Check for updates

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2026MS005775

Key Points:

- U-Surf data set modified to be structurally consistent with urban canopy scheme of Weather Research and Forecasting (WRF) regional climate model
- High-resolution simulations performed using default parameters and updated U-Surf-WRF for 13 major US cities
- Simulations using U-Surf parameters better capture range of intra-urban variability in temperature, humidity, and heat stress

Supporting Information:

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

Correspondence to:

T. Chakraborty,
tc.chakraborty@pnnl.gov

Citation:

Jiang, T., Chakraborty, T., Cheng, Y., Mukherjee, R., Zhao, L., & Martilli, A. (2026). Impact of spatially continuous urban surface properties on summertime and heatwave simulations: A multi-city analysis. *Journal of Advances in Modeling Earth Systems*, *18*, e2026MS005775. https://doi.org/10.1029/2026MS005775

Received 19 JAN 2026
Accepted 20 SEP 2026

Author Contributions:

Conceptualization: TC Chakraborty
Data curation: Timothy Jiang, TC Chakraborty, Yifan Cheng, Rohit Mukherjee
Formal analysis: Timothy Jiang, TC Chakraborty, Yifan Cheng, Rohit Mukherjee
Investigation: Timothy Jiang, TC Chakraborty

© 2026 CIEMAT (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas), Battelle Memorial Institute and The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Impact of Spatially Continuous Urban Surface Properties on Summertime and Heatwave Simulations: A Multi-City Analysis

Timothy Jiang[1], TC Chakraborty[1], Yifan Cheng[2], Rohit Mukherjee[1], Lei Zhao[2], and Alberto Martilli[3]

[1]Pacific Northwest National Laboratory, Richland, WA, USA, [2]University of Illinois Urbana-Champaign, Champaign, IL, USA, [3]CIEMAT, Madrid, Spain

Abstract Urban areas are unique in form, function, and fabric, and representing them in process-based models requires, among other things, prescribing facet-level morphological and radiative properties, which vary widely. Most urban canopy models (UCMs) prescribe these by density class or local climate zone (LCZ), assigning identical values across broad regions or worldwide. Global km-scale, facet-level urban property data sets have recently emerged, but are seldom applied in regional modeling. Here, we incorporate one such data set, U-Surf, into the Weather Research and Forecasting (WRF) model, modifying it for WRF's multi-layer UCM and releasing it as U-Surf-WRF. Considering 13 US cities, U-Surf-WRF parameters vary more between LCZs than default WRF parameters, with consistently lower impervious fraction. To determine the effects of using U-Surf-WRF, we conduct high-resolution WRF summertime seasonal and recent heatwave simulations for these 13 cities using default and U-Surf-WRF parameters. Either prescribed by LCZ or for each grid point, using U-Surf-WRF usually yields more accurate surface temperatures. It generally decreases modeled urban air temperature and increases modeled urban humidity, yielding lower heat and dry islands. Decomposing the impact of each U-Surf-WRF variable, we find that while albedo can impact daytime simulations, especially for air temperature, morphology and impervious fraction are most relevant, especially for surface temperature. This study demonstrates the impact of city-specific, facet-level urban properties in urban-resolving numerical weather simulations, with remaining gaps against observations demonstrating other challenges in representing cities in process-based models. Conversely, in WRF simulations with poorly constrained parameters, we suggest caution interpreting the magnitude and spatial distribution of urban signals.

Plain Language Summary To simulate near-surface urban weather and meteorology, models need several urban parameters, for example, building height and albedo. Many modeling studies use default parameters, which are often the same for cities throughout the world, mainly due to the unavailability of city-specific estimates. Newly emerging urban data products have worldwide coverage on the km scale but have yet to be widely taken up in practice. Here, we modify one such data set, U-Surf, for implementation into the Weather Research and Forecasting (WRF) model, naming it U-Surf-WRF. U-Surf-WRF parameters are less urban than WRF defaults for 13 major US cities. We simulate summertime heat in these cities using default properties and U-Surf-WRF. U-Surf-WRF simulated urban surface temperatures are closer to observed values, more realistically capturing the range of temperatures across the city. In general, U-Surf-WRF simulations yield higher humidity, lower temperatures, and correspondingly smaller urban heat and dry islands. We break down which U-Surf variables contribute to increased accuracy both when implemented by neighborhood type (which WRF can do by default) and when prescribed on the km scale, which involves modifying WRF. We find that building height and urban fraction data are most important overall, but that albedo is relevant during the day.

1. Introduction

Growing populations, urban expansion, and background warming have all generally increased the risk posed by urban weather extremes (Tuholske et al., 2021). Since various physical processes associated with urbanization modify local climate (Arnfield, 2003; Qian et al., 2022), urban-resolving models that can represent these processes and capture urban weather and climate signals across spatiotemporal scales are needed (Sharma et al., 2021). Such models are critical for projecting extreme weather in cities and informing mitigation and adaptation strategies (Jiang, Krayenhoff, et al., 2025; Marey et al., 2025; Zhao et al., 2021), especially relevant because different urban areas can have unique forms and functions with distinct, sometimes non-linear,

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Methodology: Timothy Jiang, TC Chakraborty, Yifan Cheng, Rohit Mukherjee, Lei Zhao, Alberto Martilli
Project administration: TC Chakraborty
Resources: TC Chakraborty
Software: Timothy Jiang, TC Chakraborty, Alberto Martilli
Supervision: TC Chakraborty
Validation: Timothy Jiang, TC Chakraborty
Visualization: Timothy Jiang
Writing – original draft: Timothy Jiang, TC Chakraborty
Writing – review & editing: Yifan Cheng, Rohit Mukherjee, Lei Zhao, Alberto Martilli

interactions with background weather and climate (Stokes & Seto, 2019; Zhao et al., 2014). Several advances have been made in this regard in the last few decades, including the incorporation of coupled urban canopy models (UCMs) in regional climate models (Hamdi et al., 2012), the development of multi-layer UCMs (Martilli et al., 2002), their integration with building energy schemes (Salamanca et al., 2010), and the addition of vegetation into UCMs (Krayenhoff et al., 2020). There has also been a push to incorporate and advance urban-resolving modeling in global Earth systems models (ESMs) (Oleson & Feddema, 2020).

Not only are distinct urban areas unique in form and function, they also differ greatly from surrounding natural landscapes in terms of radiative, morphological, and thermal properties (Chakraborty et al., 2021; Wu et al., 2024). Thus, these urban-specific parameters are important to better constrain physical processes, such as the surface energy budget, in the urban environment, and to represent interactions between urban areas and their surroundings (Hu et al., 2026). However, our ability to constrain these parameters in process-based models is currently simple, partly stemming from a dearth of available data and partly due to model structure (Cheng et al., 2025). Many regional weather and climate models, including the Weather Research & Forecasting (WRF) model, the most used such model for examining urban climate mitigation strategies (Krayenhoff et al., 2021), use coarsely prescribed urban parameters—by region, broad urban class, or both—to capture urban impacts on the surface energy budget and near-surface microclimate. Many such parameters have been estimated for individual cities on an ad hoc basis, mainly focused on regions for which ground data are available, a limiting factor in studying the full range of cities globally in all their diversity. Moreover, these parameters can often be outdated since urban properties have often evolved over time and in distinct ways across regions (Du et al., 2025; Wu et al., 2024). This is particularly relevant as computational power improves and regional climate model simulations that cover multiple urban areas are performed.

Ideally, modelers would update prescribed urban parameters to reflect properties for simulation regions and periods. However, such data are often unavailable, especially for multi-city studies over large simulation domains, and so the default parameter values are often used. Data from WUDAPT (Ching et al., 2018) (World Urban Database and Access Portal Tools) have been applied over several cities worldwide, but they often do not cover the entire city and are not available for most cities.

Even within cities, urban parameters are spatially heterogeneous, with relevance for intra-urban variability in climate hazard and exposure (Chakraborty, Newman, et al., 2023). This heterogeneity is not always represented in urban modeling, although it is usually done in the case of impervious fraction (e.g., Newman et al., 2024). For other urban properties that may also vary in a spatially continuous manner, such as facet-level radiative properties and morphology, many regional weather and climate models use zones or classes to bin this variability. For example, some ESMs and regional climate models represent urban heterogeneity using three or four density classes (e.g., Jackson et al., 2013; Sun et al., 2021). Alternatively, the local climate zone (LCZ) framework (Stewart & Oke, 2012) classifies urban development patterns into 10 urban (and seven "natural") categories. It is the state-of-the-art standard for urban-resolving regional weather and climate modeling, and can be used relatively simply in recent releases of WRF (Demuzere et al., 2022). However, radiative and morphological data by LCZ tailored for individual cities remain scarce. Thus, even within the LCZ framework (Stewart & Oke, 2012), studies often use default urban radiative and morphological parameters for each LCZ (Bernard et al., 2024). The accuracy of this assumption is questioned, given that these parameters were heuristically derived, and when setting urban parameters in WRF, users are cautioned that "The default values are probably not appropriate for any given city" and that "Users should adapt these values based on the city they are working with."

While the lack of data is a barrier, WRF does currently permit users to assign morphological properties at each grid point. However, it does not allow for gridded albedo values by default, and code has not yet been released to allow this. Yet, capturing local variability is especially important for urban albedo. Lipson et al. (2024)'s evaluation of urban canyon models showed that all models' performance improved when provided with detailed albedo information. Therefore, there is a need to constrain urban albedo values and implement them in WRF spatially continuously, that is, at each urban simulation point.

To address the need for city-specific and intra-urban surface constraints (Wang et al., 2025), several groups have recently leveraged satellite measurements and derived products to generate global spatially continuous estimates of urban parameters. This includes GLObal Building heights for Urban Studies (GLOBUS) (Kamath et al., 2024), Global Building Morphology data set for URban hydroclimate modeling (R. Li et al., 2024), Global Urban Canopy Parameters (Liao et al., 2024), Three-Dimensional Global Building Footprints (3D-GloBFP) (Che

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2024), and U-Surf (Cheng et al., 2025). These data sets are derived from various global sources and processed to make them suitable for urban climate modeling.

Most of the data sets above focus on the morphological features of cities. Given their recent development, few studies to date have applied these products, although the use of high-resolution urban morphology has been shown to increase modeled drag and heat flux (Shen et al., 2019), reducing their error against observed air temperature and wind (Sun et al., 2021). However, radiative parameters, which currently only U-Surf provides, also vary for urban facets (namely roofs, walls, and roads) and modulate urban microclimates (Best & Grimmond, 2015; Chakraborty et al., 2021). This makes U-Surf the most comprehensive data set to demonstrate the impact of observational facet-level estimates of radiative and morphological properties for urban-resolving regional weather and climate modeling.

Since U-Surf was originally developed for a single-layer UCM embedded within the Community Land Model, its morphological parameters were developed for a structurally different UCM than the multi-layer UCM within WRF. Here, we modify the U-Surf data set to make it consistent with the urban structural assumptions in WRF, releasing this adapted data set as U-Surf-WRF, and combine it with a gridded LCZ data set to simulate climatological summers and recent extreme heat events in 13 major US cities. Using these high-resolution simulations, we assess how incorporating U-Surf-WRF (both prescribed by LCZ and at individual grid points in a spatially continuous manner) alters modeled surface temperature, air temperature, humidity, and urban heat island dynamics compared to standard default configurations. Specifically, the goals of the study are to: (a) incorporate the U-Surf data set into WRF for the first time; (b) evaluate WRF simulations using U-Surf-WRF against observations; and (c) examine the mechanisms by which implementing U-Surf by LCZ and in a spatially continuous manner impacts simulations of urban heat and moisture.

In the following section, we provide an overview of the U-Surf data set, our development of U-Surf-WRF, and the experimental design to run our model simulations over multiple US cities. We then validate our simulations with the updated U-Surf-WRF data set to demonstrate its ability to capture the magnitude and spatial variability of urban heat, humidity, and heat stress signals. Finally, we discuss the potential for using spatially continuous urban parameters, such as urban albedo, within WRF and future priorities to improve urban surface constraints in process-based models.

2. Methods

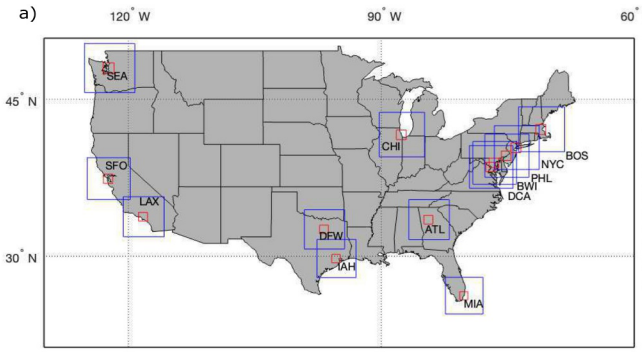
2.1. Cities and Simulation Periods

Heatwave events for 13 major US cities were simulated (Figures 1a and 1d). For each city, we identified one heatwave between May 2015 and June 2024, spinning up the model for 72, 69, 66, 63, and 60 hr to generate five ensemble members. The heatwaves were selected by identifying the 3+ consecutive days where the mean air temperature was highest for each city during this period (Meehl & Tebaldi, 2004). Cases and cities are shown in Figure 1, along with key model configuration. More details are provided in Section 2.3 below.

In addition to the targeted heatwave events, we conducted continuous seasonal simulations covering the entire summer of 2020 (1 June to 31 August) for all 13 cities. These season-long simulations were performed to establish baseline model performance and provide a robust, long-term validation against observational data sets. Given the computational expense of running high-resolution, season-long simulations, we generated a single ensemble member for 12 of the 13 cities. As an exception, five ensemble members were generated for Houston for the summer 2020 period to allow for a targeted analysis of ensemble spread and spatial continuous parameter sensitivity.

2.2. Brief Description of the U-Surf Data Set and Development of U-Surf-WRF

The recently developed U-Surf data set includes global 1 km facet-level estimates of several urban surface properties (Table 1) derived from several open-source data sets, providing data for urban surfaces such as roofs, walls, and the ground. Height-to-width ratio (H/W) is extracted from 3D-GloBFP (Che et al., 2024) and Microsoft building footprint data (Microsoft, 2022), plan area fraction (λ_p) is extracted from Microsoft building footprint data, and building height distribution is extracted from 3D-GloBFP. Radiative variables emissivity (ε) and albedo (α) are derived from ASTERv3 (Hulley et al., 2015) and Sentinel2 (Lin et al., 2022), respectively. Impervious



b)

Physics and dynamics options used in WRF

Urban physics	BEP-BEM (Salamanca et al., 2010)
Longwave	RRTM
Shortwave	Dudhia
Land	Noah
Microphysics	Thompson
Planetary boundary layer	BouLac
Resolution	$3km^2$, $1km^2$
Domain size	151x151, 100x100
AC adoption rate	50%
AC set point temperature	294.15 K

c)

Cities and corresponding heatwaves simulated

City	Abbreviation	Heatwave
Baltimore	bwi	19-21 Jul, 2019
Boston	bos	19-21 Jul, 2019
Philadelphia	phl	20-25 Jul, 2022
New York	nyc	21-23 Jul, 2021
Miami	mia	20-25 Jul, 2023
San Francisco	sfo	14-17 Oct, 2020
Seattle	sea	26-28 Jun, 2021
Dallas	dfw	18-21 Jul, 2018
Houston	iah	12-16 Aug, 2020
Los Angeles	lax	31 Aug - 8 Sep, 2022
Washington	dca	11-15 Aug, 2021
Chicago	chi	24-27 Aug, 2021
Atlanta	atl	21-27 Aug, 2023

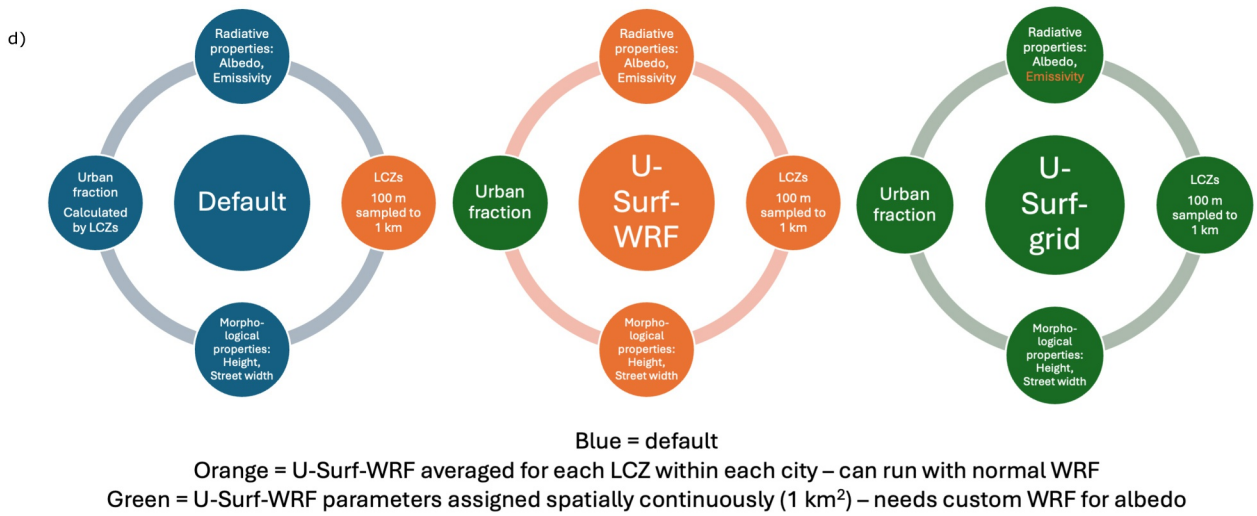


Figure 1. Experimental design and model configuration. (a) Inner (red) and outer (blue) Weather Research and Forecasting (WRF) domains for each of the 13 cities simulated. (b) Select WRF parameters and modules used in this study. (c) Cities, abbreviations, and heatwaves simulated in this study. (d) Three experiments conducted in this study. Blue indicates default WRF values were used. Orange indicates U-Surf values by local climate zone were derived (U-Surf-WRF). Green indicates spatially continuous U-Surf-WRF values were used, which we call U-Surf-grid. Note that emissivity is not provided by grid for these simulations. In addition to the heatwaves listed in (c), a continuous summer season (June–July–August 2020) was simulated for all cities for model validation.

fraction is extracted from ESA Worldcover v200 (Zanaga et al., 2022). For a complete description, the reader is referred to Cheng et al. (2025).

While ESMs such as the Community Earth System Model (CESM; Danabasoglu et al., 2020) and the US Department of Energy's Energy Exascale Earth System Model (E3SM; Golaz et al., 2022) can ingest these urban parameters grid point by grid point, WRF urban parameters are assigned by LCZ (or urban density classes in older

Table 1
Urban Surface Properties From U-Surf (Cheng et al., 2025) Used in This Study

Roof albedo	α_r
Wall albedo	α_w
Ground albedo	α_g
Roof emissivity	ε_r
Wall emissivity	ε_w
Ground emissivity	ε_g
Urban fraction	λ_u
Height-to-width ratio	H/W
Plan area fraction	λ_p
Building height	H

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

versions) through a lookup table. As such, to run WRF with U-Surf, we aggregate the spatially continuous U-Surf data into LCZs that can then be used to update the lookup tables for each city (Demuzere et al., 2020).

WRF simulates vegetated urban land such as lawns and parks separately from impervious urban surfaces. This differs from the UCM structure for which U-Surf was generated, where vegetated/pervious streets (e.g., lawns and bare soil) are considered part of the urban surface. Therefore, the H/W, λ_p, and λ_u were recalculated to make it consistent with WRF's UCM before they are ingested into WRF.

λ_u is modified by scaling it by 1 minus the fraction of pervious streets per urban area, which itself is equal to $\lambda_{rp}(1 - \lambda_p)$, where λ_{rp} is the fraction of urban ground area that is pervious (taken from U-Surf). Therefore:

$$\lambda_{u,\text{new}} = \lambda_u(1 - \lambda_{rp} + \lambda_p\lambda_{rp}) \tag{1}$$

Similarly, λ_p is modified by considering that the amount of roof area stays the same, but the urban area is being scaled as in Equation 1. That is, $\lambda_{p,\text{new}}\lambda_{u,\text{new}} = \lambda_p\lambda_u$ and so

$$\lambda_{p,\text{new}} = \frac{\lambda_p}{(1 - \lambda_{rp} + \lambda_p\lambda_{rp})} \tag{2}$$

The above is valid when the denominator of the plan area fraction is impervious area, as when λ_p is ingested by LCZ through URBPARM_LCZ.TBL. However, when λ_p is ingested into the URBPARAM variable (which allows for spatially continuous assignment of variables), WRF assumes that, consistent with NUDAPT standards, the denominator is not only impervious area but total area, urban and non-urban. Therefore, we have that:

$$\lambda_{p,\text{urbparam}} = \lambda_p\lambda_u \tag{3}$$

where λ_u is taken from U-Surf.

Finally, H/W is modified by applying Equation 5 from Cheng et al. (2025), noting that the λ_w in their equation is inversely proportional to λ_u. Substituting the new values of λ_u and λ_p given by Equations 1 and 3 and simplifying yields:

$$H/W_{\text{new}} = \frac{H/W}{1 - \lambda_{rp}} \tag{4}$$

For illustrative purposes, a comparison between the three different urban fraction fields (default, unmodified U-Surf, and modified U-Surf) and satellite imagery is presented in Figure 2 for select cities (Figure S1 in Supporting Information S1 for the remaining cities). There is more fine-scale variation in the modified U-Surf than in the unmodified U-Surf or the default fields, and suburban areas are less built-up in the modified U-Surf for most, but not all, studied cities. We validate this data set against ESA World cover (Zanaga et al., 2022) built-up area for the Chicago and Houston regions, finding that it corresponds better than both default and U-Surf impervious fraction fields (Figure S2 and Table S1 in Supporting Information S1). We also validate this data set against H/W derived from HiTAB (P. Li & Sharma, 2024) for Chicago and GLOBUS (Kamath et al., 2024) for Houston, finding that U-Surf-WRF has a better correlation than both default and U-Surf H/W, although it tends to be biased higher than U-Surf (Figure S3, Tables S2, and S3 in Supporting Information S1). The modified U-Surf data set appropriate for ingesting into WRF, which we name U-Surf-WRF, can be found at Jiang, Cheng, et al. (2025).

Figure 3 compares U-Surf-WRF against WRF default data over the 13 simulated cities in aggregate, and Figure S4 in Supporting Information S1 shows the same information for each of the 13 individual cities. Broadly, U-Surf-WRF data exhibit greater variability than default data. Urban albedo is more variable between LCZs for U-Surf-WRF for most cities, although it is similarly variable when cities are analyzed in aggregate. Urban emissivity is consistently greater for U-Surf, which is realistic given that it is generally considered underestimated in the default set of WRF parameters (Chakraborty et al., 2021).

Morphological variables tend toward a lower urban density in U-Surf-WRF, especially for LCZ 6 (open lowrise). Therefore, we hypothesize that urban heat island intensities simulated using U-Surf-WRF parameters would be

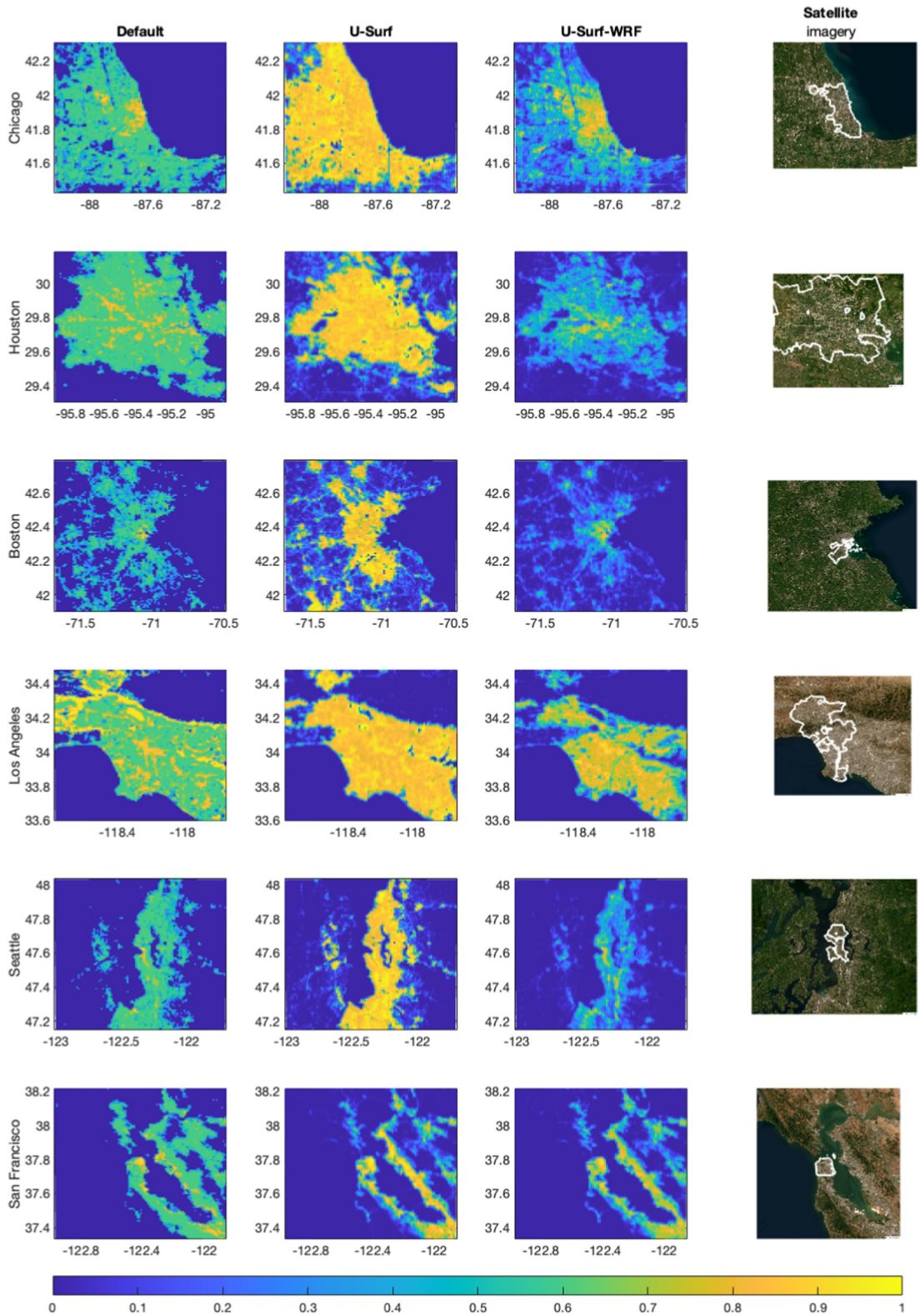


Figure 2. Urban fraction fields from default parameters, the unmodified U-Surf data set, and the U-Surf data set as modified in Section 2.2, compared with satellite imagery for select cities (sourced from Google Earth). This figure is produced for the remaining cities as Figure S1 in Supporting Information S1.

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

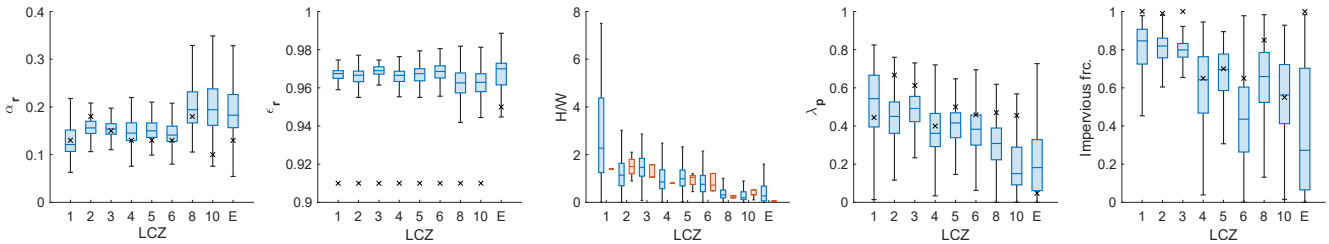


Figure 3. Select U-Surf-Weather Research and Forecasting (WRF) urban properties (blue) compared with WRF defaults (orange bars for height-to-width ratio [H/W] and *x*'s for the other properties), aggregated by local climate zone (LCZ). Each data point is one grid square within administrative city limits for all of the 13 simulated cities. The *x*-axis ticks correspond to the LCZs. See Stewart and Oke (2012) for how these LCZs are defined. This figure is reproduced for each of the 13 individual simulated cities as Figure S4 in Supporting Information S1.

lower than those using default parameters, for that LCZ. However, this is not the case for LCZ 1, where most cities examined have a greater H/W and λ_p than WRF defaults. We note that U-Surf-WRF suggests a wider distribution of H/W than WRF defaults for almost all cities and LCZs.

We also note that for some variables, the range of values within an LCZ can be large, and that this variability is not captured within the LCZ framework as implemented in WRF, except for the height-to-width ratio and impervious fraction. Therefore, in Section 2.4, we detail simulations to address this limitation of the LCZ framework.

2.3. Model Configuration and Boundary Conditions

The WRF regional weather model version 4.3.3 was used to dynamically downscale ERA5 reanalysis data to 3 km (outer grid) and then 1 km (inner grid). The coupled Building Effect Parametrization and Building Energy Model (BEP-BEM) urban canopy and building energy model was used to simulate urban effects (Martilli et al., 2002; Salamanca et al., 2010). It is a multi-layer scheme and is considered the state-of-the-art UCM in WRF. Simulation parameters are listed in Figure 1b.

For each simulation, two sets of urban parameter constraints were used: one with the default look-up table ("default" henceforth) and one with urban parameters given by the U-Surf-WRF data ("U-Surf-WRF" henceforth). The LCZ scheme (Stewart & Oke, 2012) was used to classify neighbourhoods at 100 m, and LCZ class data were extracted from Demuzere et al. (2020). While the default cases' impervious fraction fields were assigned from default values by LCZ using the w2w tool (Demuzere et al., 2021), those for the U-Surf-WRF cases were extracted from U-Surf following Section 2.2 and then resampled for each 1 km^2 grid in the inner domain. This incongruity between impervious fraction and other urban parameters fields is because it is common for WRF urban modeling practitioners to assign impervious fraction by grid point while using default properties via the URBPARM_LCZ.TBL input file. We comment further on this in Section 3.6.

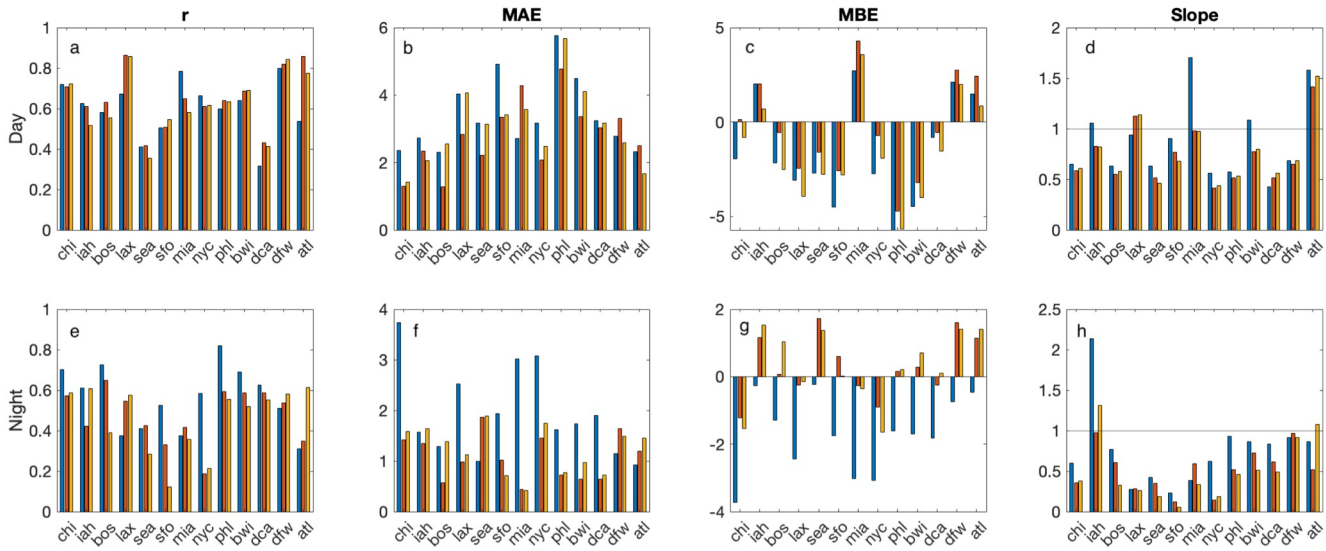
Select facet-level properties were given by U-Surf (Cheng et al., 2025), as shown in Table 1. Urban properties except H were derived by taking an urban-fraction-weighted average over each grid point within the city's administrative boundaries that corresponded to each LCZ. For example, for α_r for LCZ 2, we took the set of α_r values for all LCZ 2 grid points and weighted each point by λ_u. Building height distributions for each LCZ were derived from the λ_u-weighted distribution of the building heights for grid points in each LCZ, binned every 5 m (as is done in BEP-BEM).

In all sets of simulations, land use/land cover data for non-urban areas were derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data at a resolution of 15 arcseconds.

2.4. Simulations Assigning Spatially Continuous Urban Parameters

By default, the WRF model is able to ingest certain urban parameters spatially continuously (i.e., grid point by grid point), including impervious fraction and building height distributions. In this study, for the purposes of validation (Sections 3.1 and 3.2) and sensitivity analysis (Section 3.6), we use a modified WRF that ingests urban albedo (α_g, α_w, α_r) for each grid point as well (Jiang, Cheng, et al., 2025). Simulations were conducted with all urban properties in Table 1 averaged and assigned for each grid point in the inner domain, except for ε, which is assigned city-specific U-Surf values by LCZ as the ASTER emissivity product used in U-Surf does not resolve

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Day	Default	U-Surf-WRF	U-Surf-grid
Mean r	0.60	0.65	0.62
Mean MAE [C]	3.38	2.82	3.07
Mean MBE [C]	-1.53	-0.37	-1.45

Night	Default	U-Surf-WRF	U-Surf-grid
Mean r	0.56	0.48	0.46
Mean MAE [C]	1.96	1.07	1.22
Mean MBE [C]	-1.70	+0.30	+0.32

Figure 4. Model land surface temperature (LST) performance for 13 US cities, observations regressed as a function of model output. Simulation period is June–July–August 2020. Summary statistics are shown for MODIS observed versus Weather Research and Forecasting simulated LST, where each data point is one 1 km^2 grid point. MAE refers to Mean Absolute Error [C], MBE refers to Mean Bias Error [C], and r is the correlation coefficient. Plots of the slopes (d, h) have a reference line at slope = 1. Day (a–d) refers to 13:30 local time, and Night (e–h) refers to 1:30 local time. Tables show statistics for all cities analysed.

large spatial variability within cities (Figure 1c, Figure S4 in Supporting Information S1). These simulations are referred to as "U-Surf-grid" henceforth. While it takes effort and expertise to set up these simulations, after the first timestep it does not take materially longer for a U-Surf-grid simulation to run than the default approach.

Finally, for a more comprehensive sensitivity analysis, simulations for Chicago were conducted with only radiative ($\alpha_r, \alpha_w, \alpha_g$), morphological ($H/W, H, \lambda_p$), or impervious fraction (λ_u) variables derived from U-Surf-WRF by LCZ (with all other variables with default settings) or for each grid point (with all other variables implemented by LCZ). Further details are presented in Section 3.6.

3. Results

3.1. Surface Temperature Validation

We first examine if using U-Surf-WRF improves model simulated urban land surface temperature (LST) against best available benchmarks. We choose to use satellite-derived LST from NASA's MODIS Aqua satellite as the benchmark since it provides daily estimates at our model's native resolution (1 km). We compare modeled against observed MODIS LST for the summer of 2020. We perform the same evaluation for heatwave periods only in Figure S6 in Supporting Information S1 and refer to some of its results here.

Figure 4d shows that the range of intra-urban daytime surface temperatures in all cities is overestimated using default parameters (slope <1 indicating that the model has a greater range than observations). Using U-Surf-WRF parameters does not better capture the range of surface temperatures during the day, in most cities bringing the slope further away from 1. However, using U-Surf-WRF improves the spatial correlation of modeled surface temperature against daytime observations, increasing r by 0.03 (9 of 13 cities) (Figure 4a) and increasing spatial standard deviation (Figure S5a in Supporting Information S1), bringing it closer to observed values. Errors are reduced by 0.25°C (10 of 13 cities) (Figure 4b). Broadly, these errors tend to be underestimations (Figure 4c), with a MBE of −1.45°C and 9 of 13 cities having a negative bias for default simulations. This is not the case during heatwaves, where 10 of 13 cities have a positive (hot) bias (Figure S6 in Supporting Information S1).

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

At night, the cold bias (mean bias error = −1.70°C) is consistent across all 13 cities for default simulations, but using U-Surf-WRF reduces bias by 0.96°C (9 of 13 cities) and reduces error by 0.89°C (10 of 13 cities) (Figures 4f and 4g). Like during the day, using U-Surf-WRF results in warmer simulated LSTs in every city at night. As a whole, variability is not as well captured at night as during the day (lower r values, Figure 4e). Using U-Surf-WRF does not consistently increase r (−0.08, 6 of 13 cities) or bring the slope closer to 1 (error increased by 0.06, better in 4 of 13 cities).

Using U-Surf-grid does not appear to improve the evaluation when compared to (city-specific, LCZ-aggregated) U-Surf-WRF simulations. It does not consistently improve capturing LST variability or range at night (Figures 4e and 4h; Figure S5 in Supporting Information S1) and appears to degrade capturing LST range during the day (Figure 4a). Both during the day and at night, it increases error (+0.25°C) and bias (by 0.39°C) in most cities (9–10 of 13 cities; Figures 4b, 4c, 4f, and 4g). Therefore, we surmise that we do not see an improvement in capturing LST when U-Surf-WRF is implemented spatially continuously rather than by LCZ.

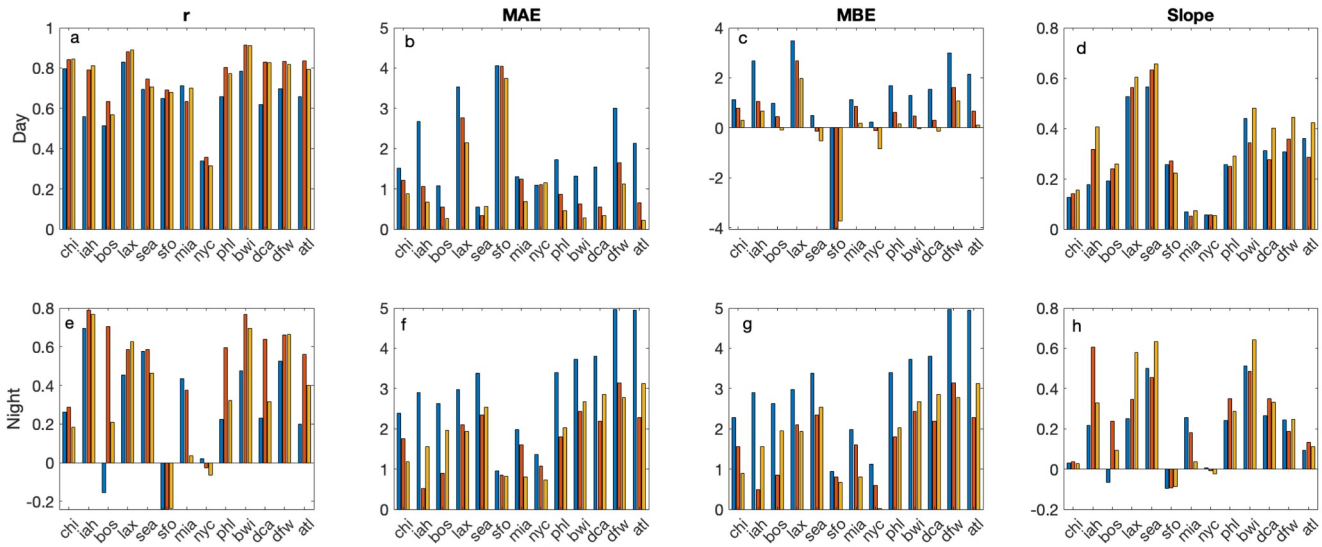
In addition to seasonal simulations, we compare modeled against observed MODIS LST for each heatwave period for each city. Figure S6 in Supporting Information S1 shows that the intra-urban variability in all cities is underestimated using default parameters, and that in most cases using the U-Surf-WRF parameters better captures the variability (r and slope closer to 1; Figures S6a and S6d in Supporting Information S1) and lowers error (Figures S6b and S6f in Supporting Information S1). Both default and U-Surf-WRF simulations tend to overestimate surface temperatures, but U-Surf-WRF tends to do so to a lesser extent (Figures S6c and S6g in Supporting Information S1).

We note that surface temperature observations have limitations including limited valid data points in some cities, cloud cover and cloud shadows obscuring surfaces, thermal anisotropy, and algorithmic uncertainties (Du et al., 2023; Yang, Xu, Wen, et al., 2024). Model limitations include lack of urban vegetation representation (Krayenhoff et al., 2021) and having only one ensemble member in these simulations. We further discuss limitations in Section 4. To address the last point about ensemble members, we perform two additional evaluations. We assess the performance of five ensemble members for the summer of 2020 for Houston in Section 3.6, and find results consistent with this section. The results in Figure S6 in Supporting Information S1 for heatwaves also have five ensemble members, but we note that they perform poorly, probably stemming from data limitations due to the shorter time periods (e.g., clouds, available pixels). We find that errors tend to be higher for the heatwave periods, with some large biases in some cases, but that the general results of the section bear out: that using U-Surf-WRF tends to bring simulations closer in line to satellite-derived LST estimates.

3.2. Simulating Urban Air Temperature Using U-Surf-WRF

Next, we turn our attention to air temperature. We evaluate against the data set of Zhang et al. (2022), which provides maximum and minimum air temperature statistically derived from in situ and satellite observations at our model's resolution, 1 km. Using the summer 2020 seasonal simulations, we evaluated model performance across the 13 cities under default, U-Surf-WRF, and U-Surf-grid configurations. Figures 5c and 5g shows that, overall, simulations are generally warmer than the Zhang et al. (2022) data set, especially at night, suggesting that we produce a stronger urban warming signal than Zhang et al. Broadly speaking, simulations also produce a greater range in air temperature than Zhang et al.'s data set, as well as greater variability (Figure S5 in Supporting Information S1). However, Zhang et al. may not fully capture urban spatial variability in their data set (Yang, Xu, Chakraborty, et al., 2024).

Compared to default simulations, U-Surf-WRF better captures variability in the large majority of cities during the day (Figures 5a, 5d, 5e, and 5h). U-Surf-WRF simulations have a lower range and variability in most cities, bringing it closer to the Zhang et al. (2022) data set. U-Surf-WRF consistently reduces bias (0.77°C, all cities) and error (0.68°C, 12 of 13 cities). Results are similar at night, with a consistent and especially large reduction in error (1.26°C, all cities) and bias (1.29°C, all cities), but slopes more in line with Zhang et al. (2022). We note that U-Surf-WRF reduces error the greatest for Houston and Dallas, sprawling cities for which the default impervious fraction greatly overestimates (Figure 2, Figure S1 in Supporting Information S1). There are many other possible controls on the performance of U-Surf within this modeling framework, including the characteristics of each individual city, its urban form, the presence of unrepresented urban greenery, and poorly constrained anthropogenic heat fluxes. Additionally, the accuracy of the validation data set may vary between cities, especially coastal cities with complex relief.



Day	Default	U-Surf-WRF	U-Surf-grid
Mean r	0.65	0.75	0.74
Mean MAE [C]	1.96	1.28	0.96
Mean MBE [C]	+1.20	+0.40	-0.07

Night	Default	U-Surf-WRF	U-Surf-grid
Mean r	0.28	0.48	0.34
Mean MAE [C]	3.03	1.77	1.92
Mean MBE [C]	+3.00	+1.71	+1.83

Figure 5. Model air temperature performance for each city, reference data regressed as a function of model output. Simulation period is June–July–August 2020. Summary statistics are shown for Zhang et al. (2022) maximum ("Day", a–d) and minimum ("Night", e–h) air temperature versus Weather Research and Forecasting simulated land surface temperature, where each data point is one 1 km^2 grid point. MAE refers to Mean Absolute Error (b, f), MBE refers to Mean Bias Error (c, g), and *r* refers to the correlation coefficient (a, e). Plots of the slopes (d, h) have a reference line at slope = 1. Tables show statistics for all cities analysed.

U-Surf-grid appears to further reduce error (0.32°C, 11 of 13 cities) and bias (0.31°C, 11 of 13 cities) in air temperature compared to U-Surf-WRF during the day, but not at night (increased error of 0.15°C, increased bias of 0.12°C, improvements in 6 of 13 cities) (Figures 5b, 5c, 5f, and 5g). It appears to capture the range better than U-Surf-WRF during the day, but not at night (Figures 5a and 5e). In terms of the magnitude of spatial variability, it performs similarly as U-Surf-WRF during the day, but substantially outperforms it at night (reduces error in standard deviation of T2 by 0.11°C, 11/13 cities). However, there is a marked decrease in *r* compared to U-Surf-WRF (−0.14, increases in 3/13 cities). As a whole, despite better capturing the spatial variability, U-Surf-grid appears to degrade performance in most metrics at night. It is surprising that U-Surf-grid often degrades simulations compared to aggregating parameters by LCZ as in U-Surf-WRF simulations. This could be due to many poorly constrained processes common to models at these scales, including existing model tuning hiding compensating errors and poor representation of processes and surface features, as discussed further in Section 4.

Evaluations of five ensemble members against in situ heatwave air temperature measurements for default and U-Surf-WRF simulations are provided in Figures S7 and S8 in Supporting Information S1 for the worst and best performing cities in the season-long evaluation (San Francisco and Boston, respectively). The in situ evaluations have a mean absolute error between 1.0 and 1.6°C, which we deem acceptable.

Having established model error for seasonal simulations, we examine air temperature in heatwaves using five ensemble members for default and U-Surf-WRF simulations. Figure 6 shows the mean 2 m air temperature, *T*2, by LCZ for each of the 13 simulated cities, along with the standard deviation within the city boundaries. As in the seasonal simulations (Figures 5c and 5g), on average, *T*2 is lower when simulated with U-Surf-WRF, and this result is robust for all simulated cities (Figure 6o). While each LCZ occupies a different place in each city (e.g., closer to or farther away from the shore), precluding inter-LCZ comparison, the difference between default and U-Surf-WRF for each LCZ is often greatest for suburban LCZ 6. This is consistent with LCZ 6 having the greatest difference between U-Surf-WRF and default parameters (Figure S4 in Supporting Information S1).

In 11 of our 13 simulated cities, the spatial variability of *T*2 is greater for simulations using U-Surf-WRF during heatwaves (0.03°C) (Figure 6n). Possible explanations include a larger range of impervious fractions throughout the city, a broader distribution of H/W within each LCZ, and a greater range of albedos between LCZs. However,

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

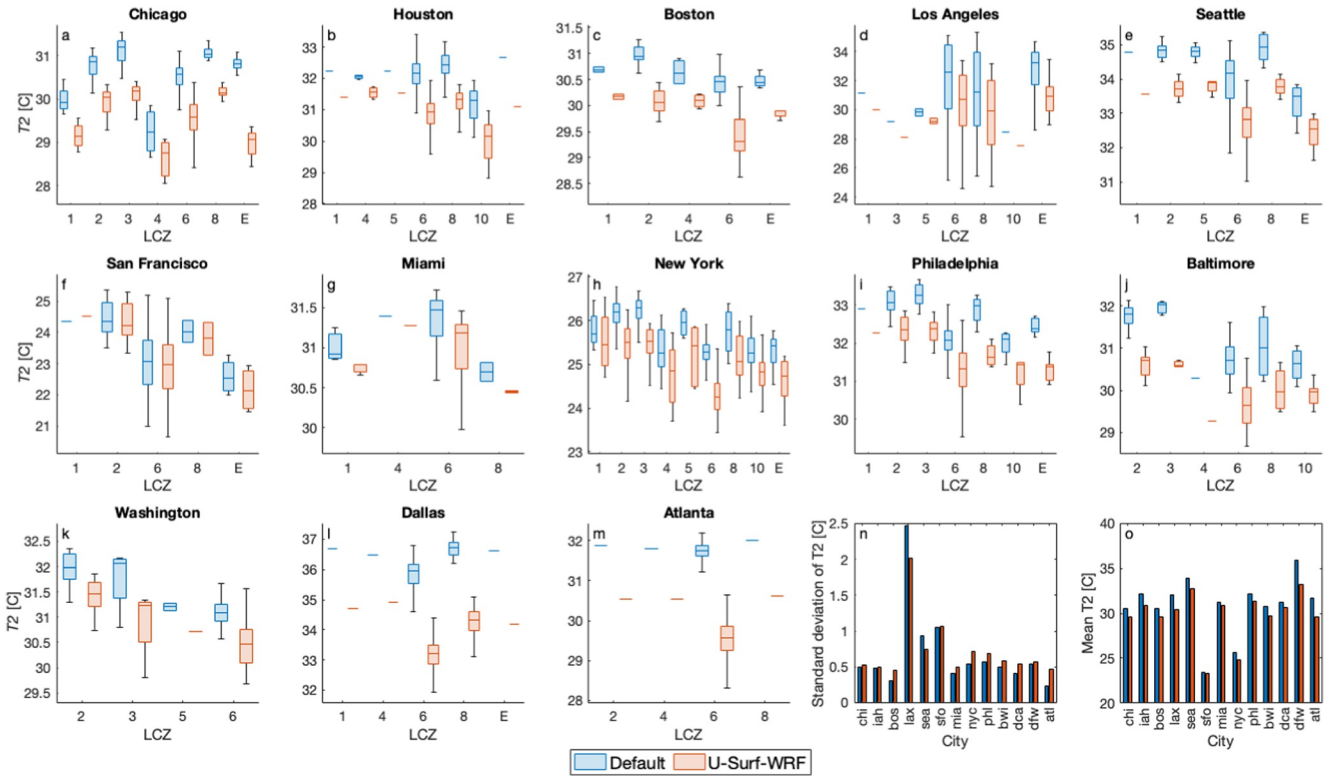


Figure 6. (a–m) Distribution of 2 m air temperature ($T2$) (°C) by local climate zone for each city, for both "default" and "U-Surf-Weather Research and Forecasting" simulations. Each point is one grid square $(1\ km^2)$ within administrative city limits. (n) Standard deviation of $T2$ for each city within administrative city limits. (o) Mean $T2$ for each city within administrative city limits. Each point is one output time (30 min).

the standard deviation of $T2$ is lower in Los Angeles under U-Surf-WRF. This is possibly because within the city boundary lie some forested and mountainous areas (Figure 2) which are cooler, so a lower urban air temperature would reduce the difference in air temperatures within the city boundaries.

Following mean $T2$, the daily maximum $T2$ exhibits similar variability for U-Surf-WRF simulations (Figure 7a). Conversely, at night, when the urban heat island effect is strongest, the minimum $T2$ varies more for default simulations (Figure 7b). This is consistent with default parameters representing heavier urban development, with corresponding stronger variability across LCZs.

3.3. Moisture and Moist Heat Estimates

Simulated heatwave moisture (vapor pressure, e) for each city is greater using U-Surf-WRF compared to the default simulations (Figure 8). This is in part due to the greater built-up land in the default parameters compared to U-Surf-WRF, which replaces transpiring, vegetated land cover. Given that U-Surf-WRF tends to bring impervious fraction closer to the values seen in more heavily vegetated areas within city boundaries, such as parks and other recreational areas, we might expect the variation in moist heat to be lower compared to default. However, the variability is similar between the two sets of simulations.

The humidex (Masterson & Richardson, 2002), a measure of humid heat and a function of both $T2$ and relative humidity, spatially varies more than $T2$. We observe that, in certain cities, the expanded urban area used by default simulations (Figure 2) results in a much drier climate within city limits, as many points are represented as more heavily urbanized than reality in the model (Figure 9). This difference is seen mainly in the suburban areas of cities (Figure 9), where open lowrise (LCZ 6) areas are common, an LCZ whose urban intensity is particularly overestimated by WRF default parameters compared to U-Surf-WRF (Figure S4 in Supporting Information S1). The lower humidity in cities is a measure of the urban dry island effect, and is consistent with observational estimates (Chakraborty et al., 2022), suggesting that, when we are concerned with moist heat, the lower simulated

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

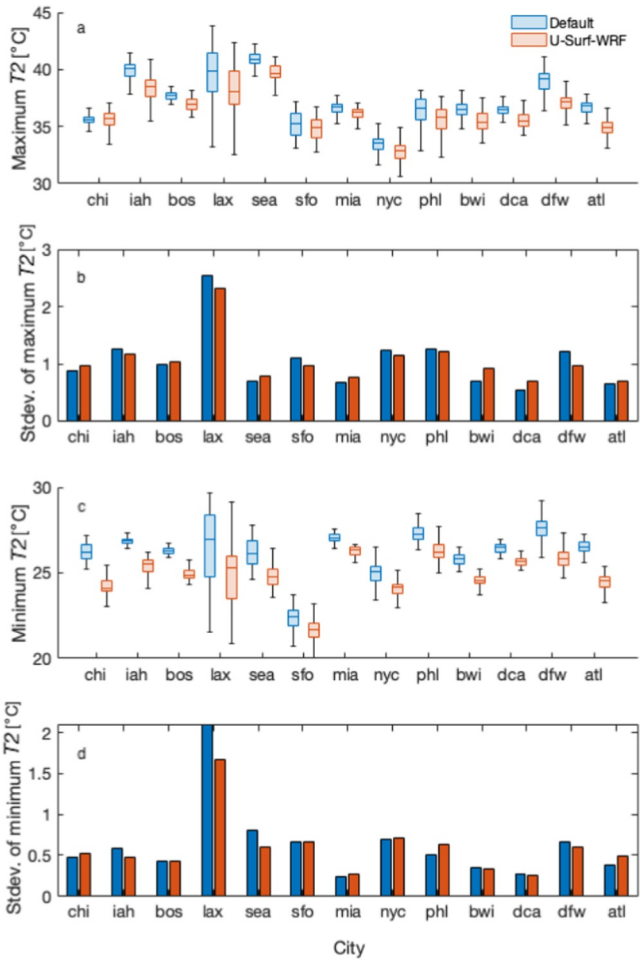


Figure 7. (a) Maximum and (c) minimum air temperature (°C) simulated using default and U-Surf-Weather Research and Forecasting urban parameters, as well as their (b, d) spatial standard deviations, averaged over all days of each city's respective extreme heat event. Each point is one 1 km^2 grid square within administrative city limits.

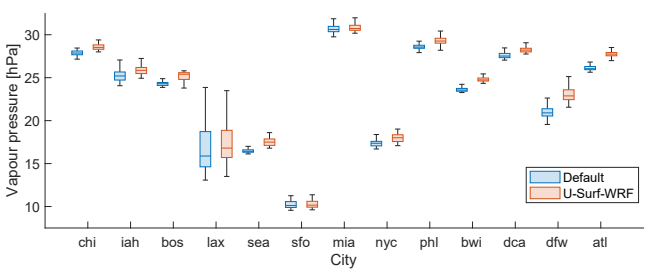


Figure 8. Vapor pressure (hPa) simulated using default and U-Surf-Weather Research and Forecasting urban parameters, averaged over all days of each city's respective extreme heat event. Each point is one 1 km^2 grid square within administrative city limits.

air temperature in suburbs when using U-Surf-WRF parameters may be partly compensated for by higher simulated humidity.

We acknowledge that, because urban observations of humidity are too sparse to capture intra-urban distributions, humidity is not validated in these simulations. Therefore, results are model-derived in this section and any results pertaining to moisture in subsequent sections. Therefore, there is greater uncertainty regarding moisture and moist heat.

3.4. Urban-Rural Differences in Temperature and Heat Stress

We compare urban (points within city limits with U-Surf-WRF impervious fraction >30%) and rural (land points in the (100 km)^2 inner domain with U-Surf-WRF impervious fraction <5%) $T2$ to extract a local urban warming signal. We note that the rural reference sites for some cities are much higher in elevation than the corresponding urban area, so we caution that this may not be a true "urban heat island." However, here we primarily wish to examine effects thereon driven by using different urban parameters, not the magnitude of the heat island. During the heat waves, we find that using default parameters often yields greater estimates of the urban heat island (Figure 10a). Estimates are especially higher at night, when the urban heat island is typically stronger (Figure 10b). This result is robust for every city analyzed.

Cities tend to be drier during the day but not at night, consistent with Meili et al. (2022) (Figures 10c and 10d). This could be due to a greater vapor pressure deficit arising from the UHI, which is stronger at night Omoto et al. (1994). Compared to default, U-Surf-WRF simulates a moister urban-rural difference, robust in all cities at night and in 11 out of 13 cities during the day. As with many outcomes examined in this section, this can be attributed in part to the lower impervious fraction in U-Surf-WRF, which would yield the diurnal pattern observed due to daytime transpiration.

When considering moist heat stress, the competing effects of a moister but cooler city when using U-Surf-WRF may offset urban-rural differences, as Section 3.3 suggests (Figures 10e and 10f). Indeed, in every city simulated, the difference between default and U-Surf-WRF simulations in an "urban humidex island" is less than that for air temperature urban heat island, although default WRF simulations still estimate higher values than U-Surf-WRF simulations.

3.5. Model Sensitivity to Individual Urban Parameters

In this section, we turn our attention to the individual U-Surf-WRF parameters' effects on $T2$ and humidity. We examine five different parameters (roof emissivity and albedo, impervious fraction, height-to-width ratio, and plan area fraction) and analyze the difference in $T2$ and e for corresponding differences in each of the parameters, over all cities' heatwave events.

Figure 11 shows that the greatest determinants of both air temperature and humidity differences between U-Surf-WRF and default simulations are H/W and impervious fraction, with R^2 at 10.5% and 10.4%, respectively, for $T2$ (10.2% and 10.5% for vapor pressure). We note that differences in H/W and impervious fraction are correlated in U-Surf-WRF, so the determination of the variance is lower than the sum of R^2 may suggest. Other variables appear to have weaker explanatory power, with R^2 between 1% and 4%. We suggest interpreting the low R^2 values not as an indication that other urban U-Surf-WRF parameter values are determining $T2$ and e changes to a greater

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

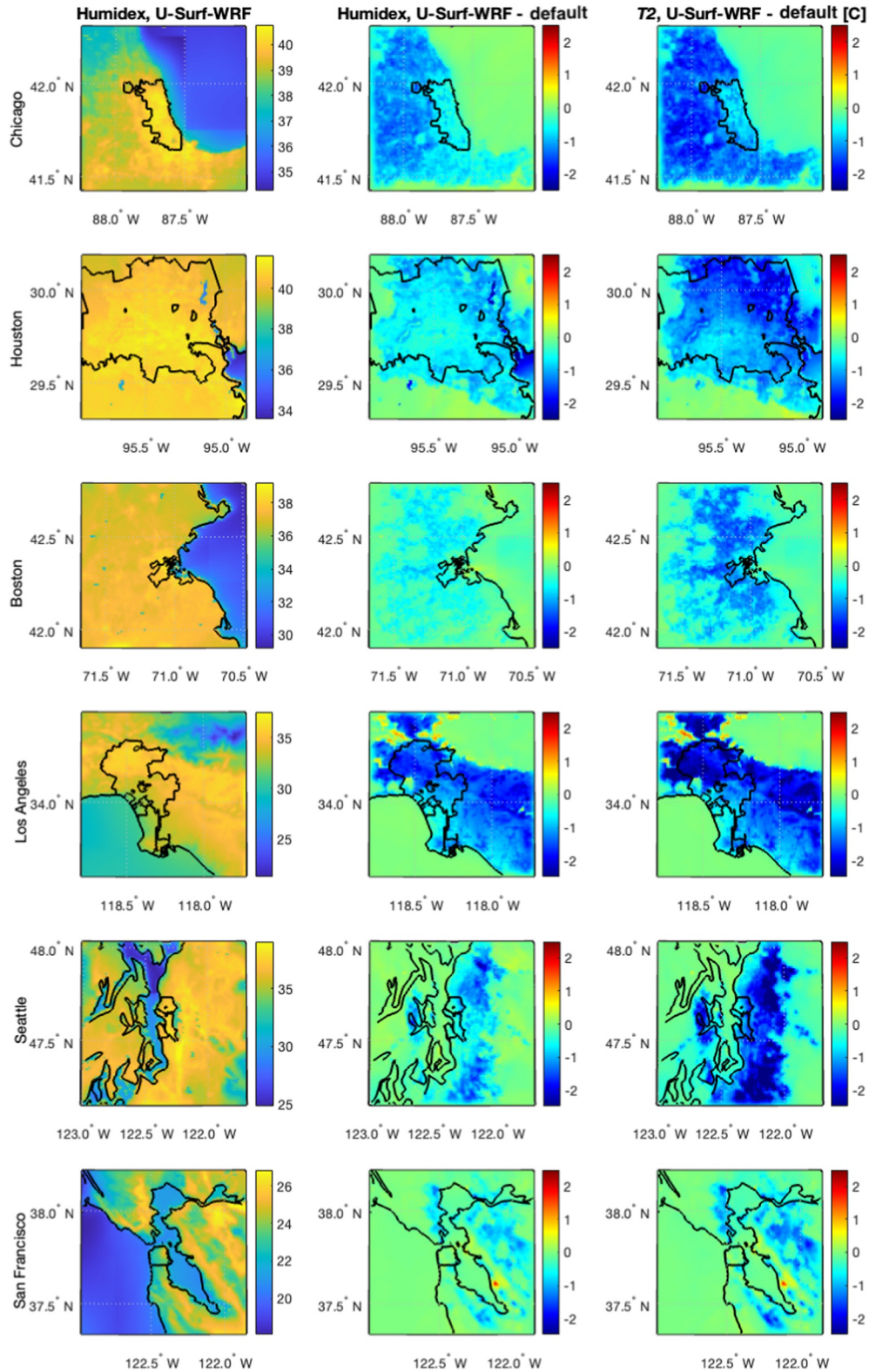


Figure 9. Humidex simulated using U-Surf-Weather Research and Forecasting (WRF) urban parameters (left); difference in humidex between U-Surf-WRF and default simulations (center); difference in air temperature between U-Surf-WRF and default simulations (right). This plot for the other seven cities is produced in as Figure S9 in Supporting Information S1.

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

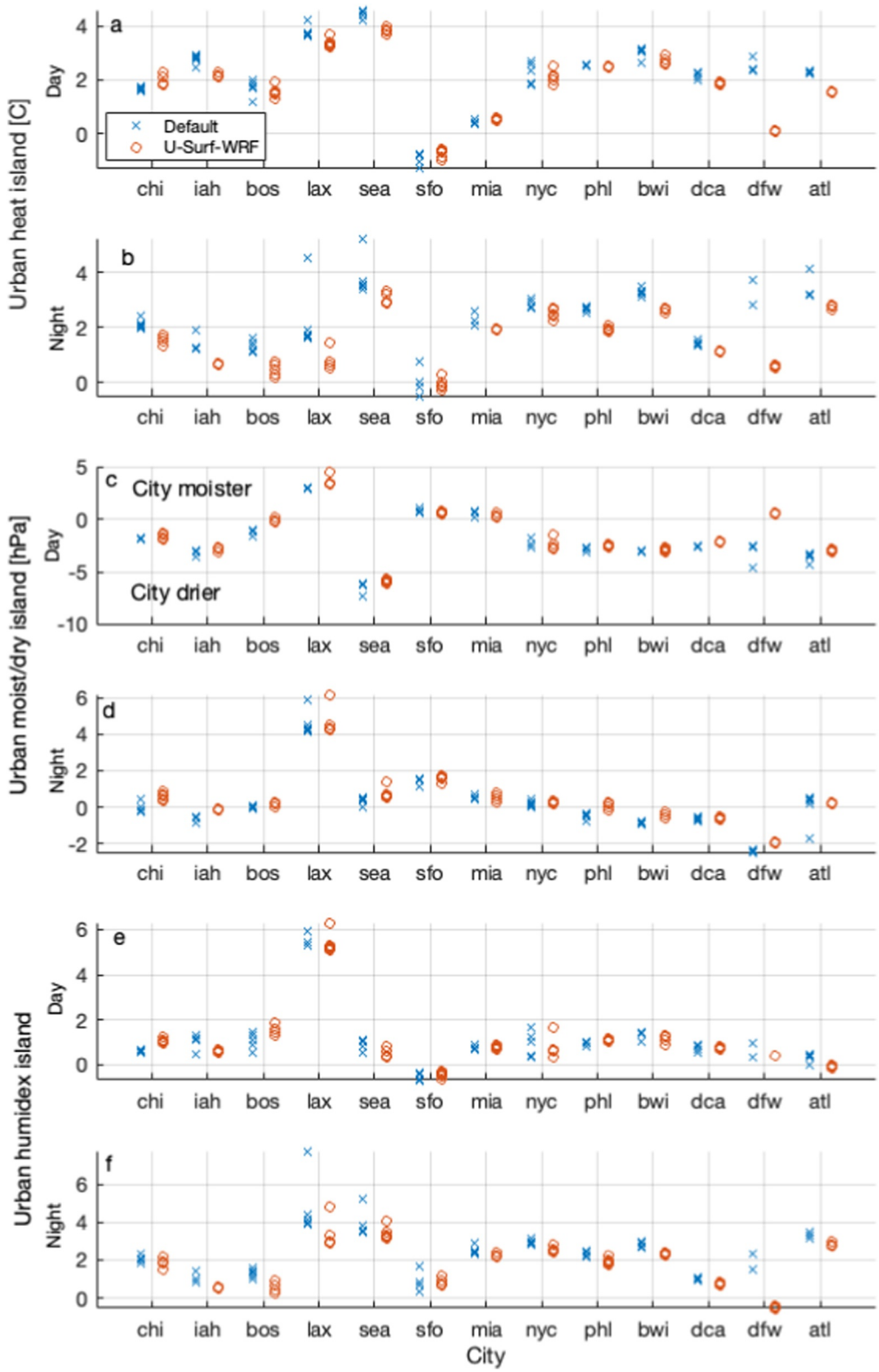


Figure 10. Urban-rural differences in (a, b) 2 m air temperature, (c, d) 2 m vapor pressure, and (e, f) 2 m humidex, during the day (a, c, e; 1:30 p.m. local time) and at night (b, d, f; 1:30 a.m. local time). Each point is one ensemble member. "Urban" and "rural" are defined in the main text.

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

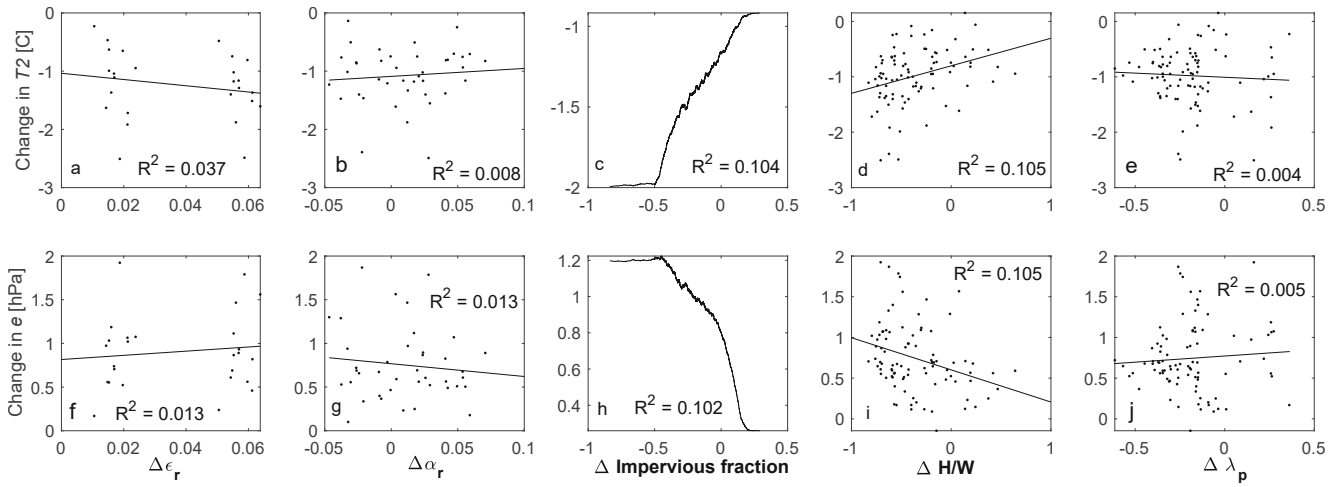


Figure 11. Changes in (a–e) 2 m air temperature ($T2$) and (f–j) 2 m vapor pressure (e) versus changes in (a, f) roof emissivity ($\Delta\epsilon_r$), (b, g) roof albedo ($\Delta\alpha_r$), (c, h) impervious fraction, (d, i) height-to-width ratio (H/W), and (e, j) plan area fraction (λ_p), for U-Surf-Weather Research and Forecasting and default simulations. Each local climate zone for each city is one data point, except for (c, h), where each 1 km^2 grid square within city limits is one data point. Panels (c, h) are processed using a 1,000-point moving mean, while all other sub-figures display only the mean y-values for each unique x-value.

extent, but as other physical processes (e.g., advection—note the intercepts in Figure 11) or noise outweighing the local U-Surf-WRF signal from changes in these parameters.

3.6. Implementing Spatially Continuous Urban Parameters Into WRF-BEP-BEM

In Sections 3.1 and 3.2, our multi-city evaluation of U-Surf-grid spanned a full season and primarily relied on single-member simulations. Here, we examine the effects of implementing U-Surf-grid in greater detail using our five-member ensemble runs. We begin by assessing the season-long ensemble for Houston (summer 2020), followed by an in-depth analysis of Chicago's specific extreme heat event. These heatwave-focused, ensembled experiments can differ from the season-long, single-member evaluations in Sections 3.1 and 3.2, and the benefit of grid-wise parameters can be event- and city-dependent.

We begin with Houston, for which the summer of 2020 is simulated for five ensemble members under default, U-Surf-WRF, and U-Surf-grid. Daily surface temperatures at 1:30 a.m. and 1:30 p.m. within city limits, as well as their spatial variability, are assessed against MODIS satellite observations. Figure S10 in Supporting Information S1 shows that surface temperatures are more accurately captured under U-Surf-grid during the day and U-Surf-WRF at night. Default parameters perform worst during both day and night.

We note that the slope in Figure S10a in Supporting Information S1 is shallower than in Figure S10b in Supporting Information S1, indicating that default parameters more severely underestimate the range in surface temperature than does U-Surf-WRF. At night, when urban effects are greatest, the U-Surf-grid simulations yield the greatest slope and r, indicating that U-Surf-grid captures the range in surface temperature better despite a higher mean absolute error (Figure S10f in Supporting Information S1). We also note the clustering of simulated nighttime LST near 25°C and 26°C with U-Surf-WRF simulations for a wide range of observed LSTs, suggesting clustering by LCZs which we do not observe in U-Surf-grid.

In terms of air temperature, we find that incorporating grid-wise urban parameters tends to very slightly increase simulated mean $T2$, similar to J. Chen et al. (2024) (Figure 12a), and we find slightly greater spatial variability when using U-Surf-grid compared to U-Surf-WRF (Figure 12d). We find that the distribution of the changes in the maximum $T2$ skew negative while the changes in the minimum $T2$ skew positive. We also note that the distribution of changes in mean $T2$ is relatively narrow and small in the mean compared to the maximum and minimum $T2$, implying that the points with the greatest deviation in maximum air temperature tend to have the least deviation in minimum air temperature, and vice versa. This would be consistent with increased urban intensity and morphology explaining most of the signal. However, impervious fraction is already continuously

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

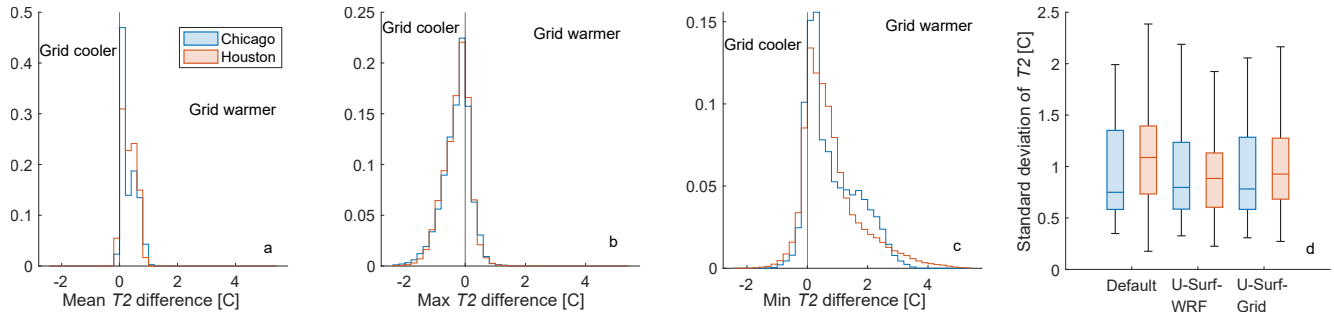


Figure 12. (a–c) Distribution of simulated $T2$ changes due to using U-Surf-grid versus U-Surf-Weather Research and Forecasting. "Max" and "min" $T2$ are taken at 1:30 p.m. and 1:30 a.m. local time, respectively. Each 1 km^2 grid square is one data point. (d) Spatial standard deviation of $T2$ for default, U-Surf-WRF, and U-Surf-grid simulations. Each 30-min time period for each of the five ensemble members is one data point.

assigned for "U-Surf-WRF" simulations. Therefore, we suggest that differences in urban morphology (i.e., height-to-width ratio and plan area fraction) are responsible for this spread.

To more comprehensively investigate the effects of assigning U-Surf-WRF properties in a spatially continuous manner, Chicago's heatwave event was simulated with different combinations of default, city-specific, and spatially continuous urban properties assignments for each variable. In addition to the three cases investigated in the rest of the manuscript (Default, U-Surf-WRF, and U-Surf-grid), we introduce additional cases to diagnose which variables affect simulations the most, incrementally implementing U-Surf-WRF parameters by LCZ or in a spatially continuous manner. A common configuration in WRF model simulations is to assign impervious fraction for each grid point but to otherwise use default urban parameters, since WRF comes by default with NLCD impervious fraction data and a relatively easy way to ingest them. Therefore, we include also this configuration (10, DDG; see Figure 13) in our analysis.

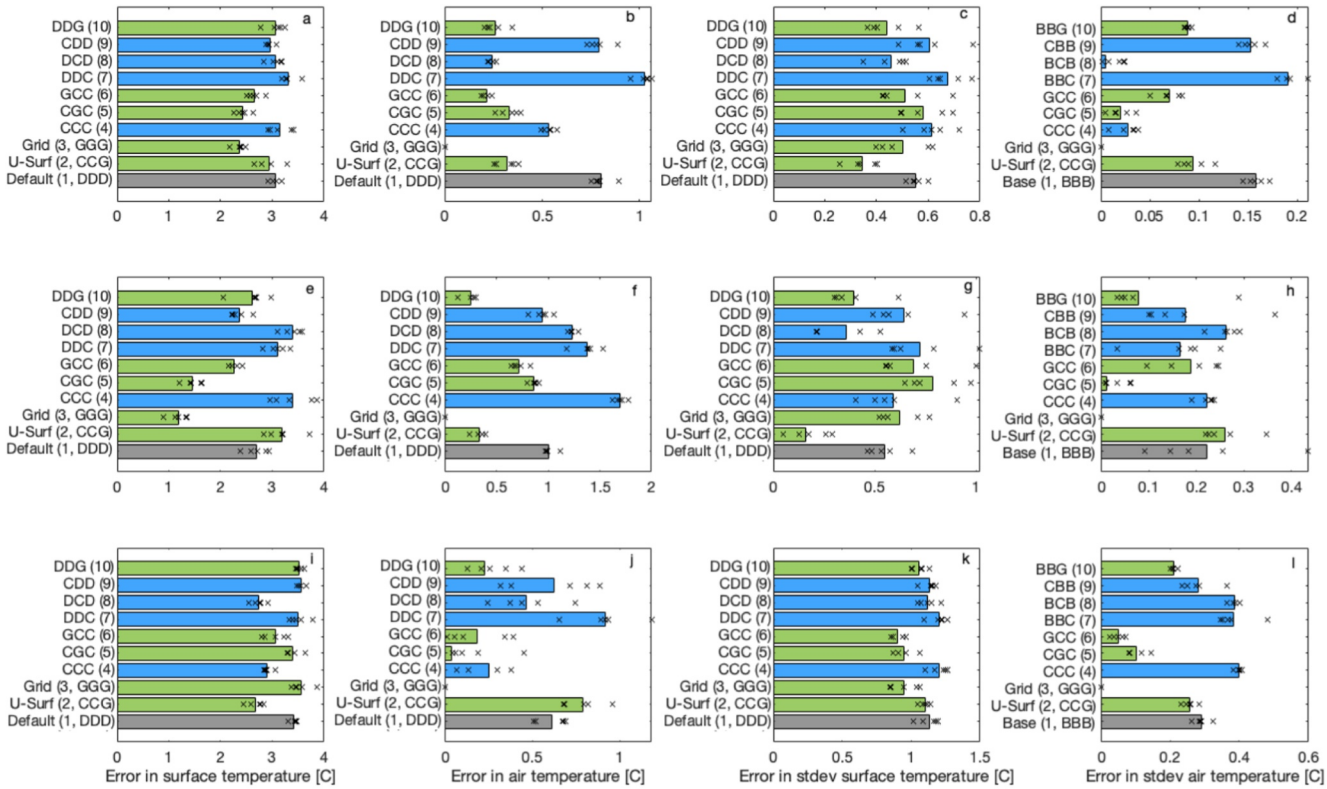
Figure 13 illustrates the relative improvement for implementing each U-Surf-WRF product, by LCZ and spatially continuously. Here, D indicates default WRF values were used in the simulations, C indicates city-specific values derived from U-Surf-WRF were used for each LCZ, and G indicates spatially continuous values. The first letter of each case's name indicates radiative parameters α and ϵ, the second letter indicates morphological parameters H, H/W, and λ_p, and the third letter indicates impervious fraction. Of special note are cases beginning with G (spatially continuous radiative parameters, cases 3 and 6): only α are spatially continuous, whereas ϵ are prescribed city-specific values by LCZ from U-Surf.

Broadly speaking, implementing spatially continuous U-Surf-WRF values (green bars) appears to reduce errors in surface temperatures to a greater extent than implementing them by LCZ (blue bars). LCZ-wise implementation of U-Surf-WRF makes LST simulations more accurate at night, but often makes them worse during the day. We find that assigning impervious fraction by grid point often makes only a marginal difference in the mean values, but that it often has a substantial effect in reducing the error in the spatial variability of the surface temperature. In contrast, LCZ-wise implementation of U-Surf-WRF properties does not improve simulated variability in surface temperature.

Spatially continuous implementation of U-Surf-WRF properties appears, on the other hand, to offer additional improvements in simulating both air and surface temperature during the day compared to benchmarks. However, at night, there appears to be no substantial improvement over LCZ-wise assignment of U-Surf-WRF properties. This is surprising since urban effects are greater at night than during the day, and presumably also the effect of spatially continuous urban parameters. It suggests that cities—or at least Chicago during this extreme heat event—are more homogeneous at night than during the day, at least within LCZs.

Since albedo acts during the day and has only a residual effect at night, we expect improvements from implementing U-Surf albedo to be greater during the day than at night. We find that this is true for both air temperature and surface temperature (Figures 13e, 13f, 13i, and 13j). However, this is not the case for the spatial variability, where benefits of spatially continuous U-Surf albedo are weak during the day, even to the point of reducing simulation accuracy. Recall that Chicago U-Surf albedo values vary to a greater extent than default values (Figure S4a in Supporting Information S1). This may suggest that the spatial variability of albedo is overestimated. In

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Case	Radiative Properties	Morphological properties	Impervious fraction
1	Default	Default	Default
2	U-Surf-WRF by LCZ	U-Surf-WRF by LCZ	U-Surf-WRF by grid
3	U-Surf-WRF by grid	U-Surf-WRF by grid	U-Surf-WRF by grid
4	U-Surf-WRF by LCZ	U-Surf-WRF by LCZ	U-Surf-WRF by LCZ
5	U-Surf-WRF by LCZ	U-Surf-WRF by grid	U-Surf-WRF by LCZ
6	U-Surf-WRF by grid	U-Surf-WRF by LCZ	U-Surf-WRF by LCZ
7	Default	Default	U-Surf-WRF by LCZ
8	Default	U-Surf-WRF by LCZ	Default
9	U-Surf-WRF by LCZ	Default	Default
10	Default	Default	U-Surf-WRF by grid

Figure 13. Added value from U-Surf-Weather Research and Forecasting (WRF) products: mean (a–d; average of daytime and nighttime values), daytime (e–h; 1:30 p.m. local time), and nighttime (i–l; 1:30 a.m. local time) absolute deviations from reference values for Chicago's extreme heat period, provided from MODIS-AQUA observations for surface temperatures and from U-Surf-grid simulations (3,GGG) for air temperatures. Table describes cases represented by each bar; see main text for an expanded description of each bar. *x*'s indicate individual ensemble members. Green bars indicate simulations with at least one component represented spatially continuously, while blue bars indicate simulations with at least one component using city-specific values by local climate zone. U-Surf-WRF is abbreviated to "U-Surf" on the *y*-axis labels.

contrast, spatial variability is better captured at night when using spatially continuously assigned albedo. This could arise if the variability is overestimated during the day but only a small part of it carries over to the night, helping simulation accuracy in terms of σ_T.

Next, we examine $T2$, noting that $T2$ comparisons are to the U-Surf-grid simulations, which therefore introduces some error in this analysis, especially at night. We find that implementing U-Surf-WRF properties by LCZ improves agreement with the benchmark compared to default for mean temperatures and especially at night, but that it worsens accuracy during the day. However, implementation of any spatially continuous parameters (green bars) outperforms corresponding LCZ-based implementations for any time of day, although this was not the case in our previous validation against observations (Section 3.2).

We also find that the common WRF urban modeling practice of assigning impervious cover in a spatially continuous manner but other urban properties set to their defaults performs (case 10) compares well to simulations with simulations using U-Surf-WRF urban properties assigned by LCZ (case 2) during the day and in the mean for

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

surface temperature. For air temperature, keeping in mind the caveats above, it yields even closer results than case 2.

In summary, this analysis shows that, at least for Chicago during this extreme heat event, implementing spatially continuous urban parameters improves simulated air and surface temperatures, in general, with morphological parameters (H/W, λ_p) being more important for surface temperatures and facet-level albedo being more important for air temperatures. We also found that the common WRF practice of assigning spatially continuous impervious fraction and otherwise using default parameters may be adequate for simulations that are mostly interested in daytime results. Further analysis over a broader range of cities during a longer time period will advance our understanding of which urban parameters simulation outputs are most sensitive to in specific scenarios (e.g., cold or warm season), helping inform efforts to constrain urban parameters in a way that is fit for purpose (Chakraborty et al., 2024).

4. Discussion

By using an ensemble of urban-resolving high-resolution WRF-BEP-BEM simulations for 13 major US cities with different sets of urban parameters, we show the importance of representing city-specific urban surface constraints derived from spatially continuous (gridded) observations in mesoscale weather simulations. In particular, we find that, using spatially continuous facet-level properties from the U-Surf-WRF data set, the spatial variability of air temperature generally increases (by 0.03°C, 11 out of 13 cities), mean air temperature and humidex decrease (1.05°C and 0.52°C, respectively; all cities), and vapor pressure increases (0.82 hPa; all cities). Correspondingly, urban-rural differences in heat and moisture decrease for all cities. When evaluated against LST observations for extreme heat events, U-Surf-WRF increases r (day: +0.09, 12 out of 13 cities; night: +0.03, 6 out of 10 cities) and decreases error (day: −0.52°C, 8 out of 13 cities; night: −0.26°C, 6 out of 10 cities); and when compared against LST observations for summer 2020, U-Surf-WRF decreases error in 10 of 13 cities both when implemented by LCZ (0.68°C) and by grid point (1.00°C) during the day. We note that the cities simulated in this study skew coastal and temperate. Arid cities contrast with their rural surroundings differently, and may be a topic of further investigation. Nonetheless, results for our selection of cities suggest that using default LCZ properties for every city does not fully leverage the LCZ framework. While the default values may be suitable for some urban modeling purposes, ideally parameters by LCZ for each city would be used. For example, an LCZ 4 neighborhood may have a dense coverage of midrise buildings punctuated by an occasional tall building in one city, and in another be scattered tall buildings with flat ground in between. These two distributions of building heights, which can be captured with U-Surf-WRF, may lead to large differences in urban weather and climate simulations.

In the past, urban modelers had only one urban class or typology to work with. Finer model resolution and data increased the number of urban classes to 3 or 4 (usually low, medium, and high-intensity), then the 11 LCZ classes. Now, the state of the field of urban modeling is progressing toward assigning properties for each simulation grid square (as "U-Surf-WRF" cases do for impervious fraction), instead of aggregating by urban class (as "U-Surf-WRF" cases do for all other variables). An increasing number of UCMs, including those in CESM and E3SM, can ingest different urban parameters for every simulated urban grid point. J. Chen et al. (2024) found that modeled air temperature was on average not substantially different between simulations using parameters by LCZ and parameters assigned for every grid point, but that modeled minimum air temperature was more accurate. Here, we comprehensively assess this approach compared to the LCZ-based approach, informing efforts to develop models that can ingest data on this level (Figure 13). We do so by modifying the WRF base code to represent gridded urban albedo. Although this is not a capability that WRF has out-of-the-box, we encourage that this be included as a default option in future releases of WRF. We provide our modifications to the WRF source code to facilitate this in Jiang, Cheng, et al. (2025).

We also perform sensitivity analyses to understand the role of changing different urban properties, including facet-level albedo, emissivity, and morphology parameters on simulated temperature and humidity. Overall, we find that impervious fraction as well as building height are most important for accurately capturing near-surface urban climate and its spatial variability on the scale of our simulations (1–100 km). The default WRF parameters tend to overestimate this impervious fraction, especially in suburban areas, and the raw U-Surf data set also severely overestimates this urban fraction (Table S1 in Supporting Information S1), because it was designed for the UCM integrated in CESM with a different canopy structure. We modify this data set to make it compatible with the structural assumptions in WRF, which we refer to as U-Surf-WRF and release for future use (Jiang,

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Cheng, et al., 2025). WRF's overestimated default urban parameters would lead to overestimations of the urban heat island and the urban dry island. These overestimations partially cancel out when calculating moist heat stress indices, though not completely (Figure 9). The overestimation of urban warming and other local meteorology in WRF is important to consider since it is the most common mesoscale model used in urban climate studies (Krayenhoff et al., 2021). With much of the literature also using these models to evaluate urban mitigation and adaptation strategies (Jiang, Krayenhoff, et al., 2025; Tan et al., 2024), we should be cautious in interpreting what these simulations tell us about the magnitudes and spatial variability of some of these effects. Moreover, these WRF simulations have also been used in conjunction with health and general socioeconomic (e.g., population distribution) data (Chakraborty, Wang, et al., 2023; K. Chen et al., 2022); and if simulations overestimate urban warming, that risks an overestimation of urban impacts on various outcomes, like disparities in heat hazard. It might be fruitful to reassess some of these results with more realistic impervious fraction as well as spatially continuous urban radiative and morphological parameters in the future.

With reference to the improvements in simulated variables seen on incorporating U-Surf-WRF, we note that UCMs often have missing and unresolved processes, in addition to parametric uncertainties (which we try to address here) and structural assumptions that are simplistic representations of 'real' urban form, function, and fabric. Many of these models have already been tuned to work reasonably well even with these uncertainties and missing processes. While we try to reduce uncertainties in the urban surface parameters using U-Surf-WRF, the existing model calibration may have been overcompensating for errors in other components. For example, U-Surf-grid degrading simulations at night while decreasing the slope (greater range/more outliers) suggests that differences in air temperature may not be mixed enough at night. Urban and boundary-layer models are most uncertain as regards anthropogenic heat and mixing at night (Sandu et al., 2013). Moreover, the effective dynamic resolution of the atmospheric model is coarser than the 1 km grid spacing, which means the LCZ-based aggregation may be acting as a low-pass filter that brings the surface input closer to the effective scale of the dynamics resolved (Skamarock, 2004). As such, even if the model performance degrades on using U-Surf-WRF and U-Surf-grid, which happens in a number of cases, that does not necessarily mean that the parameter constraints are wrong. Rather, first constraining what can be reasonably constrained using globally scalable data sets presents an opportunity to figure out how the other components of the modeling framework may be masking errors in past or default simulations. That is, we argue that the model may get the right answers for the wrong reasons, which is important to isolate in order to better understand these processes and for future model improvements.

Poor or coarse representation of other processes and parameters can explain many remaining disagreements between model simulations and benchmarks. For instance, nighttime urban air temperatures are more strongly driven by anthropogenic heat flux, which remains poorly constrained in process-based models at these scales, and for which default values by LCZ were used in all sets of simulations here.

In parallel, a major uncertainty in WRF and other urban-resolving regional climate models is the representation of urban vegetation. Multi-model comparisons have shown that models without explicit urban vegetation tend to underestimate latent heat flux (Lipson et al., 2024) and its impacts on associated meteorological variables. The representation of urban vegetation in regional climate models can be imprecise. In WRF, for example, the non-urban portion of urban grids are classified by default as "cropland mosaic," a mix of forests, grassland, shrubs, and cropland, when in practice different urban greenery is managed in a variety of different manners, varying from irrigated lawns and golf courses to unmanaged forests to croplands.

Additionally, there often lacks urban vegetation data beyond a general "pervious" fraction. U-Surf represents urban parks, lawns, yards, and other vegetated surfaces into one category, "pervious road," simulated in CTSM's urban model similarly to bare soil. Street trees, distinct from forests and parks, are particularly affected by this dearth of data, even though they modulate several aspects of urban climate (Coleman et al., 2021; Gromke & Ruck, 2009; Krayenhoff et al., 2020; Salmond et al., 2016). However, while U-Surf provides the fraction of urban ground that is vegetated, it makes no attempt to capture street trees apart from other urban vegetation, possibly explaining why, even with U-Surf-WRF implemented, spatial variability in LST is still poorly captured (Figures S6d and S6h in Supporting Information S1). While street tree and urban vegetation databases exist for individual cities, they are not universally available, and access and format limitations restrict their applicability for global or even regional-scale studies. Given this state of the modeling landscape, we urge the development of high-resolution street-tree and vegetation subtype products, possibly by leveraging satellite data, that UCMs with

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the capacity to resolve urban vegetation impacts can employ. In the same vein, we also encourage the development of UCMs that can capture the complex dynamics of urban vegetation in all its forms.

5. Conclusions

As satellite capabilities increase, providing weather and climate modelers with a wealth of urban data, it is important to assess the value of undertaking the effort to implement them and to make sure they improve simulation fidelity. It is not entirely clear under which circumstances finer-resolution prescription of surface properties would result in more accurate simulations. In this study, we showed that implementing the satellite-derived U-Surf-WRF urban canyon parameters yielded greater model fidelity for most metrics and times of day, while improving outcomes less in other circumstances. For example, implementing U-Surf-WRF by LCZ substantially increased modeled LST correlation with observations during the day, but implementing U-Surf-grid appeared to degrade some of the improvements. However, implementing U-Surf-WRF by grid points yielded a large improvement in modeled air temperature even compared to U-Surf-WRF by LCZ. With U-Surf-WRF, urban spatial variability was enhanced by 0.15°C (11 out of 13 cities), corresponding well with observations. The range was also brought better in line with observations, although still overestimated. U-Surf-WRF proved especially important when investigating urban-rural differences, where the magnitude of the effect is small, consistently yielding smaller nocturnal heat islands and daytime dry islands.

As with models, data must be fit for purpose (Chakraborty et al., 2024). We found that bulk estimates of average urban air temperature may be well simulated through the common WRF practice of using default urban parameters but assigning urban fraction grid by grid. However, we demonstrated that incorporating varied urban parameters may be valuable in applications where intra-urban variability is of interest, especially of surface temperatures. We also envision applications investigating intra-city inequitable exposure to heat, especially during extreme heat events, which have disproportionate and nonlinear impacts on health and comfort (Gasparrini et al., 2015).

While the algorithms employed to generate global-scale urban parameter data sets are not adapted to any particular city or neighborhood, it offers a baseline level of variability beyond what using default parameters offers. That is, city-specific data sets and observations will continue to be useful, but for national, continental, and global-scale studies, global-scale data sets derived from satellite observations can offer a level of accuracy in urban representation in regional climate models that generic LCZ parameters cannot. Therefore, we encourage the urban climate community to incorporate these spatially continuous estimates of urban canopy parameters in their simulations so as to more accurately capture the magnitude of urban heat, dry, and heat stress islands as well as their within-city variability. We also hope that the WRF community will continue to improve the model's capability to fully utilize such data sets for high-fidelity urban climate modeling. To facilitate this, we provide a global 1 km U-Surf-WRF product that is structurally consistent with WRF's urban canopy structure and code modifications to run WRF with grid-wise facet-level urban albedo and morphology parameters (Jiang, Cheng, et al., 2025).

Conflict of Interest

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

Availability Statement

Code from WRF version 4.3.3, data from U-Surf can be accessed through links in the bibliography. U-Surf data that have been modified for use with WRF, as described in Section 2.2, can be accessed at Jiang, Cheng, et al. (2025). Modifications to WRF to allow it to ingest grid-by-grid urban albedo can also be accessed at Jiang, Cheng, et al. (2025).

Acknowledgments

This work was supported jointly by the Biological and Environmental Research program of the US Department of Energy (DOE), Office of Science through an Early Career Research Program Award to T.C. and a U.S. National Aeronautics and Space Administration (NASA) grant through the AMT and LCLUC programs (Grant 80NSSC25K7322) to L.Z. and T.C. A.M.'s contribution has been supported by the CROCUS project funded by DOE BER (Grant DE-FOA-0002581). Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the DOE under Contract No. DE-AC02-05CH11231 using NERSC award BER-ERCAP0031684. Computational resources for this research were also provided by COMPASS, a coastal research program funded by the Biological and Environmental Science Program of the DOE Office of Science. Finally, the authors would like to thank three anonymous reviewers, whose comments at every stage proved constructive and helpful in improving this manuscript.

References

Arnfield, A. J. (2003). Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, *23*(1), 1–26. https://doi.org/10.1002/joc.859

Bernard, J., Bocher, E., Gousseff, M., Leconte, F., & Le Saux Wiederhold, E. (2024). A generic algorithm to automatically classify urban fabric according to the local climate zone system: Implementation in GeoClimate 0.0.1 and application to French cities. *Geoscientific Model Development*, *17*(5), 2077–2116. https://doi.org/10.5194/gmd-17-2077-2024

Best, M. J., & Grimmond, C. S. B. (2015). Key conclusions of the first international urban land surface model comparison project. *Bulletin of the American Meteorological Society*, *96*(5), 805–819. https://doi.org/10.1175/bams-d-14-00122.1

Chakraborty, T., 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., 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., Venter, Z. S., Demuzere, M., Zhan, W., Gao, J., Zhao, L., & Qian, Y. (2024). Large disagreements in estimates of urban land across scales and their implications. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-52241-5

Chakraborty, T., 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

Chakraborty, T., Wang, J., Qian, Y., Pringle, W., Yang, Z., & Xue, P. (2023). Urban versus lake impacts on heat stress and its disparities in a shoreline city. *GeoHealth*, *7*(11), e2023GH000869. https://doi.org/10.1029/2023GH000869

Che, Y., Li, X., Liu, X., Wang, Y., Liao, W., Zheng, X., et al. (2024). 3D-GloBFP: The first global three-dimensional building footprint dataset. *Earth System Science Data*, *16*(11), 5357–5374. https://doi.org/10.5194/essd-16-5357-2024

Chen, J., Dong, N., Liu, Z., Chen, Y., Luo, M., & Huang, H. (2024). Local temperature impact of urban heat mitigation strategy based on WRF integrating urban canopy parameters and local climate zones. *Building and Environment*, *267*, 112257. https://doi.org/10.1016/j.buildenv.2024.112257

Chen, K., Newman, A. J., Huang, M., Coon, C., Darrow, L. A., Strickland, M. J., & Holmes, H. A. (2022). Estimating heat-related exposures and urban heat Island impacts: A case study for the 2012 Chicago heatwave. *GeoHealth*, *6*(1), e2021GH000535. https://doi.org/10.1029/2021GH000535

Cheng, Y., Zhao, L., Chakraborty, T., 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

Ching, J., Mills, G., Bechtel, B., See, L., Feddema, J., Wang, X., et al. (2018). WUDAPT: An urban weather, climate, and environmental modeling infrastructure for the anthropocene. *Bulletin of the American Meteorological Society*, *99*(9), 1907–1924. https://doi.org/10.1175/BAMS-D-16-0236.1

Coleman, A. F., Ryan, R. L., Eisenman, T. S., Locke, D. H., & Harper, R. W. (2021). The influence of street trees on pedestrian perceptions of safety: Results from environmental justice areas of Massachusetts US. *Urban Forestry and Urban Greening*, *64*, 127258. https://doi.org/10.1016/j.ufug.2021.127258

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Demuzere, M., Argüeso, D., Zonato, A., & Kittner, J. (2022). W2W: A Python package that injects WUDAPT's LocalClimate Zone information in WRF. *Journal of Open Source Software*, *7*(76), 4432. https://doi.org/10.21105/joss.04432

Demuzere, M., Hankey, S., Mills, G., Zhang, W., Lu, T., & Bechtel, B. (2020). Combining expert and crowd-sourced training data to map urban form and functions for the continental US. *Scientific Data*, *7*(1), 264. https://doi.org/10.1038/s41597-020-00605-z

Demuzere, M., Kittner, J., & Bechtel, B. (2021). LCZ generator: A web application to create Local Climate Zone maps. *Frontiers in Environmental Science*, *9*, 637455. https://doi.org/10.3389/fenvs.2021.637455

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, H., Zhan, W., Zhou, B., Ju, Y., Liu, Z., Middel, A., et al. (2025). Exacerbated heat stress induced by urban browning in the Global South. *Nature Cities*, *2*(2), 1–13. https://doi.org/10.1038/s44284-024-00184-9

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., et al. (2015). Mortality risk attributable to high and low ambient temperature: A multicountry observational study. *The Lancet*, *386*(9991), 369–375. https://doi.org/10.1016/S0140-6736(14)62114-0

Golaz, J., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., et al. (2022). The DOE E3SM model version 2: Overview of the physical model and initial model evaluation. *Journal of Advances in Modeling Earth Systems*, *14*(12), e2022MS003156. https://doi.org/10.1029/2022MS003156

Gromke, C., & Ruck, B. (2009). On the impact of trees on dispersion processes of traffic emissions in street canyons. *Boundary-Layer Meteorology*, *131*(1), 19–34. https://doi.org/10.1007/s10546-008-9301-2

Hamdi, R., Degrauwe, D., & Termonia, P. (2012). Coupling the Town Energy Balance (TEB) scheme to an operational limited-area NWP model: Evaluation for a highly urbanized area in Belgium. *Weather and Forecasting*, *27*(2), 323–344. https://doi.org/10.1175/WAF-D-11-00064.1

Hu, H., Zhang, X., He, C., Fertig, E., Zarrin, R., Neyestani, S. E., & Zhang, J. (2026). Quantifying urban morphology-induced uncertainty in urban meteorology and heat stress simulations in Southern California. *Journal of Geophysical Research: Atmospheres*, *131*(4), e2025JD045318. https://doi.org/10.1029/2025JD045318

Hulley, G. C., Hook, S. J., Abbott, E., Malakar, N., Islam, T., & Abrams, M. (2015). The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale. *Geophysical Research Letters*, *42*(19), 7966–7976. https://doi.org/10.1002/2015GL065564

Jackson, T. L., Feddema, J. J., Oleson, K. W., Bonan, G. B., & Bauer, J. T. (2013). Parameterization of urban characteristics for global climate modeling. In *Geography of climate change* (pp. 135–152). Routledge.

Jiang, T., Cheng, Y., Mukherjee, R., Chakraborty, T., Zhao, L., & Martilli, A. (2025). U-Surf-WRF: Modifications to U-Surf and the WRF model for compatibility [Dataset and Software]. *Zenodo*. https://doi.org/10.5281/zenodo.15217359

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

Kamath, H. G., Singh, M., Malviya, N., Martilli, A., He, L., Aliaga, D., et al. (2024). GLObal Building heights for Urban Studies (UT-GLOBUS) for city- and street- scale urban simulations: Development and first applications. *Scientific Data*, *11*(1), 886. https://doi.org/10.1038/s41597-024-03719-w

Krayenhoff, E. S., Broadbent, A. M., Zhao, L., Georgescu, M., Middel, A., Voogt, J. A., et al. (2021). Cooling hot cities: A systematic and critical review of the numerical modelling literature. *Environmental Research Letters*, *16*(5), 053007. https://doi.org/10.1088/1748-9326/abdcf1

Krayenhoff, E. S., Jiang, T., Christen, A., Martilli, A., Oke, T. R., Bailey, B. N., et al. (2020). A multi-layer urban canopy meteorological model with trees (BEP-Tree): Street tree impacts on pedestrian-level climate. *Urban Climate*, *32*, 100590. https://doi.org/10.1016/j.uclim.2020.100590

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Li, P., & Sharma, A. (2024). Detailed height mapping of trees and buildings (HiTAB) in Chicago and its implications to urban climate studies. *Environmental Research Letters*, *19*(9), 094013. https://doi.org/10.1088/1748-9326/ad661a

Li, R., Sun, T., Ghaffarian, S., Tsamados, M., & Ni, G. (2024). GLAMOUR: GLobAl building MOrphology dataset for URban hydroclimate modelling. *Scientific Data*, *11*(1), 618. https://doi.org/10.1038/s41597-024-03446-2

Liao, W., Li, Y., Liu, X., Wang, Y., Che, Y., Shao, L., et al. (2024). GloUCP: A global 1 km spatially continuous urban canopy parameters for the WRF model. *Earth System Science Data Discussions*, 1–21. https://doi.org/10.5194/essd-2024-408

Lin, X., Wu, S., Chen, B., Lin, Z., Yan, Z., Chen, X., et al. (2022). Estimating 10-m land surface albedo from Sentinel-2 satellite observations using a direct estimation approach with Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, *194*, 1–20. https://doi.org/10.1016/j.isprsjprs.2022.09.016

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

Marey, A., Wang, L. L., Gaur, A., Lu, H., Leroyer, S., & Belair, S. (2025). Urban climate simulation for extreme heat events—A comparison between WRF and GEM. *Urban Climate*, *63*, 102570. https://doi.org/10.1016/j.uclim.2025.102570

Martilli, A., Clappier, A., & Rotach, M. W. (2002). An urban surface exchange parameterisation for mesoscale models. *Boundary-Layer Meteorology*, *104*(2), 261–304. https://doi.org/10.1023/A:1016099921195

Masterson, J., & Richardson, F. (2002). Humidex: A method of quantifying human discomfort due to excessive heat and humidity.

Meehl, G. A., & Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, *305*(5686), 994–997. https://doi.org/10.1126/science.1098704

Meili, N., Paschalis, A., Manoli, G., & Fatichi, S. (2022). Diurnal and seasonal patterns of global urban dry islands. *Environmental Research Letters*, *17*(5), 054044. https://doi.org/10.1088/1748-9326/ac68f8

Microsoft. (2022). Microsoft planetary computer. Retrieved from https://github.com/microsoft/GlobalMLBuildingFootprints

Newman, A. J., Kalb, C., Chakraborty, T., Fitch, A., Darrow, L. A., Warren, J. L., et al. (2024). The high-resolution urban meteorology for impacts dataset (HUMID) daily for the conterminous United States. *Scientific Data*, *11*(1), 1321. https://doi.org/10.1038/s41597-024-04086-2

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

Omoto, Y., Hamotani, K., & Um, H. H. (1994). Recent changes in trends of humidity of Japanese cities. *Journal of Japan Society of Hydrology and Water Resources*, *7*(2), 106–113. https://doi.org/10.3178/jjshwr.7.2_106

Qian, Y., Chakraborty, T., 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

Salamanca, F., Krpo, A., Martilli, A., & Clappier, A. (2010). A new building energy model coupled with an urban canopy parameterization for urban climate simulations—Part I. Formulation, verification, and sensitivity analysis of the model. *Theoretical and Applied Climatology*, *99*(3–4), 331–344. https://doi.org/10.1007/s00704-009-0142-9

Salmond, J. A., Tadaki, M., Vardoulakis, S., Arbuthnott, K., Coutts, A., Demuzere, M., et al. (2016). Health and climate related ecosystem services provided by street trees in the urban environment. *Environmental Health*, *15*(S1), S36. https://doi.org/10.1186/s12940-016-0103-6

Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., & Balsamo, G. (2013). Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models? *Journal of Advances in Modeling Earth Systems*, *5*(2), 117–133. https://doi.org/10.1002/jame.20013

Sharma, A., Wuebbles, D. J., & Kotamarthi, R. (2021). The need for urban-resolving climate modeling across scales. *AGU Advances*, *2*(1), e2020AV000271. https://doi.org/10.1029/2020AV000271

Shen, C., Chen, X., Dai, W., Li, X., Wu, J., Fan, Q., et al. (2019). Impacts of high-resolution urban canopy parameters within the WRF model on dynamical and thermal fields over Guangzhou, China. https://doi.org/10.1175/JAMC-D-18-0114.1

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

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, *93*(12), 1879–1900. https://doi.org/10.1175/bams-d-11-00019.1

Stokes, E. C., & Seto, K. C. (2019). Characterizing and measuring urban landscapes for sustainability. *Environmental Research Letters*, *14*(4), 045002. https://doi.org/10.1088/1748-9326/aafab8

Sun, Y., Zhang, N., Miao, S., Kong, F., Zhang, Y., & Li, N. (2021). Urban morphological parameters of the main cities in China and their application in the WRF model. *Journal of Advances in Modeling Earth Systems*, *13*(8), e2020MS002382. https://doi.org/10.1029/2020MS002382

Tan, H., Kotamarthi, R., Martilli, A., Fytanidis, D., & Dai, X. (2024). Impact of urban parameters on microclimate conditions and building energy consumption over Chicago metropolitan area during an extreme heatwave event. https://doi.org/10.20944/preprints202407.2459.v1

Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., et al. (2021). Global urban population exposure to extreme heat. *Proceedings of the National Academy of Sciences*, *118*(41), e2024792118. https://doi.org/10.1073/pnas.2024792118

Wang, C., Zhao, Y., Li, Q., Wang, Z.-H., & Fan, J. (2025). Ultrafine-resolution urban climate modeling: Resolving processes across scales. *Journal of Advances in Modeling Earth Systems*, *17*(6), e2025MS005053. https://doi.org/10.1029/2025MS005053

Wu, S., Lin, X., Bian, Z., Lipson, M., Lafortezza, R., Liu, Q., et al. (2024). Satellite observations reveal a decreasing albedo trend of global cities over the past 35 years. *Remote Sensing of Environment*, *303*, 114003. https://doi.org/10.1016/j.rse.2024.114003

Yang, Q., Xu, Y., Chakraborty, T., Du, M., Hu, T., Zhang, L., et al. (2024). A global urban heat island intensity dataset: Generation, comparison, and analysis. *Remote Sensing of Environment*, *312*, 114343. https://doi.org/10.1016/j.rse.2024.114343

Yang, Q., Xu, Y., Wen, D., Hu, T., Chakraborty, T., Liu, Y., et al. (2024). Satellite clear-sky observations overestimate surface urban heat islands in humid cities. *Geophysical Research Letters*, *51*(2). https://doi.org/10.1029/2023gl106995

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., et al. (2022). ESA WorldCover 10 m 2021 v200.

Zhang, T., Zhou, Y., Zhao, K., Zhu, Z., Chen, G., Hu, J., & Wang, L. (2022). A global dataset of daily maximum and minimum near-surface air temperature at 1 km resolution over land (2003–2020). *Earth System Science Data*, *14*(12), 5637–5649. https://doi.org/10.5194/essd-14-5637-2022

Zhao, L., Lee, X., Smith, R. B., & Oleson, K. (2014). Strong contributions of local background climate to urban heat islands. *Nature*, *511*(7508), 216–219. https://doi.org/10.1038/nature13462

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

19422466, 2026, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2026MS005775, Wiley Online Library on [27/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License