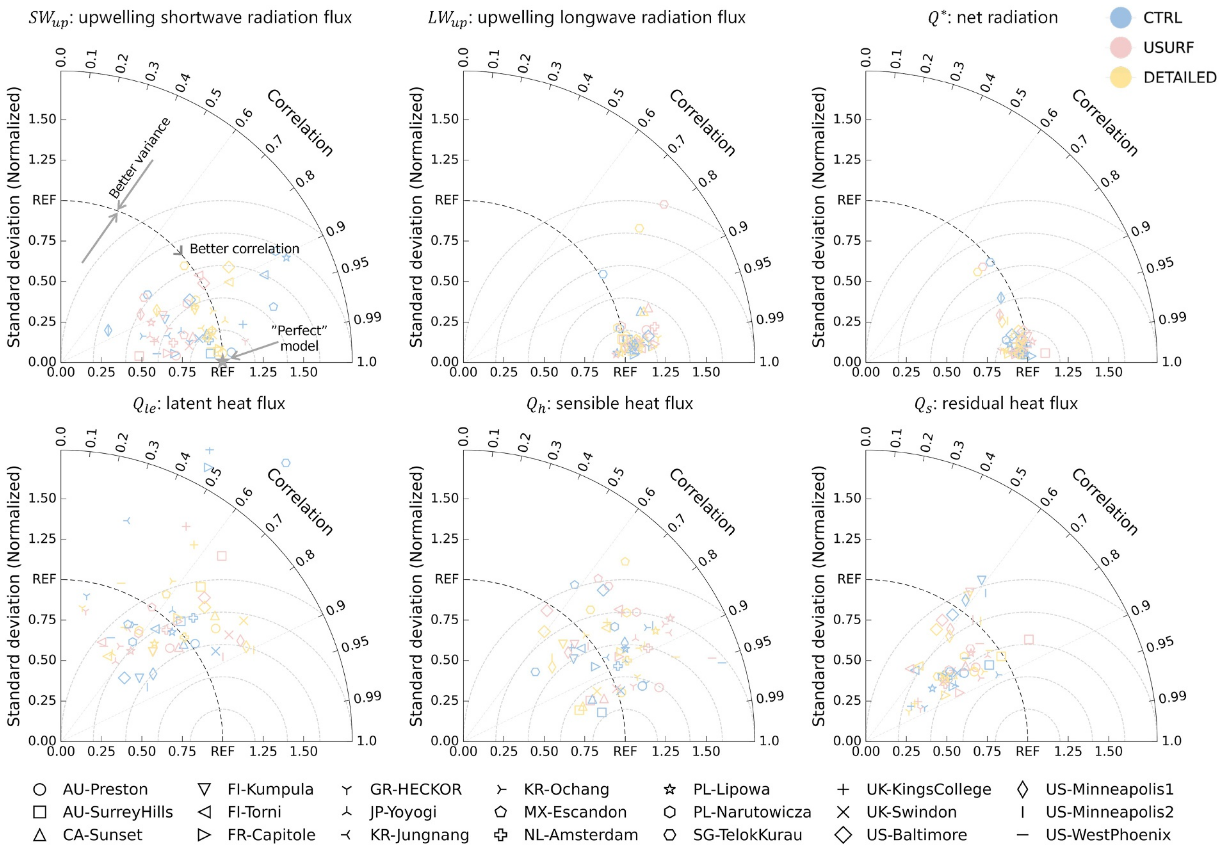

Figure 7. Taylor diagram comparing modeled results from the SP_CTRL_UHR (blue points), SP_USURF_UHR (red points), and DETAILED (yellow points) simulations with observations at 21 Urban-PLUMBER sites. Each data point with different shape represents an individual site. Models with better performance are closer to the gray star at the bottom of the diagram.

demonstrates that even when high-resolution urban inputs are aggregated to coarser resolutions, enhanced urban representation could still maintain sufficient detail to capture urban meteorological signals. Across regions with more dense urban station coverage, including North America, Europe and East/Southeast Asia, most sites exhibit lower RMSE in at least one of the variables (2 m air temperature, RH). North America and East/Southeast Asia improve at 83.3% (15/18) and 69.2% (27/39) of stations, with 46.7% and 88.9% of these improvements statistically significant. Europe also has a relatively high improvement rate (67.4%, 31 of 46 stations) though a lower significance ratio (16.1%). In sparsely sampled regions including South America, Africa and Oceania, improvement rates are 88.9%, 100.0% and 25.0%, respectively, with South America showing the highest fraction of significant improvements (87.5%).

Among the 85 stations that demonstrated improved performance, temperature enhancements occur more frequently than those in RH. Specifically, 26 stations show improvements in both variables, 42 stations improve only in temperature, and 17 stations improve only in RH. This distribution indicates that temperature simulations benefit more extensively from enhanced urban property representation. Climate models typically generate larger biases and variability when simulating humidity and precipitation compared to temperature (Randall et al., 2007). More specifically, in CLMU, RH is calculated as a function of near-surface air temperature and specific humidity (Text S3 in Supporting Information S1). Thus, when temperatures are overestimated, RH will be systematically underestimated. In station measurements, local factors such as moisture availability and measuring uncertainties can decouple the temperature-relative humidity dependency, which explains the asymmetrical performance in the improvements.

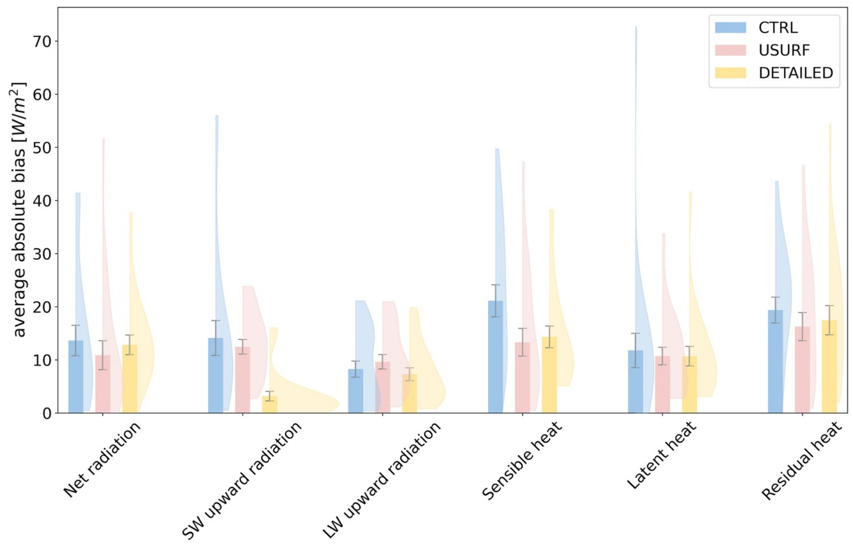

Figure 8. Comparison of modeled and observed urban energy fluxes. Bars show the average absolute difference between simulations and observations, calculated over the available observational period and averaged across 21 Urban-PLUMBER flux tower sites. The KDE curves above the bars represent the distribution of absolute biases across individual sites.

4.2. Reduced Biases in Surface Energy Fluxes

In this section, we implement additional point-scale simulations, referred to as "DETAILED," to evaluate the simulated urban surface energy fluxes against tower observations. The "DETAILED" configuration employs a set of hybrid parameters that combines morphological parameters and albedo values directly from the Urban-PLUMBER site metadata, and emissivity and thermal parameters from Oleson and Feddema (2020).

Compared to the CTRL runs, the USURF simulations demonstrate limited improvement in capturing the short-term variability of near-surface variables at the daily scale (averaged across available time steps), as observed across 21 Urban-PLUMBER sites (Figure 7). The performance enhancements vary between surface flux variables. For a perfect model, Pearson correlation coefficient (R) should be 1, the distance (d_σ) between normalized standard deviation (σ) and the ideal value 1, and normalized centered-root-mean-square error (cRMSE) should both be 0. Thus, improvement is demonstrated as decrease in d_σ and cRMSE, along with increase in R. Net radiation (Q^*) has marginal yet consistent improvements across 55% of sites. When comparing USURF to CTRL runs, on average, Q^* shows a 0.3% decrease in d_σ, 0.4% increase in R and 1.3% reduction in cRMSE. However, the rest of the flux terms show mixed results. Upwelling shortwave radiation (SW_{up}) demonstrates a 2.5% reduction in cRMSE and 0.2% increase in R despite a 16.2% increase in d_σ. Upwelling longwave radiation (LW_{up}) shows a 34.0% reduction in cRMSE but with 5.2% increase in d_σ and 0.5% decrease in R. Similarly, latent heat flux (Q_{le}) shows a 3.4% decrease in d_σ, a 2.5% reduction in cRMSE yet lower correlation coefficients decreasing from 0.630 to 0.606 (3.7%). The residual flux (Q_s) exhibits similar patterns with a 4.4% decrease in d_σ, a reduction of 1.1% in cRMSE, yet a slight reduction in R (0.4%). Sensible heat flux (Q_h) has the poorest performance with a 8.8% increase in d_σ, a 13.3% increase in cRMSE and 2.4% decrease in R.

The DETAILED simulations, despite incorporating the landscape descriptions directly from the site metadata, do not improve substantially beyond either CTRL or USURF run. On the contrary, it also demonstrates mixed behaviors in the above three metrics. The modeled fluctuations at this high temporal resolution remain inconsistent with observations, indicating that the urban surface property representation alone is not sufficient to resolve short-timescale dynamics. Additionally, accurate simulations of short-term flux dynamics might require better representation of other input variables such as precipitation, soil states and vegetation characteristics (Lipson et al., 2024) as well as urban-specific physics/processes.

When evaluated using long-term averages, USURF runs demonstrate a clear reduction in average absolute bias across multiple energy flux components, except for upward longwave radiation (Figure 8; simulation results by site are shown in Figure S12 in Supporting Information S1). The remaining five flux terms (Q^*, SW_{up}, Q_h, Q_{le},

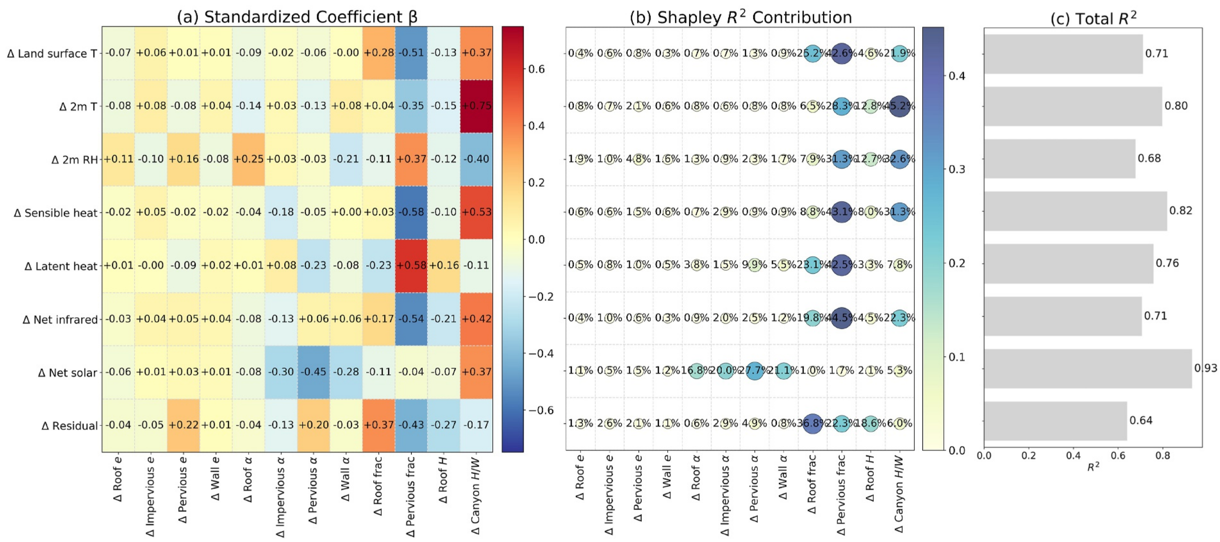

Figure 9. Sensitivity and contribution of urban canopy parameter (UCP) to simulated changes in surface meteorology and energy fluxes. Panel (a) and (b) show the standardized coefficients β and the Shapley R^2 contribution of explained variance from each parameter, respectively. The sum of each row is equal to 100% in panel (b). Panel (c) presents the total R^2 explained by all UCPs differences combined. Results are derived from the pair of 0.03125° × 0.03125° land-only simulations over CONUS, as described in Table 2.

Q_s) exhibit reductions of 20.53%, 11.81%, 36.94%, 9.14% and 14.47% in average absolute bias compared to CTRL runs. In addition, the kernel density curves in the USURF runs also shift toward lower bias values and show a narrower distribution than CTRL runs. This convergence suggests enhanced model stability and reduced uncertainty in long-term large-scale energy balance calculations. Notably, USURF runs show comparable performance to DETAILED runs across most energy flux terms, indicating that the satellite-derived urban surface properties in U-Surf data set can achieve accuracy levels approaching parameters obtained through direct field measurements or site-specific calibration. More accurate modeling of energy partitioning has important implications for land-atmosphere coupling and boundary layer dynamics (Abbas et al., 2024; Findell et al., 2024). For instance, a higher Bowen ratio, indicating that more energy is channeled into sensible heat flux, could lead to deeper boundary layer growth and higher mixing layer heights (Lin et al., 2022), which could modify regional circulation patterns, cloud formation and precipitation (Mejia et al., 2024; Sui et al., 2024). Thus, the systematic bias reduction across energy flux components suggests that incorporating high-resolution, accurate urban surface properties could substantially improve models' performance in representing urban climate impacts and feedbacks. Note that the impacts of accurate urban representation shown here come from land-only simulations. These effects should also alter land-atmosphere interactions in coupled runs.

4.3. Contributions of Input Parameters to Simulated Changes

To further examine how changes in urban canopy parameters regulate the surface and near-surface meteorology and energy fluxes in the model, we analyzed the pair of UHR simulations over CONUS (0.03125° × 0.03125°) using 5 year annual average values. We extracted differences in the input parameters and modeled variables, and restricted the analysis to grid cells located within major urban areas defined by the U.S. Census Bureau (2024) to fit a multivariate linear regression model (Figure 9). The total explained variance (Figure 9c) indicates that differences in UCPs account for a large portion of the modeled variability. Standardized coefficients indicate that changes in both pervious surface fraction and canyon height-to-width ratio exhibit strong control over surface energy partitioning, but through distinct mechanisms. Increase in pervious area reduces sensible heat flux and enhances latent heat flux substantially, with corresponding cooling effects on surface and 2 m-height air temperature. In contrast, increase in canyon height-to-width ratio alters the radiative and aerodynamic structure of the urban canopy, enhancing radiation trapping and reducing ventilation, which tends to channel more energy into sensible heat flux, leading to warmer near-surface air temperature. Changes in albedo primarily affect net shortwave radiation by altering the solar absorption and reflection, among which pervious albedo contributes most strongly. Emissivity shows limited direct influence on longwave radiation, indicating that variations in

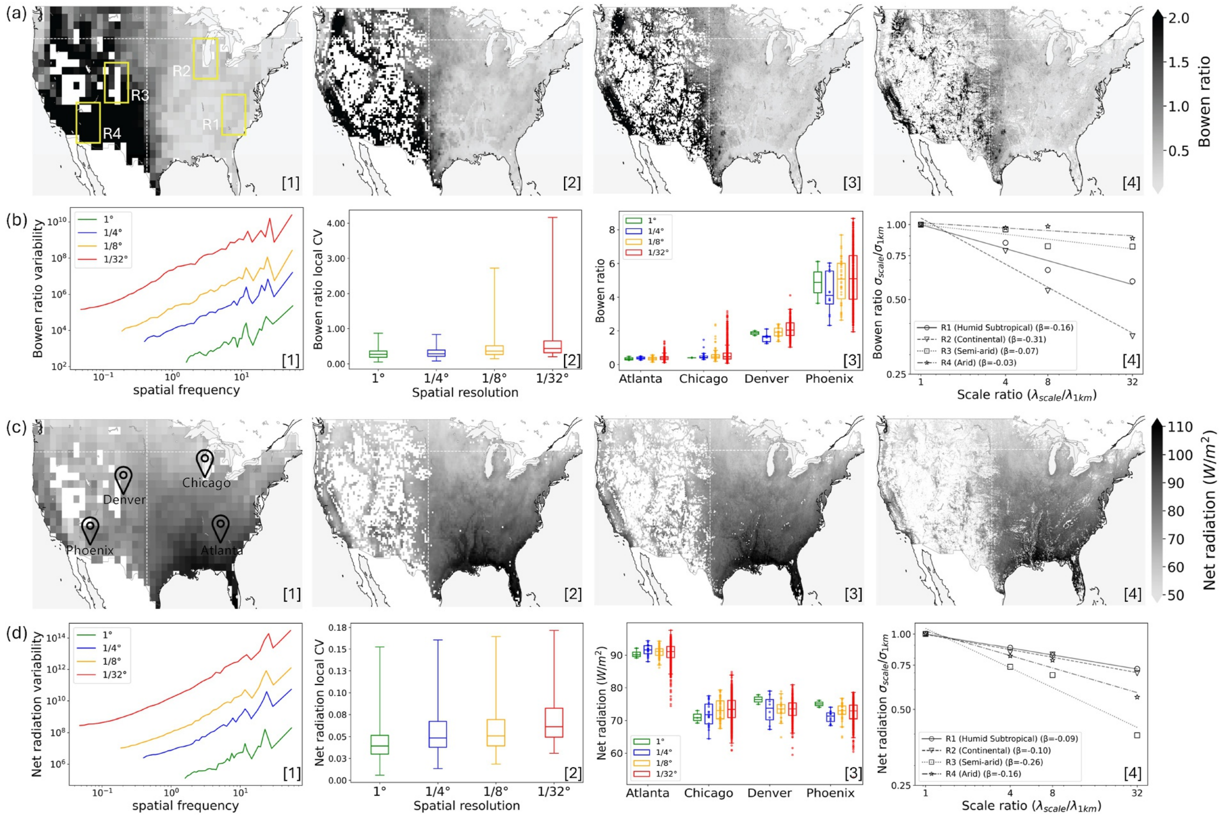

Figure 10. 2010–2014 annual mean (a) Bowen ratio and (c) net radiation from USURF_ simulations at four spatial resolutions: nominal 1°, 1/4°, 1/8°, and 1/32°, shown from left to right. Panels (b) and (d) present multi-aspect analyses of Bowen ratio and net radiation across scales respectively, including [1] Fourier power spectrum and [2] local coefficient of variation (CV) over CONUS, [3] spatial distribution in four representative cities selected in four labeled 5° × 5° regions (R1-R4), [4] information loss (log-scale) in above regions. Colors indicate different spatial resolutions in [1–3], marker shapes represent different regions in [4].

longwave radiation are dominated by changes in surface temperature rather than emissivity itself. The Shapley variance decomposition further confirms these patterns by quantifying the relative contributions of each individual UCP to the explained spatial variance of modeled responses (Figure 9b).

4.4. Impacts of Spatial Resolution

In the spatial scaling analysis, we select two representative surface energy flux variables, Bowen ratio and net radiation, to illustrate these effects. The consistent atmospheric forcing across all four resolutions ensures that large-scale spatial patterns for both variables are preserved across scales (Figures 10a and 10c), thus, we can isolate and examine the specific impacts of spatial upscaling.

Analyses across the entire CONUS region show consistent patterns in Fourier power spectrum (Figures 10b1 and 10d1) and local coefficient of variation distributions (Figures 10b2 and 10d2)]. The power spectrum curves shift upward (from green to red curves) with increasing spatial resolution, particularly at higher spatial frequency ($\geq 10^{-1}$), indicating that finer-resolution simulations can substantially capture more small-scale variability for both Bowen ratio and net radiation. At the same time, higher resolution simulations produce increased coefficient of variation for both variables. Bowen ratio, which responds more sensitively to surface property heterogeneity, also shows wider spreads at higher resolution, indicating higher local contrast in the energy partitioning between surface sensible and latent heat fluxes. In contrast, net radiation is more spatially smooth, displaying less variation in the spread of local CV. We also examine the spatial distributions of Bowen ratio and net radiation in four representative cities: Atlanta, GA; Chicago, IL; Denver, CO; and Phoenix, AZ, located in four climate regions: R1 (humid subtropical), R2 (continental), R3 (semi-arid) and R4 (arid). In general, the intra- and inter-city differences are more pronounced in higher resolution simulations (Figures 10b3 and 10d3). Notably, the median and extreme values within the same city vary substantially across spatial scales. This scale-dependent variation

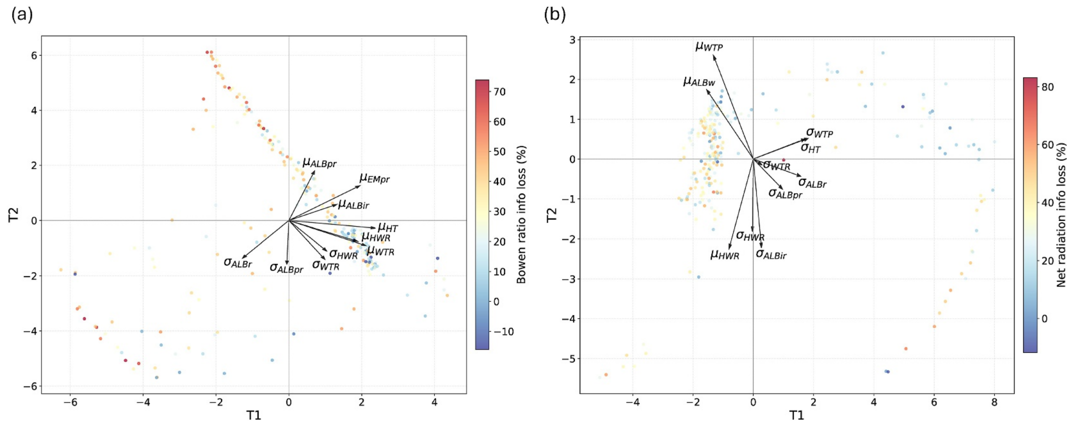

Figure 11. Supervised principal component analysis (sPCA) biplot identifying the top 10 most important physical drivers (urban surface properties) underlying the spatial information loss across urban pixels over CONUS. BR stands for Bowen ratio, NR represents net radiation. σ represents standard deviation and μ represents mean value. Urban properties include emissivity (EM), albedo (ALB), building height (HT), canyon height-to-width ratio (HWR), roof fraction (WTR), pervious fraction (WTP). Subscripts denote urban facets: roof (r), impervious canyon (ir), pervious canyon (pr), wall (w). Displayed arrows are selected using the Variable Importance in Projection, while arrow direction and length are based on sPCA loadings. Arrow length indicates the magnitudes of each driver's contribution to principal components (or latent vectors; T_1, T_2), while arrow direction shows the positive or negative correlation between drivers and PCs. Color of the data points indicate the information loss due to spatial aggregation averaged across resolutions.

introduces large uncertainties when validating coarse-resolution model outputs against point observations such as those from weather stations. The large cross-scale discrepancies observed in surface energy fluxes under land-only configurations suggest that the spatial resolution effects could be amplified in coupled simulations where land-atmosphere feedbacks are taken into account.

The information loss patterns for Bowen ratio and net radiation vary across four climate regions R1-R4. We find strongest resolution-dependency (most negative slope) of Bowen ratio and net radiation in continental ($\beta = -0.31$) and semi-arid ($\beta = -0.26$) regions, respectively. Compared to the finest resolution of 1/32°, the information loss γ of Bowen ratio and net radiation reaches up to 64.5% and 60.5% at 1 deg, respectively (Figures 10b4 and 10d4).

The CONUS-wide sPCA analysis reveals the dominant UCP drivers associated with the spatial information loss in Bowen ratio (Figure 11a) and net radiation (Figure 11b). For the Bowen ratio, the first latent variable T_1 aligns primarily with mean roof fraction (μ_{WTR}), roof height (μ_{HT}) and canyon height-to-width ratio (μ_{HWR}), and they also carry the largest VIP scores (Text S4 in Supporting Information S1), indicating that urban morphology dominates the scaling information loss in the energy partitioning between latent and sensible heat. The close alignment of μ_{WTR} and σ_{WTR} also suggests that urban pixels with higher roof coverage tend to exhibit more spatial heterogeneity, which co-vary and jointly influence the information loss. The variance in roof (σ_{ALBr}) and pervious albedo (σ_{ALBpr}) are displayed in different directions, meaning that albedo heterogeneity operates separately from morphology-driven effects. As for the net radiation, the top VIP variables (contributors) become pervious canyon floor fraction (μ_{WTP}) and variance in building height (σ_{HT}) and albedo values (σ_{ALBpr}; σ_{ALBr}). Several urban surface properties cluster in similar directions, indicating that the spatial heterogeneity in morphology and albedo combined together is central to the information loss. The radiative budget appears more sensitive to sub-grid variation among different urban facets, including roof, wall, impervious and pervious canyon floor.

In addition to the regional analysis across scales, the point-scale simulation analysis also reveals complex scale-dependent responses in surface energy fluxes that have significant implications on land-atmosphere interactions. We find the non-monotonic responses of the average mean absolute error to the changes in spatial resolution (Figure 12; simulation results by site are shown in Figure S13 in Supporting Information S1), indicating that enhanced urban property representations do not uniformly improve all aspects of surface energy representation. Sensible heat, latent heat and stored plus anthropogenic heat show a more linear relationship between bias and resolution, that is, a systematic improvement with higher spatial resolution. However, this might be a result of land-only simulation configurations, without resolved vertical mixing, convection and advection. The radiation

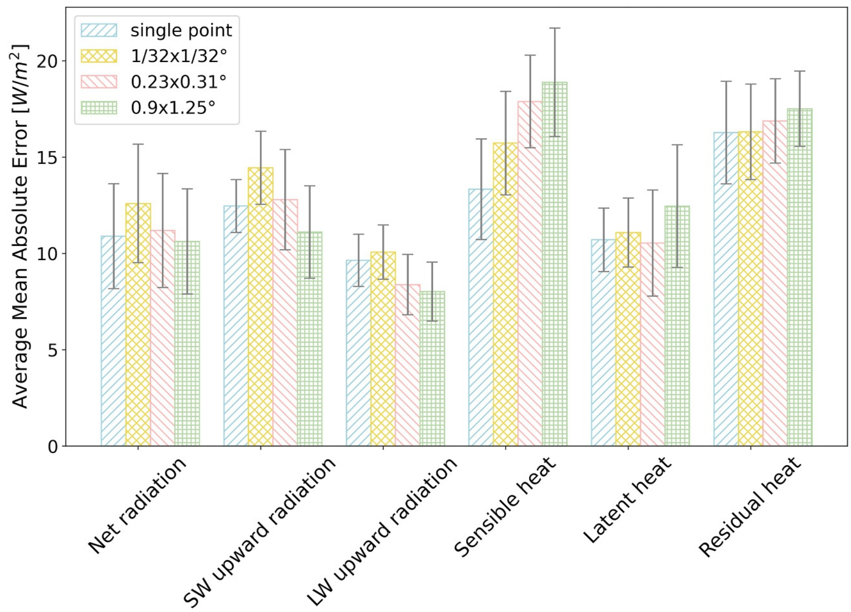

Figure 12. Inter-resolution comparison of modeled and observed urban energy fluxes using U-Surf based urban parameters. Bars show the average absolute difference between USURF simulations and Urban-PLUMBER flux tower observations, calculated over the available observational period and averaged across 21 sites using four spatial scales described in Table 3.

terms exhibit more irregular patterns (Figure 12), which points to more complex interactions between surface properties and radiative transfer processes. These non-monotonic responses may reflect the tradeoff effects in how the radiative and morphological parameters modulate the radiative absorption and reflection processes. For example, for building surfaces with higher albedo values due to the installation of reflective materials, the reflected solar radiation could still be trapped in the deep canyon due to a high canyon height-to-width ratio. Another important source of uncertainty arises from the scale mismatch between the turbulence and radiative characteristics at those local sites. Both U-Surf based and Urban-PLUMBER site characteristics are intended to represent the turbulent flux footprint, which often reflects a relatively large spatial average (radius ∼500 m) that varies with wind direction and atmospheric dispersion. In contrast, the radiation fluxes typically represent a smaller local area centered around the tower (Lipson et al., 2022, 2024; Sailor, 2011).

4.5. Discussion

Our results demonstrate that improved UCPs can, in general, enhance CESM's performance in simulating urban surface meteorology and energy fluxes. This enhancement has multiple critical implications for ESMs' urban-resolving capabilities. First, the overall enhanced capability to capture urban temperature and humidity signals in our results indicates enhanced model capability in assessing urban heat stress and its attribution to local urban contribution. This is important for both accurate urban climate impact assessments (e.g., on public health and energy) and effective urban adaptation/mitigation planning. Second, better-resolved intra-urban heterogeneity enables better understanding, modeling, and prediction of different environmental stressors across communities (Chakraborty et al., 2023), which is essential for guiding targeted urban planning and equitable climate adaptation strategies. Third, incorporating high-resolution, spatially-explicit, and realistic representation of urban canopy parameters, such as roof albedo and emissivity, roof fraction, and pervious fraction, enables enhanced model capacity to test more realistic and novel urban climate strategies such as albedo management (Ouyang et al., 2022; Sun et al., 2024), urban greenery (Alexander et al., 2024; J. Chen et al., 2024), and new building materials (K. Zhang et al., 2025). Last, accurate, kilometer-scale urban surface constraints allow ESMs to resolve finer-scale processes, especially over the urban areas with complex terrain features (e.g., coastal or mountainous regions), as models are moving toward kilometric resolutions. For example, the much smoother gradients shown in USURF runs compared to CTRL runs in coastal urban areas (Figure 3d) can lead to distinct feedbacks in land-atmosphere(-ocean) coupled simulations, which will have important implications on cloud and precipitation formation, and heatwave propagation.

On the other hand, we note that the performance in model improvement varies across regions and in some occasions large discrepancies with observations still exist. These limitations can likely be attributed to three primary factors.

First, CLMU's multiple parameterizations were originally calibrated toward the default urban surface biophysical input data, potentially limiting its performance with new high-resolution input data that capture more distinct and detailed surface properties. Enhanced urban representation without additional model optimization/calibration therefore does not guarantee improved model performance and future work needs to be done on systematic model calibration and parameter tuning to fully exploit the capabilities of the high-resolution urban input data set.

Second, we acknowledge the limitations of current urban process parametrization in CLMU. Urban vegetation is known to significantly modulate both absolute magnitude and seasonal variability in urban climate responses (Chakraborty & Lee, 2019a), yet remains absent from current CLMU and most operational regional climate models. Explicit representation of urban vegetation structure and function could substantially improve the modeling of energy partitioning in urban environments, as well as temperature and humidity estimates (Chakraborty et al., 2022). Additionally, current urban thermal parameters rely on outdated categorical values that inadequately represent modern building materials and construction practices. However, improving parameters alone does not guarantee better performance (Grimmond et al., 2011) as CLMU represents heat conduction using a one-dimensional form based on Fourier's Law (Oleson & Feddema, 2020; Oleson et al., 2010), which does not resolve three-dimensional heat pathways in real-world buildings. Therefore, systematic parameter optimization in the context of the simplified building energy model would potentially be a pathway to improve the model performance. In the longer term, advances in the heat transfer scheme beyond the 1D treatment along with improved realistic local building thermal properties will be needed to better support the modeling of building energy and anthropogenic heat fluxes, which could improve the diurnal cycle simulation accuracy. The current urban property input data is temporally static, not taking land use and land cover change or evolving urban morphology into consideration. Incorporating temporally dynamic urban properties would enable investigation of urban-rural differences over time and improve representation of seasonal urban climate responses, which will become increasingly important for long-term climate projections under urbanization scenarios. Finally, there are uncertainties in the satellite-derived algorithms and products incorporated into U-Surf, especially when it comes to resolving urban signals. Improving these requires new algorithm development from the remote sensing community and better urban-scale data sets to calibrate/benchmark them.

Third, while the observations and reanalysis data used in this study provide valuable references to evaluate the model performance on resolving urban-scale signals, it is challenging to obtain a "perfect" urban-resolving benchmark for model validation. For example, satellite-derived LST product can differ from modeled skin temperature due to the thermal anisotropy (Krayenhoff & Voogt, 2016), where urban thermal anisotropy could bias LST by up to 5 K (H. Du et al., 2023). The urban signals in PRISM air temperature data could be diluted by the dominance of non-urban weather stations, and thus are underestimated compared to modeled results.

5. Conclusions

Despite the emerging importance of resolving urban areas for high resolution ESM simulations, both parametric and process simplifications in such models significantly limit this capability. Here, we develop a spatially continuous urban-resolving modeling framework by employing the newly developed U-Surf data set to replace traditional categorical mapping of urban canopy parameters with high-resolution and observationally constrained urban surface properties. This advancement addresses uncertainties associated with unresolved urban heterogeneity and enables the direct incorporation of realistic urban characteristics even at convection permitting scale, filling a key gap in current kilometer-scale Earth system modeling.

Validation against satellite observations, reanalysis data, weather stations, and flux tower measurements confirmed that this enhanced representation of urban properties allows for better capture of urban meteorology, both in magnitude and spatial variation. The U-Surf simulation demonstrates systematic bias reductions of up to 1.89 K for LST and 0.18 K for near-surface air temperature across diverse climate regimes. Enhanced performance also extends to RH representation across urban weather stations as well as the multi-year averages of important surface energy balance components. Finally, we perform a spatial scaling analysis across scales, which reveals valuable insights into how urban processes depend on model resolution and the physical mechanisms underlying scale-dependent responses. We demonstrate that information loss during spatial aggregation varies across background climate regimes, where the variability in radiative properties accounts for the majority of sensitivity in humid and temperate regions, while urban morphology dominates in affecting responses in semi-arid and arid regions. The climate regime-dependency of scale sensitivity further confirms that uniform parameterizations may inadequately represent diverse urban environments and scale-aware urban representations that account for regional climate controls are critical for high-resolution modeling.

This study establishes a clear pathway for integrating detailed urban surface representation into global-scale climate models and enables global-scale urban weather and climate experiments and systematic intra- and inter-city comparisons. While our study uses land-only simulations, we note that the improved property implementation would also impact coupled simulations, where urban effects could propagate beyond city boundaries through land-atmosphere interactions. As global climate models and ESMs advance toward kilometer-scale simulations, these high-resolution urban surface properties can be seamlessly integrated into land models while coupled with more advanced atmosphere modeling infrastructure to capture urban feedback. This urban-resolving framework represents a major advancement that will transform our ability to simulate and understand urban processes across scales in global ESMs, enabling next-generation urban climate modeling, and thus better science-based tools, to understand an increasingly urbanizing world.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The CTSM code (version tag ctsm5.2.015) used in this study is publicly available at CTSM Development Team (2024). The simulations are performed using the Derecho system managed by the NSF National Center for Atmospheric Research (Computational and Information Systems Laboratory, 2023). The bias-corrected ERA5 atmospheric forcing is publicly available at Cucchi et al. (2020). The U-Surf data set is publicly available at Cheng et al. (2024). The analysis and plotting scripts are publicly available in the author's GitHub directory: Cheng et al. (2026). Simulation outputs, post-processed data and validation data sets are publicly available at Cheng, Zhao, Oleson, et al. (2025).

Acknowledgments
L.Z., Y.C. and J.Y. acknowledge the support by the U.S. National Science Foundation (CAREER Award Grant 2145362) and the Institute for Sustainability, Energy, and Environment at the University of Illinois Urbana-Champaign. L.Z. and Y.C. acknowledge the high-performance computing support from the Derecho system (https://doi.org/10.5065/qx9a-pg09) provided by the NSF National Center for Atmospheric Research (NCAR), sponsored by the National Science Foundation. L.Z. and T.C. also acknowledge the support by the U.S. National Aeronautics and Space Administration (NASA) through the AMT and LCLUC programs (Grant 80NSSC25K7322). T.C.'s contribution was also supported by the Earth and Environmental Systems Modeling program area of the U.S. Department of Energy (DOE) Office of Science's Biological and Environmental Research program through a DOE Early Career Award. The Pacific Northwest National Laboratory is operated for the DOE by the Battelle Memorial Institute under contract DE-AC05-76RL01830. K.O. acknowledges funding from 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. C.H. acknowledges the support by the NASA Land Cover Land Use Change (LCLUC) Grant 80NSSC25K7493 Subaward 2432603. K.Z. acknowledges support by the HMEI Environmental Fellows Program at Princeton University.

References

Abbas, S. S., Behrendt, A., Branch, O., & Wulfmeyer, V. (2024). Relationships between surface fluxes and boundary layer dynamics: Statistics at the Land-Atmosphere Feedback Observatory (LAFO). *EGUsphere*, 1–24. https://doi.org/10.5194/egusphere-2024-3878

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statistics*, *2*(1), 97–106. https://doi.org/10.1002/wics.51

Alexander, G. A., Voter, C. B., Wright, D. B., & Loheide II, S. P. (2024). Urban ecohydrology: Accounting for sub-grid lateral water and energy transfers in a land surface model. *Water Resources Research*, *60*(3), e2023WR035511. https://doi.org/10.1029/2023WR035511

Battini, F., Menapace, A., Stradiotti, G., Zanfei, A., Nicolosi, F. F., Dalla Torre, D., et al. (2025). Urban smart energy systems from a climate change perspective: Technical, economic and environmental optimization analysis. *Smart Energy*, *18*, 100180. https://doi.org/10.1016/j.segy.2025.100180

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

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., et al. (2013). The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geoscientific Model Development*, *6*(3), 687–720. https://doi.org/10.5194/gmd-6-687-2013

Caldwell, P. M., Mametjanov, A., Tang, Q., Van Roekel, L. P., Golaz, J.-C., Lin, W., et al. (2019). The DOE E3SM coupled model version 1: Description and results at high resolution. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4095–4146. https://doi.org/10.1029/2019MS001870

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., et al. (2021). Convection-permitting simulations with the E3SM global atmosphere model. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002544. https://doi.org/10.1029/2021MS002544

Chakraborty, T., & Lee, X. (2019a). Land cover regulates the spatial variability of temperature response to the direct radiative effect of aerosols. *Geophysical Research Letters*, *46*(15), 8995–9003. https://doi.org/10.1029/2019GL083812

Chakraborty, T. C., & Lee, X. (2019b). 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

Chakraborty, T. C., Lee, X., Ermida, S., & Zhan, W. (2021). 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

Chakraborty, T. C., Li, J., Li, L., Bisht, G., Leung, L. R., & Kraucunas, I. (2025). *Demonstrating use of E3SM with an urban canopy parameterization for modeling urban impact on local-to-regional heat extremes (DOE/SC-CM-25-001)*. U.S. Department of Energy, Earth and Environmental Systems Modeling Program. Retrieved from https://eesm.science.energy.gov/sites/default/files/2025-01/FY2025_1st_Quarter_Metrics.pdf

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

Chakraborty, T. C., & Qian, Y. (2024). Urbanization exacerbates continental- to regional-scale warming. *One Earth*, *7*(8), 1387–1401. https://doi.org/10.1016/j.oneear.2024.05.005

Chakraborty, T. C., Venter, Z. S., Qian, Y., & Lee, X. (2022). Lower urban humidity moderates outdoor heat stress. *AGU Advances*, *3*(5), e2022AV000729. https://doi.org/10.1029/2022AV000729

Charrondière, C., & Stiperski, I. (2024). Spectral scaling of unstably stratified atmospheric flows: Turbulence anisotropy and the low-frequency spread. *Quarterly Journal of the Royal Meteorological Society*, *150*(764), 4196–4216. https://doi.org/10.1002/qj.4811

Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., et al. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. *Nature Communications*, *11*(1), 537. https://doi.org/10.1038/s41467-020-14386-x

Chen, J., Bach, P. M., Nice, K. A., & Leitão, J. P. (2024). Investigating the efficacy of a fast urban climate model for spatial planning of green and blue spaces for heat mitigation. *Science of the Total Environment*, *955*, 176925. https://doi.org/10.1016/j.scitotenv.2024.176925

Chen, K., Boomsma, J., & Holmes, H. A. (2023). A multiscale analysis of heatwaves and urban heat islands in the western U.S. during the summer of 2021. *Scientific Reports*, *13*(1), 9570. https://doi.org/10.1038/s41598-023-35621-7

Cheng, Y., Zhao, L., Chakraborty, T., Oleson, K., Demuzere, M., Liu, X., et al. (2024). U-Surf: A global 1 km spatially continuous urban surface property dataset for kilometer-scale urban-resolving Earth system modeling (Version 1.1) [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.11247598

Cheng, Y., Zhao, L., Chakraborty, T. C., Oleson, K., Demuzere, M., Liu, X., et al. (2025). U-Surf: A global 1 km spatially continuous urban surface property dataset for kilometer-scale urban-resolving Earth system modeling. *Earth System Science Data*, *17*(5), 2147–2174. https://doi.org/10.5194/essd-17-2147-2025

Cheng, Y., Zhao, L., Oleson, K., Chakraborty, T., Zhang, K., He, C., & Yang, J. (2025). Importance of spatially continuous urban surface properties in urban-resolving Earth system modeling (Version 1.0) [Dataset]. *Figshare*. https://doi.org/10.6084/m9.figshare.29640872

Cheng, Y., Zhao, L., Oleson, K., Chakraborty, T., Zhang, K., He, C., & Yang, J. (2026). yifanc17/Urban_resolving_CESM (Version 1.1) [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.18968484

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., et al. (2019). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, *11*(1), 185–209. https://doi.org/10.1029/2018MS001369

Computational and Information Systems Laboratory. (2023). Derecho: HPE Cray EX System (University Community Computing) [Dataset]. *NSF National Center for Atmospheric Research*. https://doi.org/10.5065/qx9a-pg09

CTSM Development Team. (2024). ESCOMP/CTSM [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.3739617

Cucchi, M., Weedon, G., Amici, A., Bellouin, N., Lange, S., Schmied, H. M., et al. (2020). Near surface meteorological variables from 1979 to 2019 derived from bias-corrected reanalysis (Version 1.1) [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.20d54e34

Demuzere, M., Coutts, A. M., Göhler, M., Broadbent, A. M., Wouters, H., van Lipzig, N. P. M., & Gebert, L. (2014). 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

Demuzere, M., Oleson, K., Coutts, A. M., Pigeon, G., & van Lipzig, N. P. M. (2013). Simulating the surface energy balance over two contrasting urban environments using the Community Land Model Urban. *International Journal of Climatology*, *33*(15), 3182–3205. https://doi.org/10.1002/joc.3656

Du, H., Zhan, W., Liu, Z., Scott Krayenhoff, E., Chakraborty, T., Zhao, L., et al. (2023). Global mapping of urban thermal anisotropy reveals substantial potential biases for remotely sensed urban climates. *Science Bulletin*, *68*(16), 1809–1818. https://doi.org/10.1016/j.scib.2023.06.032

Du, M., Li, N., Hu, T., Yang, Q., Chakraborty, T., Venter, Z., & Yao, R. (2024). Daytime cooling efficiencies of urban trees derived from land surface temperature are much higher than those for air temperature. *Environmental Research Letters*, *19*(4), 044037. https://doi.org/10.1088/1748-9326/ad30a3

Findell, K. L., Yin, Z., Seo, E., Dirmeyer, P. A., Arnold, N. P., Chaney, N., et al. (2024). Accurate assessment of land–atmosphere coupling in climate models requires high-frequency data output. *Geoscientific Model Development*, *17*(4), 1869–1883. https://doi.org/10.5194/gmd-17-1869-2024

Grimmond, C. S. B., Blackett, M., Best, M. J., Baik, J.-J., Belcher, S. E., Beringer, J., et al. (2011). Initial results from Phase 2 of the international urban energy balance model comparison. *International Journal of Climatology*, *31*(2), 244–272. https://doi.org/10.1002/joc.2227

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development*, *9*(11), 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Hou, A., Ni, G., Yang, H., & Lei, Z. (2013). Numerical analysis on the contribution of urbanization to wind stilling: An example over the greater Beijing metropolitan area. *Journal of Applied Meteorology and Climatology*, *52*(5), 1105–1115. https://doi.org/10.1175/JAMC-D-12-013.1

Hsu, A., Sheriff, G., Chakraborty, T., & Manya, D. (2021). Disproportionate exposure to urban heat island intensity across major US cities. *Nature Communications*, *12*(1), 2721. https://doi.org/10.1038/s41467-021-22799-5

Jackson, T. L., Feddema, J. J., Oleson, K. W., Bonan, G. B., & Bauer, J. T. (2010). Parameterization of urban characteristics for global climate modeling. *Annals of the Association of American Geographers*, *100*(4), 848–865. https://doi.org/10.1080/00045608.2010.497328

Jiang, T., Krayenhoff, E. S., Martilli, A., Nazarian, N., Stone, B., & Voogt, J. A. (2025). Prioritizing urban heat adaptation infrastructure based on multiple outcomes: Comfort, health, and energy. *Proceedings of the National Academy of Sciences*, *122*(19), e2411144122. https://doi.org/10.1073/pnas.2411144122

Kotthaus, S., & Grimmond, C. S. B. (2014). Energy exchange in a dense urban environment – Part II: Impact of spatial heterogeneity of the surface. *Urban Climate*, *10*, 281–307. https://doi.org/10.1016/j.uclim.2013.10.001

Krayenhoff, E. S., Moustaoui, M., Broadbent, A. M., Gupta, V., & Georgescu, M. (2018). Diurnal interaction between urban expansion, climate change and adaptation in US cities. *Nature Climate Change*, *8*(12), 1097–1103. https://doi.org/10.1038/s41558-018-0320-9

Krayenhoff, E. S., & Voogt, J. A. (2016). Daytime thermal anisotropy of urban neighbourhoods: Morphological causation. *Remote Sensing*, *8*(2), 108. https://doi.org/10.3390/rs8020108

Lawrence, D., Fisher, R., Koven, C., Oleson, K., Swenson, S., Vertenstein, M., et al. (2018). *Technical Description of version 5.0 of the Community Land Model (CLM)*. NCAR Technical Note NCAR/TN-541+STR. National Center for Atmospheric Research. Retrieved from http://www.cesm.ucar.edu/models/cesm2/land/CLM50_Tech_Note.pdf

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018MS001583

Li, D., & Bou-Zeid, E. R. (2013). 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*(9), 2051–2064. https://doi.org/10.1175/JAMC-D-13-02.1

Li, L., Bisht, G., Hao, D., & Leung, L. R. (2024). Global 1 km land surface parameters for kilometer-scale Earth system modeling. *Earth System Science Data*, *16*(4), 2007–2032. https://doi.org/10.5194/essd-16-2007-2024

Li, X., Zhou, Y., Eom, J., Yu, S., & Asrar, G. R. (2019). Projecting global urban area growth through 2100 based on historical time series data and future shared socioeconomic pathways. *Earth's Future*, *7*(4), 351–362. https://doi.org/10.1029/2019EF001152

Li, X. C., Zhao, L., Oleson, K., Zhou, Y., Qin, Y., Zhang, K., & Fang, B. (2024). 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*(4), e2023MS004107. https://doi.org/10.1029/2023MS004107

Li, X. C., Zhao, L., Qin, Y., Oleson, K., & Zhang, Y. (2024). Elevated urban energy risks due to climate-driven biophysical feedbacks. *Nature Climate Change*, *14*(10), 1056–1063. https://doi.org/10.1038/s41558-024-02108-w

Liao, W., Liu, X., Li, D., Luo, M., Wang, D., Wang, S., et al. (2018). Stronger contributions of urbanization to heat wave trends in wet climates. *Geophysical Research Letters*, *45*(20), 11310–11317. https://doi.org/10.1029/2018GL079679

Lin, H., Li, Y., & Zhao, L. (2022). Partitioning of sensible and latent heat fluxes in different vegetation types and their spatiotemporal variations based on 203 FLUXNET sites. *Journal of Geophysical Research: Atmospheres*, *127*(21), e2022JD037142. https://doi.org/10.1029/2022JD037142

Lipson, M., Grimmond, S., Best, M., Chow, W. T. L., Christen, A., Chrysoulakis, N., et al. (2022). Harmonized gap-filled datasets from 20 urban flux tower sites. *Earth System Science Data*, *14*(11), 5157–5178. https://doi.org/10.5194/essd-14-5157-2022

Lipson, M. J., Grimmond, S., Best, M., Abramowitz, G., Coutts, A., Tapper, N., et al. (2024). Evaluation of 30 urban land surface models in the Urban-PLUMBER project: Phase 1 results. *Quarterly Journal of the Royal Meteorological Society*, *150*(758), 126–169. https://doi.org/10.1002/qj.4589

Lo, M.-H., & Famiglietti, J. S. (2013). Irrigation in California's Central Valley strengthens the southwestern U.S. water cycle. *Geophysical Research Letters*, *40*(2), 301–306. https://doi.org/10.1002/grl.50108

Mejia, J. F., Henao, J. J., & Eslami, E. (2024). Role of clouds in the Urban Heat Island and extreme heat: Houston-Galveston metropolitan area case. *Journal of Geophysical Research: Atmospheres*, *129*(17), e2024JD041243. https://doi.org/10.1029/2024JD041243

Moon, J.-Y., Streffing, J., Lee, S.-S., Semmler, T., Andrés-Martínez, M., Chen, J., et al. (2025). Earth's future climate and its variability simulated at 9 km global resolution. *Earth System Dynamics*, *16*(4), 1103–1134. https://doi.org/10.5194/esd-16-1103-2025

Moraglia, G., Pryor, S. C., & Crippa, P. (2024). Quantifying the impacts of an urban area on clouds and precipitation patterns: A modeling perspective. *Journal of Geophysical Research: Atmospheres*, *129*(12), e2024JD041402. https://doi.org/10.1029/2024JD041402

Moreno-Chamarro, E., Arsouze, T., Acosta, M., Bretonnière, P.-A., Castrillo, M., Ferrer, E., et al. (2025). The very-high-resolution configuration of the EC-Earth global model for HighResMIP. *Geoscientific Model Development*, *18*(2), 461–482. https://doi.org/10.5194/gmd-18-461-2025

Nastrom, G. D., Gage, K. S., & Jasperson, W. H. (1984). Kinetic energy spectrum of large-and mesoscale atmospheric processes. *Nature*, *310*(5972), 36–38. https://doi.org/10.1038/310036a0

Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., & Grimmond, C. S. B. (2008). An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. *Journal of Applied Meteorology and Climatology*, *47*(4), 1038–1060. https://doi.org/10.1175/2007JAMC1597.1

Oleson, K. W., Bonan, G. B., Feddema, J. J., Vertenstein, M., & Kluzek, E. (2010). *Technical description of an urban parameterization for the community land model (CLMU)*. NCAR Technical Note NCAR/TN-480+STR (p. 169). National Center for Atmospheric Research. https://doi.org/10.5065/D6K35RM9

Oleson, K. W., & Feddema, J. (2020). Parameterization and surface data improvements and new capabilities for the Community Land Model Urban (CLMU). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2018MS001586. https://doi.org/10.1029/2018MS001586

Ouyang, Z., Sciusco, P., Jiao, T., Feron, S., Lei, C., Li, F., et al. (2022). Albedo changes caused by future urbanization contribute to global warming. *Nature Communications*, *13*(1), 3800. https://doi.org/10.1038/s41467-022-31558-z

Oyedele, O. F., & Lubbe, S. (2015). The construction of a partial least-squares biplot. *Journal of Applied Statistics*, *42*(11), 2449–2460. https://doi.org/10.1080/02664763.2015.1043858

PRISM Group. (2014). *PRISM weather data*. Oregon State University. Retrieved from https://prism.oregonstate.edu

Qian, Y., Chakraborty, T. C., Li, J., Li, D., He, C., Sarangi, C., et al. (2022). Urbanization impact on regional climate and extreme weather: Current understanding, uncertainties, and future research directions. *Advances in Atmospheric Sciences*, *39*(6), 819–860. https://doi.org/10.1007/s00376-021-1371-9

Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R., et al. (2025). Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5 and NEMOv3.4. *Geoscientific Model Development*, *18*(1), 33–69. https://doi.org/10.5194/gmd-18-33-2025

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., et al. (2007). Climate models and their evaluation. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, et al. (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Roberts, M. J., Reed, K. A., Bao, Q., Barsugli, J. J., Camargo, S. J., Caron, L.-P., et al. (2025). High-Resolution Model Intercomparison Project phase 2 (HighResMIP2) towards CMIP7. *Geoscientific Model Development*, *18*(4), 1307–1332. https://doi.org/10.5194/gmd-18-1307-2025

Sailor, D. J. (2011). A review of methods for estimating anthropogenic heat and moisture emissions in the urban environment. *International Journal of Climatology*, *31*(2), 189–199. https://doi.org/10.1002/joc.2106

Scott, A. A., Waugh, D. W., & Zaitchik, B. F. (2018). Reduced Urban Heat Island intensity under warmer conditions. *Environmental Research Letters*, *13*(6), 064003. https://doi.org/10.1088/1748-9326/aabd6c

Simpson, I. R., Lawrence, D. M., Swenson, S. C., Hannay, C., McKinnon, K. A., & Truesdale, J. E. (2022). Improvements in wintertime surface temperature variability in the Community Earth System Model version 2 (CESM2) related to the representation of snow density. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002880. https://doi.org/10.1029/2021MS002880

Singh, J., Karmakar, S., PaiMazumder, D., Ghosh, S., & Niyogi, D. (2020). Urbanization alters rainfall extremes over the contiguous United States. *Environmental Research Letters*, *15*(7), 074033. https://doi.org/10.1088/1748-9326/ab8980

Skamarock, W. C. (2004). Evaluating mesoscale NWP models using kinetic energy spectra. *Monthly Weather Review*, *132*(12), 3019–3032. https://doi.org/10.1175/MWR2830.1

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S.-H., & Ringler, T. D. (2012). A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering. *Monthly Weather Review*, *140*(9), 3090–3105. https://doi.org/10.1175/MWR-D-11-00215.1

Su, J., Jiao, L., & Xu, G. (2025). Intensified exposure to compound extreme heat and ozone pollution in summer across Chinese cities. *npj Climate and Atmospheric Science*, *8*(1), 1–8. https://doi.org/10.1038/s41612-025-00966-5

Sui, X., Yang, Z.-L., Shepherd, M., & Niyogi, D. (2024). Global scale assessment of urban precipitation anomalies. *Proceedings of the National Academy of Sciences*, *121*(38), e2311496121. https://doi.org/10.1073/pnas.2311496121

Sun, Y., Fang, B., Oleson, K. W., Zhao, L., Topping, D. O., Schultz, D. M., & Zheng, Z. (2024). Improving urban climate adaptation modeling in the Community Earth System Model (CESM) through transient urban surface albedo representation. *Journal of Advances in Modeling Earth Systems*, *16*(12), e2024MS004380. https://doi.org/10.1029/2024MS004380

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Tian, J., Zhang, Y., Klein, S. A., Öktem, R., & Wang, L. (2022). How does land cover and its heterogeneity length scales affect the formation of summertime shallow cumulus clouds in observations from the US Southern Great Plains? *Geophysical Research Letters*, *49*(7), e2021GL097070. https://doi.org/10.1029/2021GL097070

Tian, J., Zhang, Y., Knupp, K., Pangle, P., & Comstock, J. (2024). Summertime continental shallow cumulus cloud detection using GOES-16 satellite and ground-based ceilometer at north Alabama. *Journal of Geophysical Research: Atmospheres*, *129*(24), e2024JD041512. https://doi.org/10.1029/2024JD041512

Torelló-Sentelles, H., Marra, F., Koukoula, M., Villarini, G., & Peleg, N. (2024). Intensification and changing spatial extent of heavy rainfall in urban areas. *Earth's Future*, *12*(9), e2024EF004505. https://doi.org/10.1029/2024EF004505

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. United Nations. Retrieved from https://population.un.org/wup/assets/WUP2018-Report.pdf

United States Census Bureau. (2024). TIGER/Line shapefiles (2024 edition) [Dataset]. *United States Census Bureau*. Retrieved from https://www2.census.gov/geo/tiger/TIGER2024/

Vahmani, P., & Hogue, T. S. (2015). Urban irrigation effects on WRF-UCM summertime forecast skill over the Los Angeles metropolitan area. *Journal of Geophysical Research: Atmospheres*, *120*(19), 9869–9881. https://doi.org/10.1002/2015JD023239

Venter, Z. S., Chakraborty, T., & Lee, X. (2021). Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms. *Science Advances*, *7*(22), eabb9569. https://doi.org/10.1126/sciadv.abb9569

Vergopolan, N., Sheffield, J., Chaney, N. W., Pan, M., Beck, H. E., Ferguson, C. R., et al. (2022). High-resolution soil moisture data reveal complex multi-scale spatial variability across the United States. *Geophysical Research Letters*, *49*(15), e2022GL098586. https://doi.org/10.1029/2022GL098586

Vo, T. T., Hu, L., Xue, L., Li, Q., & Chen, S. (2023). Urban effects on local cloud patterns. *Proceedings of the National Academy of Sciences*, *120*(21), e2216765120. https://doi.org/10.1073/pnas.2216765120

Wang, J., Nikolaou, N., an der Heiden, M., & Irrgang, C. (2024). High-resolution modeling and projection of heat-related mortality in Germany under climate change. *Communications Medicine*, *4*(1), 1–8. https://doi.org/10.1038/s43856-024-00643-3

Wu, S., Chan, T. O., Zhang, W., Ning, G., Wang, P., Tong, X., et al. (2021). Increasing compound heat and precipitation extremes elevated by urbanization in South China. *Frontiers in Earth Science*, *9*, 636777. https://doi.org/10.3389/feart.2021.636777

Yan, H., Gao, Y., Wilby, R., Yu, D., Wright, N., Yin, J., et al. (2024). Urbanization further intensifies short-duration rainfall extremes in a warmer climate. *Geophysical Research Letters*, *51*(5), e2024GL108565. https://doi.org/10.1029/2024GL108565

Yang, J., Zhao, L., & Oleson, K. (2023). Large humidity effects on urban heat exposure and cooling challenges under climate change. *Environmental Research Letters*, *18*(4), 044024. https://doi.org/10.1088/1748-9326/acc475

Young, A., Bhattacharya, B., Daniëls, E., & Zevenbergen, C. (2025). Integrating WRF forecasts at different scales for pluvial flood forecasting using a rainfall threshold approach and a real-time flood model. *Journal of Hydrology*, *656*, 132891. https://doi.org/10.1016/j.jhydrol.2025.132891

Zhang, J., Zhang, K., Liu, J., & Ban-Weiss, G. (2016). Revisiting the climate impacts of cool roofs around the globe using an Earth system model. *Environmental Research Letters*, *11*(8), 084014. https://doi.org/10.1088/1748-9326/11/8/084014

Zhang, K., Cao, C., Chu, H., Zhao, L., Zhao, J., & Lee, X. (2023). Increased heat risk in wet climate induced by urban humid heat. *Nature*, *617*(7962), 738–742. https://doi.org/10.1038/s41586-023-05911-1

Zhang, K., Zhao, L., Oleson, K., Li, X. C., & Lee, X. (2025). Enhancing urban thermal environment and energy sustainability with temperature-adaptive radiative roofs. *Earth's Future*, *13*(1), e2024EF005246. https://doi.org/10.1029/2024EF005246

Zhang, S., Fu, H., Wu, L., Li, Y., Wang, H., Zeng, Y., et al. (2020). Optimizing high-resolution Community Earth System Model on a heterogeneous many-core supercomputing platform. *Geoscientific Model Development*, *13*(10), 4809–4829. https://doi.org/10.5194/gmd-13-4809-2020

Zhang, W., Villarini, G., Vecchi, G. A., & Smith, J. A. (2018). Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. *Nature*, *563*(7731), 384–388. https://doi.org/10.1038/s41586-018-0676-z

Zhao, L., Oleson, K., Bou-Zeid, E., Krayenhoff, E. S., Bray, A., Zhu, Q., et al. (2021). Global multi-model projections of local urban climates. *Nature Climate Change*, *11*(2), 152–157. https://doi.org/10.1038/s41558-020-00958-8

Zhao, L., Oppenheimer, M., Zhu, Q., Baldwin, J. W., Ebi, K. L., Bou-Zeid, E., et al. (2018). Interactions between urban heat islands and heat waves. *Environmental Research Letters*, *13*(3), 034003. https://doi.org/10.1088/1748-9326/aa9f73

Zhu, X., Ni, G., Cong, Z., Sun, T., & Li, D. (2016). Impacts of surface heterogeneity on dry planetary boundary layers in an urban-rural setting. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12164–12179. https://doi.org/10.1002/2016JD024982

References From the Supporting Information

Porson, A., Harman, I. N., Bohnenstengel, S. I., & Belcher, S. E. (2009). How many facets are needed to represent the surface energy balance of an urban area? *Boundary-Layer Meteorology*, *132*(1), 107–128. https://doi.org/10.1007/s10546-009-9392-4

Sun, Y., Oleson, K. W., Zhao, L., Mills, G., He, C., Demuzere, M., et al. (2025). Enhancing global-scale urban land cover representation using local climate zones in the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, *17*(11), e2025MS004934. https://doi.org/10.1029/2025MS004934