



Global multi-model projections of local urban climates

Lei Zhao^{1,2}✉, Keith Oleson^{1,3}, Elie Bou-Zeid^{1,4}, E. Scott Krayenhoff^{1,5}, Andrew Bray^{1,6},
Qing Zhu^{1,7}, Zhonghua Zheng^{1,8}, Chen Chen^{1,8} and Michael Oppenheimer^{1,9}

Effective urban planning for climate-driven risks relies on robust climate projections specific to built landscapes. Such projections are absent because of a near-universal lack of urban representation in global-scale Earth system models. Here, we combine climate modelling and data-driven approaches to provide global multi-model projections of urban climates over the twenty-first century. The results demonstrate the inter-model robustness of specific levels of urban warming over certain regions under climate change. Under a high-emissions scenario, cities in the United States, Middle East, northern Central Asia, northeastern China and inland South America and Africa are estimated to experience substantial warming of more than 4 K—larger than regional warming—by the end of the century, with high inter-model confidence. Our findings highlight the critical need for multi-model global projections of local urban climates for climate-sensitive development and support green infrastructure intervention as an effective means of reducing urban heat stress on large scales.

Cities are where major human-perceived climate change impacts occur^{1,2}. Occupying only ~3% of the Earth's land surface³, urban areas accommodate more than 50% of the world's population, and this percentage is projected to increase to 70% by 2050 (ref. ⁴). Many globally recognized environmental problems, such as heat stress, water scarcity, air pollution and energy security, are amplified in built areas through the uniqueness of urban climates and high population density^{5–7}. Despite the importance of urban climatic impacts, state-of-the-art Earth system models (ESMs) and general circulation models (GCMs) that participate in the Coupled Model Intercomparison Project (CMIP)⁸ almost universally lack a representation of urban areas. A primary reason is the legacy of early versions of GCMs designed for global-scale purposes in which urban areas were too small to resolve and did not cause discernible changes in large-scale dynamics and feedbacks. Therefore, the multi-model projections of global surface climate to date are essentially projections of 'non-urban' climate⁹. While the construction of multi-model ensembles is recognized as the best practice for deterministic or probabilistic climate projections with robustness and uncertainty characterization^{9,10}, using those non-urban projections for the assessment of climate-driven risks to human society in built environments (such as energy expenditures, labour productivity, city conflicts, human mortality and morbidity) is not an adequate alternative. Urban impacts derived from these ensembles are highly biased.

To address this shortcoming, recent efforts have either physically downscaled climate projections from one or two GCMs using regional climate models (RCMs) that have urban land schemes over limited regions at high computational cost^{11–13}, or have implemented urban representation in a GCM¹⁴. However, GCMs, although relatively consistent for global average results, exhibit large inter-model variability for regional climate projections, and

this variability alters the local urban response. Thus, urban forecasts based on a single GCM are insufficient. Projections of urban climate using GCMs and downscaling have shown large discrepancies or even conflicting results for the same regions^{15,16}. This is in large part due to (1) the internal variability of the climate system and (2) uncertainty in model parameters and structure and associated variation at the regional scale. For decadal or longer time horizons, the dominant sources of uncertainty at the regional scale are model parameter and structural uncertainty, in addition to scenario uncertainty¹⁷. The role of internal variability on local urban climate can be addressed by multi-member ensemble simulations using a single GCM, but the parameter and structural uncertainty are difficult to address because of the absence of multi-model urban simulations¹⁸. Additionally, the high computational demand of dynamical downscaling limits long-term and global-scale simulations, leading to a strong geographic bias in the regions studied toward China, Europe and North America, with few studies in polar regions, South America, Africa and Oceania¹⁹. This is a substantial omission, because the largest global cities are projected to be concentrated in Africa and Southeast Asia by 2100 (ref. ²⁰). To plan ahead for future climate-driven risks to cities and construct effective solutions, robust projections of local-scale urban-specific climates with quantitative characterization of confidence and uncertainty remain a critical research gap.

We have developed a method that combines fully coupled climate simulations and reduced-order modelling to provide the first global multi-model projections of local urban climates under climate change scenarios, with assessment of the associated robustness and uncertainties. We used the Community Earth System Model (CESM)²¹ of the US National Center for Atmospheric Research (NCAR) to build a reduced-order urban climate emulator, which was then applied to 26 CMIP Phase 5 (CMIP5) ESMs to generate

¹Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. ²National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL, USA. ³Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. ⁴Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA. ⁵School of Environmental Sciences, University of Guelph, Guelph, Ontario, Canada. ⁶Department of Mathematics, Reed College, Portland, OR, USA. ⁷Climate and Ecosystem Sciences Division, Climate Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁸Centre for Climate Research, Singapore, Singapore. ⁹School of Public and International Affairs and Department of Geosciences, Princeton University, Princeton, NJ, USA. ✉e-mail: leizhao@illinois.edu

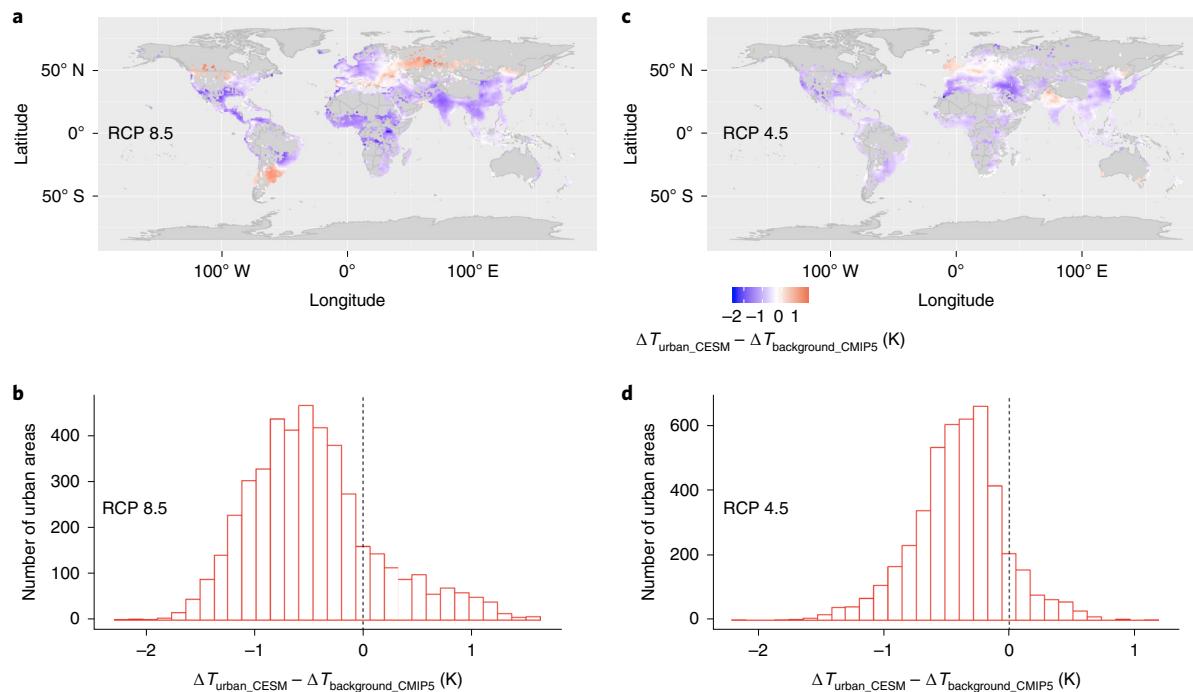


Fig. 1 | Difference between projected urban warming and background regional warming for JJA. Difference between CESM-modelled urban warming ($\Delta T_{\text{urban, CESM}}$) and background regional warming ($\Delta T_{\text{background, CMIP5}}$) projected by CMIP5 multi-model ensemble means for JJA. Δ refers to changes between the end of the century and the beginning of the century (2091–2100 versus 2006–2015). **a,c**, Global maps of the temperature difference between urban warming and regional warming under RCP 8.5 (**a**) and RCP 4.5 (**c**). The colours indicate grid cells that have urban land (total of 4,439 grid cells). Each coloured point represents decadal mean JJA subgrid urban 2 m air temperature difference relative to the 0.9° latitude $\times 1.25^\circ$ longitude model grid cell in which it is embedded. Dark grey indicates grid cells without urban land. **b,d**, Histograms of the temperature difference between urban warming and regional warming under RCP 8.5 (**b**) and RCP 4.5 (**d**).

urban projections under two representative concentration pathway (RCP) scenarios, namely RCP 4.5 and RCP 8.5. The CESM is the only CMIP5 model that has a physically based urban land parameterization²² and that has been evaluated against observed urban surface energy fluxes, surface temperature and air temperature over sites across the globe^{23–29}. The emulator is a location-dependent regression model with all of the atmospheric fields that force the urban land model in the CESM as inputs, and the urban climatic response as outputs (see Methods). We used 95 years (2006–2100) of monthly average atmospheric forcings and urban outputs from CESM simulations as the training sets to build the emulator, and we then applied it to all available ESMs in CMIP5 over the same time range. Although the focus of this study was on the mean urban climate, the whole emulator framework could be further developed to capture finer temporal scales (for example, daily or hourly), if the requisite training data were available. We evaluated both the CESM dynamic simulations and the emulator, with the results demonstrating its credibility (Extended Data Figs. 1–3 and Methods). We then analysed the difference between the end (2091–2100) and beginning (2006–2015) of the century to highlight urban warming signals for two seasons (June–August (JJA) and December–February (DJF)), and documented the validity of the emulator over the full range of available projections. Because the emulator framework outputs the data of the full time period (2006–2100), urban climate projections for any other timescales of interest (such as mid-century or sooner) can be obtained from the data. Urban planners and decision-makers in any city could therefore have access to city-specific projections for any planning horizon they need.

Our emulator uses atmospheric forcings from other ESM simulations in CMIP5 to drive a statistical, rather than dynamical, version of the CESM urban model. The variability and uncertainty in

large-scale dynamics and feedbacks among various models are presumably preserved in the large-scale atmospheric forcings, and this is particularly true given the small fraction of a typical ESM grid square that is occupied by urban area. Therefore, the inter-model variability in our emulated urban projections captures the impacts of large-scale climatology on the local-scale urban climate.

Our results confirm that the background regional/local warming (changes in grid cell mean temperatures between the end and beginning of the century) projected by the traditional CMIP5 models fails to capture the spatiotemporal variations of urban warming (changes in urban subgrid temperatures between the end and beginning of the century) forecast by the CESM under climate change (Fig. 1 and Extended Data Fig. 4). CESM-projected local urban warmings deviate from the regional background (non-urban) warming produced by the CMIP5 ensemble means under RCP 8.5 and RCP 4.5 by the end of the century (Fig. 1a,c). The urban warming effect relative to the surrounding regional warming is not uniform, either spatiotemporally or between models. Thus, urban warming projections cannot be represented with an anomaly added to regional background warming projections by ESMs, as this misrepresents the realistic warming signals for a large number of metropoles and smaller cities globally (Fig. 1b,d).

This misrepresentation is overcome in our multi-model ensemble projections of urban-specific warming, which show strong and robust climatic and geographic signals (Figs. 2 and 3). By the end of the century, the multi-model mean urban warming for JJA is 0.7–6.8, 0.7–7.6 and 0.7–6.7 K for monthly mean diurnal average (T_a), maximum (T_{max}) and minimum (T_{min}) air temperature, respectively, depending on the region and RCP scenario (Figs. 2 and 3). The magnitudes of the three warming signals (ΔT_a , ΔT_{max} and ΔT_{min}) are substantially less in the mitigated scenario (RCP 4.5) compared with

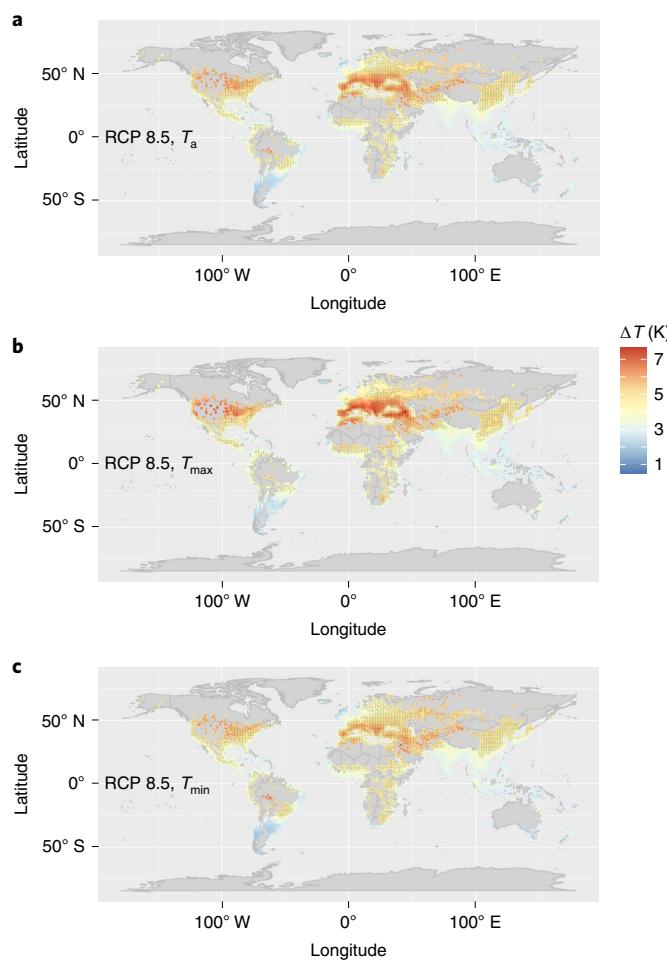


Fig. 2 | Multi-model ensemble mean urban warming for JJA under RCP 8.5. **a–c**, Seasonal mean urban warming between 2006–2015 and 2091–2100 for the diurnal average temperature T_a (**a**), the diurnal maximum temperature T_{\max} (**b**) and the diurnal minimum temperature T_{\min} (**c**) under RCP 8.5. Stippling indicates substantial change ($\Delta T \geq 4$ K) with high inter-model robustness ($\text{SNR} > 2.5$).

the high-emissions scenario (RCP 8.5). Specifically, the global mean urban warming in JJA reaches 4.4 K by the end of the century under RCP 8.5, which is 2.5 K or 130% higher than that under RCP 4.5. This confirms the crucial role of mitigating non-local greenhouse gas emissions for local-scale urban temperature benefits¹¹. By the end of the century, the mid-to-northern part of the United States, southern Canada, Europe, the Middle East, northern Central Asia and northwestern China exhibit the most pronounced urban warming during both daytime and night-time. Inland South America also shows strong night-time warming. This spatial variability is likely a result of the combined effects of regional background warming, differences in synoptic-scale dynamics, biophysical processes associated with urban surface energy balance^{16,23,30,31} and the sensitivity of the urban response to large-scale climate variability. Further research is needed for a quantitative attribution of these dynamics in future work. Our ensemble-projected urban daytime and night-time warming and their spatial gradients are in broad agreement with recent studies focused on the contiguous United States¹¹ (CONUS; Extended Data Fig. 2) as well as over European¹² and Australian³² cities using direct dynamical downscaling with all the mean absolute differences less than 0.5 K (see Methods), further demonstrating the high credibility of our emulator. Globally,

there is a weak diurnal effect in the ensemble mean projection. The global mean daytime urban warming is slightly stronger than the night-time one, with about 0.2 K difference between ΔT_{\max} (Figs. 2b and 3b) and ΔT_{\min} (Figs. 2c and 3c) under both scenarios. However, given that night-time urban warmth shows more dependence on the morphological aspects of a city^{23,33,34}, ΔT_{\min} here may be underestimated by the emulator due to the lack of urban growth in the CESM.

Our results also show strong urban warming in DJF by the end of the century (Extended Data Fig. 5). Consistent with JJA, the warming in DJF is largely mitigated under RCP 4.5 compared with RCP 8.5. The night-time warming is slightly stronger than daytime, with the global mean night-time value ~0.2 K higher than the daytime value under RCP 8.5. The most pronounced urban warming in DJF during daytime and night-time occurs in high-latitude cities in the Northern Hemisphere, reaffirming more rapid warming in the Arctic region³⁵. For example, observational evidence suggests that the city of Anchorage in Alaska is experiencing climate change at twice the rate of cities at mid-latitudes³⁶.

The spatial patterns of model local agreement on urban temperature change are shown in Extended Data Fig. 6, based on the signal-to-noise ratio (SNR), a measure of robustness. Model agreement increases under RCP 8.5 compared with RCP 4.5, especially in the middle and high latitudes of the Northern Hemisphere (Extended Data Fig. 6). This is because models project stronger urban warming under RCP 8.5, yet similar inter-model variability to RCP 4.5. We have used stippling on the maps of urban warming (ΔT ; Figs. 2 and 3 and Extended Data Fig. 5) to indicate cities with strong warming ($\Delta T \geq 4$ K under RCP 8.5 and $\Delta T \geq 1.5$ K under RCP 4.5) with high inter-model agreement ($\text{SNR} > 2.5$). Under RCP 8.5, our emulator projects with high confidence that a large number of cities globally will experience more than 4 K of warming during daytime and night-time in JJA by the end of century (Fig. 2b,c). These areas include the CONUS, the Middle East, the northern part of Central Asia, northeastern China and some inland cities of South America and Africa. Despite a substantially mitigated urban warming globally and fewer 'stippled' areas under RCP 4.5, the models still agree that a large number of cities (especially at night-time) will experience warming of more than 1.5 K, the target set by the Paris Agreement (Fig. 3b,c). Cities with substantial projected warming in DJF with high robustness are shown in Extended Data Fig. 5. Arctic cities are particularly impacted under RCP 8.5.

We find a near-universal decrease in urban relative humidity (RH) in JJA by the end of century (Fig. 4; –6 to 3% under RCP 4.5 and –13 to 6% under RCP 8.5). Compared with RCP 8.5, the reduction in RH is markedly smaller under RCP 4.5. Previous studies reported marginal RH change over non-urban surfaces under climate change, because both partial and saturation water vapour pressure increase under warming^{37,38}. This is not the case for urban surfaces. The saturation vapour pressure increases considerably over these surfaces due to warming. The increase in partial water vapour pressure, on the other hand, is limited by the large fraction of impervious surfaces in cities that hinders evaporation despite the elevated ability of warmer air to contain water vapour. The increase in partial pressure therefore remains smaller than over non-urban surfaces; urban-specific humidity during JJA is projected to increase slightly (by ~10–20%) by the end of the century under both scenarios (Extended Data Fig. 7). These results are in broad agreement with a recent downscaling study in Germany¹². Empirical evidence exists for such a decrease in observed urban RH in the last decade^{39–41}. Most coastal cities are projected to have indiscernible or no decrease in RH, because of the larger water availability that allows the increase in urban partial vapour pressure to keep up with its rural counterpart (Fig. 4). In DJF, global urban RH changes do not show a decreasing pattern as spatially consistent as in JJA (Extended Data Fig. 8).

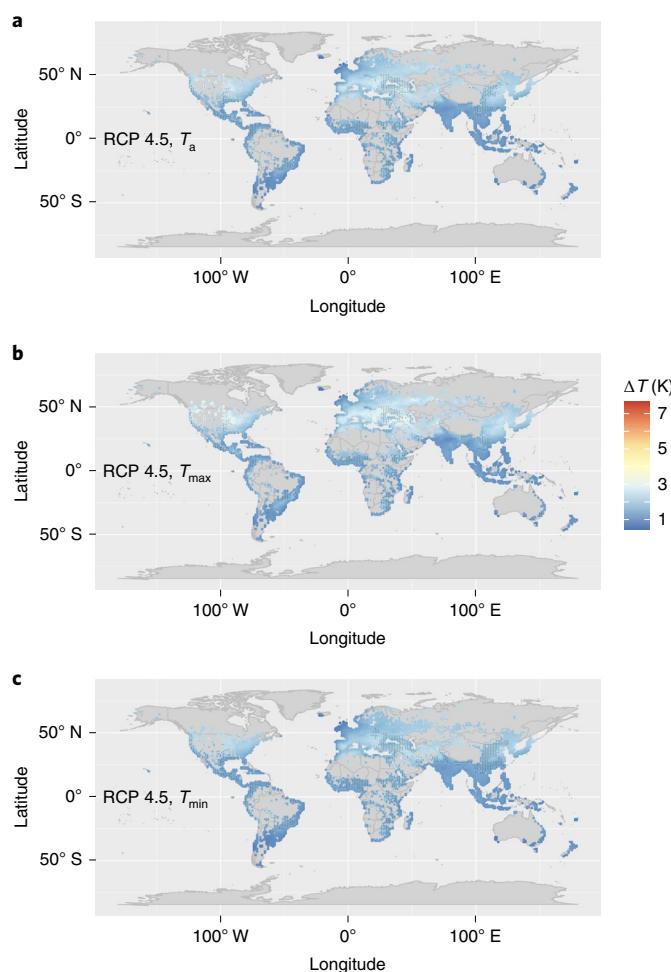


Fig. 3 | Multi-model ensemble mean urban warming for JJA under RCP 4.5. **a–c**, Seasonal mean urban warming between 2006–2015 and 2091–2100 for the diurnal average temperature T_a (**a**), the diurnal maximum temperature T_{\max} (**b**) and the diurnal minimum temperature T_{\min} (**c**) under RCP 4.5. Stippling indicates substantial change ($\Delta T \geq 1.5$ K) with high inter-model robustness ($\text{SNR} > 2.5$).

In general, the model agreement with respect to local urban RH is weaker than it is for temperature, indicating larger inter-model variability for RH projections (Extended Data Fig. 9). Regions with higher robustness coincide with, to some extent, the regions of large RH changes (Fig. 4). The stippling in Fig. 4 indicates urban areas with substantial projected RH change ($\text{abs}(\Delta \text{RH}) > 5\%$ under RCP 8.5 or $\text{abs}(\Delta \text{RH}) > 2.5\%$ under RCP 4.5) and high inter-model robustness ($\text{SNR} > 1$). Its spatial coverage is markedly smaller than the corresponding stippling for temperature (Figs. 2 and 3).

The human-perceived impacts of climate change are primary factors that motivate the growing research focus on urban climates. We further evaluated two commonly used measures of human-perceived heat stress in which humidity is factored in with air temperature: the wet-bulb temperature (T_{wb}) and the US National Weather Service version of the heat index (HI). The former is preferable as it is a measure inherently tied to thermodynamics rather than an empirical index and has the added advantage of serving as an indicator of the efficiency of evaporative cooling. We estimate that the HI increases faster than air temperature over almost all the cities globally under both scenarios (Extended Data Fig. 10), indicating a stronger human-perceived heat burden change for future urban residents than would be concluded from the temperature change

alone. This effect is particularly pronounced over mid-to-low latitude cities. Reanalysis and GCM simulations have shown a similar trend of larger increases in HI compared with air temperature alone over non-urban landscapes under climate change³⁷. Fortunately, heat stress can be notably mitigated through evaporative cooling in both present-day and future warmer climates. Because T_{wb} represents the lowest temperature that can be attained by a surface evaporating at its potential rate at constant pressure⁴², the difference between T_a and T_{wb} represents the potential efficiency of evaporative cooling from green infrastructure or human sweating. We illustrate this multi-model mean potential evaporative cooling efficiency over global urban areas in Fig. 5. In general, green infrastructure would be less effective in evaporative cooling in wetter cities than in dryer cities under both present-day and future climates. This supports green infrastructure intervention as an effective means of reducing urban heat stress for dryer cities. Despite being less effective in wetter cities, green infrastructure would still provide cooling and, in addition, more shading (by trees). Under climate change, the potential evaporative cooling efficiency from green infrastructure is projected to increase in most of the urban areas globally except some coastal cities (Fig. 5c), with a global average increase of 16.6% by the end of the century under RCP 8.5. This increase is largely caused by the multi-model projected drying trend in urban RH for global cities. Our results suggest that efforts to increase urban green infrastructure would potentially produce more effective cooling on a large scale in future warmer climates.

We note three main limitations to this study. First, the effect of future urbanization on urban climate is not embedded in the current emulator. Because the urban fraction and properties in each grid cell are time-invariant in the current version of the CESM (CESM version 2), the training sets produced by CESM simulations do not contain the effects of urban development. Our emulated results represent projections of climate change impacts on urban climates, and likely the lower bounds of potential urban warming

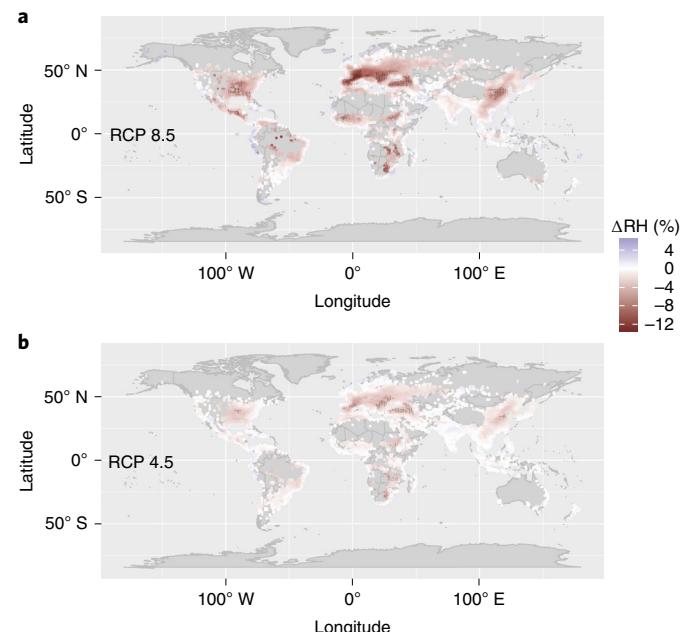


Fig. 4 | Multi-model mean urban relative humidity change for JJA. **a,b**, Seasonal mean urban RH changes between 2006–2015 and 2091–2100 under RCP 8.5 (**a**) and RCP 4.5 (**b**). Stippling indicates substantial change ($\text{abs}(\Delta \text{RH}) > 5\%$ under RCP 8.5 or $\text{abs}(\Delta \text{RH}) > 2.5\%$ under RCP 4.5) with high inter-model robustness ($\text{SNR} > 1$).

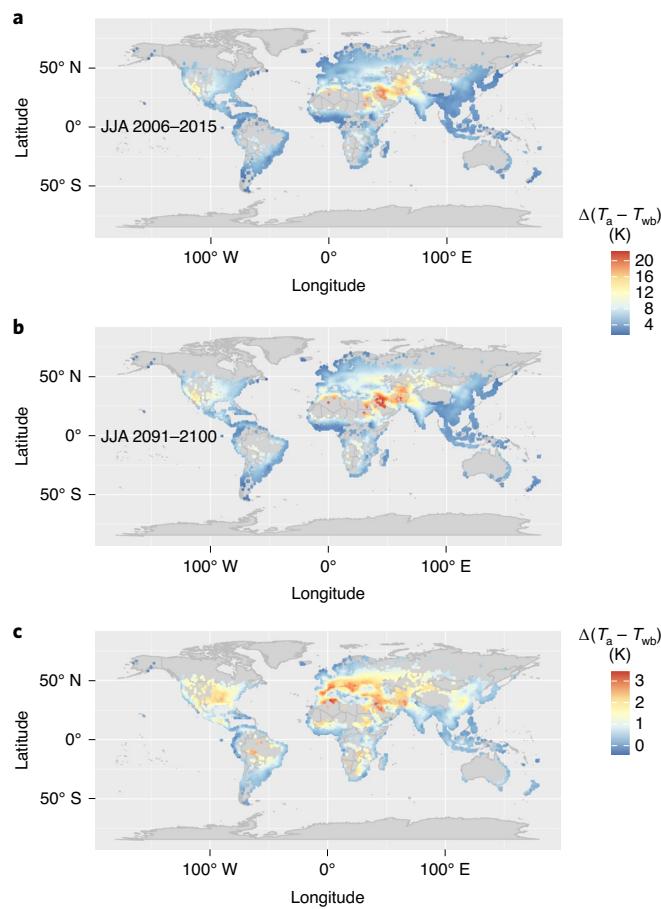


Fig. 5 | Multi-model mean urban evaporative cooling efficiency.

Multi-model mean urban evaporative cooling efficiency from green infrastructure or human sweating as measured by $T_a - T_{wb}$ for JJA under RCP 8.5. **a–c**, JJA mean during 2006–2015 (**a**), JJA mean during 2091–2100 (**b**) and JJA difference between 2006–2015 and 2091–2100 (**c**).

for growing cities in future climates, unless heat mitigation is explicitly taken into account in urban development^{11,32,43}. Importantly, our emulator strategy can be extended to include the repercussions of urban development. One could implement a future urban scenario in the CESM and use the urbanization-enabled CESM to train the emulator; it would then be able to capture both urbanization and urban physical processes in each urban grid cell. Second, the urban projections in this study are based on a single urban model, that is, the urban parameterization in the CESM. Developing emulators based on multiple urban models would help assess associated variability in future work. However, unlike ESMs, urban parameterizations do not have many internal dynamics and non-linearities, and as such cannot drift far from each other if forced by identical atmospheric forcing^{44,45}. This is supported by the comparison of the CESM and Weather Research and Forecasting (WRF) model results for the CONUS (Extended Data Fig. 2); despite having two different urban land schemes, the two models project similar warming when the atmospheric forcing is the same, with a mean absolute difference of 0.47 K (much smaller than the urban warming signals of either models). The major variability is from the diverse large-scale atmospheric forcings that are produced by different ESMs. Therefore, the multi-model results presented here are expected to encompass the potential variability resulting from urban models. Third, this emulator predicts city-scale urban climate rather than high-resolution intra-city values that other downscaling (dynamical and statistical)

techniques are capable of producing. Our framework operates at the city-to-global scale but lacks intra-city details that RCMs^{11–13,32,46,47} and microscale urban models^{48–50} can produce. The emulator can, however, be used to select future periods and regions where heat hazards are particularly intense and where downscaling to the ~1 km scale may be beneficial; it can thus guide dynamical downscaling efforts. If future generations of ESMs have finer-scale intra-city representations, one could easily train an emulator on those ESMs to provide multi-model high-resolution urban projections.

This study proposes a novel reduced-order urban climate emulator anchored in process-based simulations and physics-guided machine learning. Our framework makes urban areas globally comparable and facilitates understanding the impacts of climate change on local urban climates. The emulated global multi-model projections of local urban climates over the full time period (2006–2100) are provided (see Data availability), for those interested in extracting time periods of interest. We highlight (1) the necessity of modelling local urban-specific climates and quantifying robustness and uncertainties, and (2) the need for multi-model global projections of climate change-driven heat stress in urban locations. Our results can accurately account for the human-perceived impacts of climate change in population hotspots. Our emulator strategy is not restricted to urban systems; similar emulators can be developed for other systems such as rivers and lakes, glaciers or vegetated ecosystems. This approach can be adopted in other broad-scale climate change impacts, vulnerability and adaptation applications to reduce the need for computationally expensive (and maybe unattainable) downscaling, or to guide research for higher-resolution downscaling where information at finer spatial scales might be beneficial.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41558-020-00958-8>.

Received: 30 March 2020; Accepted: 27 October 2020;

Published online: 4 January 2021

References

1. Grimm, N. B. et al. Global change and the ecology of cities. *Science* **319**, 756–760 (2008).
2. Mora, C. et al. Global risk of deadly heat. *Nat. Clim. Change* **7**, 501–506 (2017).
3. Schneider, A., Friedl, M. A. & Potere, D. A new map of global urban extent from MODIS satellite data. *Environ. Res. Lett.* **4**, 044003 (2009).
4. Heilig, G. K. *World Urbanization Prospects: The 2011 Revision* (United Nations, 2012).
5. Cao, C. et al. Urban heat islands in China enhanced by haze pollution. *Nat. Commun.* **7**, 12509 (2016).
6. Zhao, L. et al. Interactions between urban heat islands and heat waves. *Environ. Res. Lett.* **13**, 034003 (2018).
7. Baklanov, A. et al. From urban meteorology, climate and environment research to integrated city services. *Urban Clim.* **23**, 330–341 (2018).
8. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
9. Knutti, R. & Sedláček, J. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Change* **3**, 369–373 (2013).
10. Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. R. Soc. A* **365**, 2053–2075 (2007).
11. Krayenhoff, E. S., Moustakou, M., Broadbent, A. M., Gupta, V. & Georgescu, M. Diurnal interaction between urban expansion, climate change and adaptation in US cities. *Nat. Clim. Change* **8**, 1097–1103 (2018).
12. Langendijk, G. S., Rechid, D. & Jacob, D. Urban areas and urban-rural contrasts under climate change: what does the EURO-CORDEX ensemble tell us?—Investigating near surface humidity in Berlin and its surroundings. *Atmosphere* **10**, 730 (2019).
13. Daniel, M. et al. Benefits of explicit urban parameterization in regional climate modeling to study climate and city interactions. *Clim. Dyn.* **52**, 2745–2764 (2019).

14. Li, D., Malyshev, S. & Sheviakova, E. Exploring historical and future urban climate in the Earth System Modeling framework: 1. Model development and evaluation. *J. Adv. Model. Earth Syst.* **8**, 917–935 (2016).
15. McCarthy, M. P., Best, M. J. & Betts, R. A. Climate change in cities due to global warming and urban effects. *Geophys. Res. Lett.* **37**, L09705 (2010).
16. Oleson, K. Contrasts between urban and rural climate in CCSM4 CMIP5 climate change scenarios. *J. Clim.* **25**, 1390–1412 (2011).
17. Hawkins, E. & Sutton, R. The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* **90**, 1095–1108 (2009).
18. Oleson, K. W., Anderson, G. B., Jones, B., McGinnis, S. A. & Sanderson, B. Avoided climate impacts of urban and rural heat and cold waves over the U.S. using large climate model ensembles for RCP8.5 and RCP4.5. *Climatic Change* **146**, 377–392 (2018).
19. Chapman, S., Watson, J. E. M., Salazar, A., Thatcher, M. & McAlpine, C. A. The impact of urbanization and climate change on urban temperatures: a systematic review. *Landsc. Ecol.* **32**, 1921–1935 (2017).
20. Hoornweg, D. & Pope, K. Population predictions for the world's largest cities in the 21st century. *Environ. Urban.* <https://doi.org/10.1177/0956247816663557> (2016).
21. Hurrell, J. W. et al. The Community Earth System Model: a framework for collaborative research. *Bull. Am. Meteorol. Soc.* **94**, 1339–1360 (2013).
22. Oleson, K., Bonan, G., Feddema, J., Vertenstein, M. & Kluzeck, E. in *Technical Description of an Urban Parameterization for the Community Land Model (CLMU)* 169 (National Center for Atmospheric Research, 2010).
23. Zhao, L., Lee, X., Smith, R. B. & Oleson, K. Strong contributions of local background climate to urban heat islands. *Nature* **511**, 216–219 (2014).
24. Demuzere, M. et al. Impact of urban canopy models and external parameters on the modelled urban energy balance in a tropical city. *Q. J. R. Meteorol. Soc.* **143**, 1581–1596 (2017).
25. Karsisto, P. et al. Seasonal surface urban energy balance and wintertime stability simulated using three land-surface models in the high-latitude city Helsinki. *Q. J. R. Meteorol. Soc.* **142**, 401–417 (2016).
26. Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M. & Grimmond, C. S. B. An urban parameterization for a global climate model. Part I: formulation and evaluation for two cities. *J. Appl. Meteorol. Climatol.* **47**, 1038–1060 (2008).
27. Demuzere, M., Oleson, K., Coutts, A. M., Pigeon, G. & van Lipzig, N. P. M. Simulating the surface energy balance over two contrasting urban environments using the community land model urban. *Int. J. Climatol.* **33**, 3182–3205 (2013).
28. Demuzere, M., De Ridder, K. & Van Lipzig, N. P. M. Modeling the energy balance in Marseille: sensitivity to roughness length parameterizations and thermal admittance. *J. Geophys. Res. Atmos.* **113**, D16120 (2008).
29. Fitria, R., Kim, D., Baik, J. & Choi, M. Impact of biophysical mechanisms on urban heat island associated with climate variation and urban morphology. *Sci. Rep.* **9**, 1–13 (2019).
30. Fischer, E. M., Oleson, K. W. & Lawrence, D. M. Contrasting urban and rural heat stress responses to climate change. *Geophys. Res. Lett.* **39**, L03705 (2012).
31. Oleson, K. W., Bonan, G. B., Feddema, J. & Jackson, T. An examination of urban heat island characteristics in a global climate model. *Int. J. Climatol.* **31**, 1848–1865 (2011).
32. Argüeso, D., Evans, J. P., Fita, L. & Bormann, K. J. Temperature response to future urbanization and climate change. *Clim. Dyn.* **42**, 2183–2199 (2014).
33. Oke, T. R. The energetic basis of the urban heat island. *Q. J. R. Meteorol. Soc.* **108**, 1–24 (1982).
34. Oleson, K. W., Bonan, G. B., Feddema, J. & Vertenstein, M. An urban parameterization for a global climate model. Part II: sensitivity to input parameters and the simulated urban heat island in offline simulations. *J. Appl. Meteorol. Climatol.* **47**, 1061–1076 (2008).
35. Screen, J. A. & Simmonds, I. The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature* **464**, 1334–1337 (2010).
36. Markon, C. et al. in *Fourth National Climate Assessment. Volume II: Impacts, Risks, and Adaptation in the United States*, 1185–1241 (US Global Change Research Program, 2018); <https://doi.org/10.7930/NCA4.2018.CH26>
37. Li, J., Chen, Y. D., Gan, T. Y. & Lau, N.-C. Elevated increases in human-perceived temperature under climate warming. *Nat. Clim. Change* **8**, 43–47 (2018).
38. Willett, K. M., Gillett, N. P., Jones, P. D. & Thorne, P. W. Attribution of observed surface humidity changes to human influence. *Nature* **449**, 710–716 (2007).
39. Luo, M. & Lau, N.-C. Urban expansion and drying climate in an urban agglomeration of east China. *Geophys. Res. Lett.* **46**, 6868–6877 (2019).
40. Lokoshchenko, M. A. Urban heat island and urban dry island in Moscow and their centennial changes. *J. Appl. Meteorol. Climatol.* **56**, 2729–2745 (2017).
41. Moriwaki, R., Watanabe, K. & Morimoto, K. Urban dry island phenomenon and its impact on cloud base level. *J. Jpn. Soc. Civil Eng.* **1**, 521–529 (2013).
42. Brutsaert, W. *Evaporation into the Atmosphere: Theory, History and Applications* (Springer, 1982).
43. Zhao, L., Lee, X. & Schultz, N. M. A wedge strategy for mitigation of urban warming in future climate scenarios. *Atmos. Chem. Phys.* **17**, 9067–9080 (2017).
44. Grimmond, C. S. B. et al. The international urban energy balance models comparison project: first results from phase 1. *J. Appl. Meteorol. Climatol.* **49**, 1268–1292 (2010).
45. Grimmond, C. S. B. et al. Initial results from Phase 2 of the international urban energy balance model comparison. *Int. J. Climatol.* **31**, 244–272 (2011).
46. Matte, D., Larsen, M. A. D., Christensen, O. B. & Christensen, J. H. Robustness and scalability of regional climate projections over europe. *Front. Environ. Sci.* **6**, 163 (2019).
47. Christensen, J. H., Larsen, M. A. D., Christensen, O. B., Drews, M. & Stendel, M. Robustness of european climate projections from dynamical downscaling. *Clim. Dyn.* **53**, 4857–4869 (2019).
48. Gromke, C. et al. CFD analysis of transpirational cooling by vegetation: case study for specific meteorological conditions during a heat wave in Arnhem, Netherlands. *Build. Environ.* **83**, 11–26 (2015).
49. Middel, A., Chhetri, N. & Quay, R. Urban forestry and cool roofs: assessment of heat mitigation strategies in Phoenix residential neighborhoods. *Urban For. Urban Green.* **14**, 178–186 (2015).
50. Huang, H.-Y., Margulis, S. A., Chu, C. R. & Tsai, H.-C. Investigation of the impacts of vegetation distribution and evaporative cooling on synthetic urban daytime climate using a coupled LES—LSM model. *Hydrol. Process.* **25**, 1574–1586 (2011).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

The fundamental rationale of our methodology was to use the atmospheric forcings from various ESMs to drive the CESM urban model in a statistical way rather than a physical way.

Community Earth System Model. The NCAR's CESM⁵¹ is a fully coupled ESM consisting of seven prognostic components, namely atmosphere (Community Atmosphere Model, CAM), land, land ice, ocean, sea ice, river and a coupler that exchanges fluxes between components. The land–atmosphere interactions are represented in its land component, the Community Land Model (CLM)⁵¹. The CLM employs nested hierarchical levels to represent land surface heterogeneity. For the first hierarchy, each grid cell consists of up to five subgrid 'land units': urban, vegetated, wetland, glacier and lake. These land units within a single grid cell are driven by the same atmospheric forcing at the reference height (from the lowest atmospheric layer in the CAM), because at this forcing height the air is well mixed. This architecture is an important feature as it allows isolation of local climate response (for example, urban climate response) in different landscapes to a given atmospheric forcing (for example, background climate forcing). To simulate the effects of different landscapes on local climate, the different landscapes in a grid cell must be forced with the same atmospheric forcing. This modelling structure is, to some extent, similar to those mesoscale modelling studies performed using the WRF model, in which the forcing meteorology is usually a background regional atmosphere coming from reanalysis data or GCM-modelled atmospheric data.

The CESM is the only model in CMIP5 that has a physically based urban land parameterization (Community Land Model Urban, CLMU) as part of its land model that is sufficiently detailed for the purposes of the study. This urban scheme is based on an 'urban canyon' conceptual structure that comprises roof, wall (sunlit and shaded) and canyon floor (pervious and impervious) facets⁵². A global urban surface dataset specifically developed for global climate modelling⁵² is embedded in the CLMU. This dataset prescribes the present-day urban extent and properties, including thermal (for example, heat capacity and thermal conductivity), radiative (for example, albedo and emissivity) and morphological (for example, building height-to-street width ratio, roof areal fraction, average building height and pervious ground fraction) characteristics for every grid cell having an 'urban' subgrid land unit. The urban spatial extent is derived from a population density dataset at 1 km spatial resolution. The urban property data are compiled by synthesizing a variety of datasets, including satellite products, a global database of tall buildings, local building codes data and other municipal documentation, and validated against imagery from Google Earth (see ref. ⁵² for details). With this high-dimensional urban dataset and the prognostic forcings from the CAM, the CESM directly outputs state and flux variables and the building energy use for the urban land unit in each grid cell, laying the groundwork for assessing the urban climate on a large scale.

The CLMU uses a simplified scheme to dynamically capture the building energy use (primarily space heating and air conditioning) and associated waste heat in urban areas⁵². The internal boundary conditions for roofs and walls are determined by an approximation of the internal building temperature held between prescribed maximum and minimum temperatures. Building interior maximum and minimum thermostat settings are prescribed by region from ref. ⁵². The amount of energy required to be added to increase the interior building temperature to the minimum temperature and the amount required to be removed to reduce the interior building temperature to the maximum temperature are designated as the space-heating and air-conditioning fluxes, respectively. The heat removed by air conditioning is added as waste heat (sensible heat) to the canyon floor. Waste heat from inefficiencies in the heating and air-conditioning equipment and from energy lost in the conversion of primary energy sources to end-use energy is also added as sensible heat to the canyon floor. The anthropogenic heat flux due to traffic and human metabolism is not parameterized in the current version of the CLMU.

The CLM–CAM coupling in the model is a two-way dynamic interaction. In a fully coupled CESM simulation, the biogeochemical component of the CLM captures leaf and stem area and canopy height of vegetated surfaces prognostically in response to climate change⁵³. One limitation of the current urban configuration in the CESM is that the urban extent and properties are static, and thus urbanization effects are not represented in CESM transient climate simulations. Therefore, this study provides global urban climate projections under greenhouse gas forcing only, but an emulator including the urban expansion effects can be developed when the needed CESM simulations become available. Note that a future urban scenario with urban expansion (not densification or other morphological changes) would not affect the urban subgrid outputs in the CESM. Its effects on grid cell means are also minimal due to the small fraction (2%) of urban areas on the Earth's land surface, even in grid cells with large cities. Compared with large-scale climate change, the impacts of urbanization on local urban climates are minimal, as is evidenced by the recent study incorporating urban expansion and densification scenarios into their WRF downscaling simulations for the United States¹¹.

Reduced-order urban climate emulator. The urban climate emulator is a location-dependent (grid cell) regression model based on atmospheric forcings,

time and their interactions. For the scope of this study, we built and applied the emulator on a monthly scale. It ingests the monthly average atmospheric forcings, month of year, two-way interaction terms and location as the inputs, and then outputs the monthly mean urban climatic response R_{urb} :

$$R_{urb} = f_{lat, long}(\mathbf{AF}, m, I) = \beta_{lat, long} \times \mathbf{AF} + \gamma_{lat, long} \times m + \theta_{lat, long} \times (\mathbf{AF} \times m) \quad (1)$$

where \mathbf{AF} is the matrix of monthly average atmospheric forcings from the atmosphere component of the models, m is the matrix of month-of-year indicator, I represents the two-way interaction terms between \mathbf{AF} and m , essentially accounting for the seasonality of the atmospheric forcing variables, lat and long are the latitude and longitude specifying the grid cell geolocation, respectively, and $f_{lat, long}$ denotes the location-dependent regression functions as specified by the model coefficients ($\beta_{lat, long}$, $\gamma_{lat, long}$ and $\theta_{lat, long}$). For \mathbf{AF} , we incorporated all the atmospheric forcing fields that drive the CLM in the CESM except three variables (lightning frequency, nitrogen deposition rate and aerosol deposition rate) that do not affect urban climate. This ensures that the emulator mimics the dynamic interaction between the land and atmosphere in a CESM simulation. These atmospheric forcing fields include the atmospheric bottom level (lowest height) in the atmosphere component of the model, incident shortwave and longwave radiation, precipitation (liquid and solid), atmospheric temperature, pressure, specific humidity and wind speed (zonal and meridional) at the forcing height. The latitude- and longitude-dependent coefficients of the regression equation ($\beta_{lat, long}$, $\gamma_{lat, long}$ and $\theta_{lat, long}$) make the emulator location-specific so that it maps from atmospheric forcings to local urban climate specifically for each urban land unit. The impacts of urban surface characteristics on its climate, and its specific response to that climate, are embedded in these coefficients and the location-specific regression function f . Therefore, applying the emulator to other CMIP5 ESMs does not need the high-dimensional global urban dataset as long as various model latitude/longitude grids align. Compared with the original CLMU, the emulator requires far less input datasets and computation, making it a reduced-order model of urban climates. Higher-capacity models for function f (including the support vector regression and random forest models) were also investigated, found to add little skill in model prediction and therefore deemed unnecessary.

The outputs R_{urb} in this study included the monthly average urban 2 m diurnal mean (T_0), the diurnal maximum (T_{max}) and minimum (T_{min}) air temperatures, and urban 2 m relative humidity (RH). The emulator can be expanded to include any variables that the CLM produces for the urban land unit.

We used fully coupled CESM simulations to provide training datasets to fit the emulator. The model fitting was conducted on each urban land unit to produce location-specific regression functions $f_{lat, long}$. To apply the emulator to other CMIP5 ESMs, consistency of the atmospheric forcings across models is required. Because all the other ESMs do not have urban representation in their domain, one could argue that the atmospheric forcings in the CESM might already be 'contaminated' by the impacts of urban land units in the CLM. Both our simulations (not shown) and previous studies have demonstrated the indiscernible effects of urban areas on grid cell-level atmospheric variables^{53,55,56}, because the model is normally run at a coarse spatial resolution. To minimize any potential biases introduced into the emulator, we conducted two additional simulations identical to the CESM CMIP5 RCP 4.5 (ref. ⁵⁵) and RCP 8.5 (ref. ⁵⁶) runs from 2006 to 2100 but without urban land units represented (replaced with bare soil land units in the grid cells) at a resolution of 0.9° latitude × 1.25° longitude. We compared the atmospheric forcings generated from these two simulations with those from the original CESM simulations (that is, ones with urban land units) and found indiscernible differences (<0.3% difference in 50 atmospheric variables checked, including all forcing fields used in the emulator). We used the entire range (95 years) of the monthly atmospheric forcings from these two simulations without urban land, and the CLM urban outputs from the original CESM–CMIP5 runs with urban land over the same time span (2006–2100) as the training sets to build the emulator. This emulator method, however, is not limited to monthly outputs. A daily or hourly scale emulator could be trained on daily or hourly data of the CESM simulations, and then applied to other ESMs as long as their daily or hourly atmospheric forcings are available.

The emulator does not assume time-stationary relationships between urban response and the atmospheric forcings. If the emulator were trained on a shorter timespan, the regression functions would be different. When the emulator is applied to other ESMs, it needs to be applied to the same time range to avoid extrapolation.

Using atmospheric forcings from 'fully coupled' CESM runs to drive the CLM is not equivalent to an 'offline' simulation. An offline simulation implies no two-way interactions between the atmosphere and land. In other words, any changes to land surface state variables would not affect the state of the atmosphere. Typical examples of offline simulations include the CLM forced by site observations or reanalysis. However, forcing the CLM with the atmosphere fields from fully coupled CESM runs (with the same CLM) is actually a shortcut to the fully coupled mode and should be considered as a retrieval of the land surface variables from the fully coupled runs. It produces nearly identical outputs to those obtained from the fully coupled CESM runs, because the two-way land–atmosphere interactions and impacts of large-scale feedbacks (for example, large-scale dynamics, ocean–air feedbacks, carbon climate feedbacks

and dynamic land use/land cover change) and regional dynamics (for example, influence of topography on regional atmospheric circulation and atmospheric rivers) represented in a fully coupled configuration have been preserved in the atmospheric forcings^{6,43}. In addition, this emulator framework is not restricted to GCM/ESM. One could train the emulator using an RCM for a given region (or city) to reproduce mesoscale and sub-city-scale variability, if finer-scale dynamics are important in the application.

Validation of the CLMU and urban climate emulator. We conducted a thorough validation of the whole emulator framework. We first evaluated the process-based CLMU simulation against a ground-based observation dataset and mesoscale modelling simulations. We then validated the statistical robustness of the emulator.

Validation of the CLMU. The CESM-CLMU was evaluated against both in situ and remote sensing observations in previous efforts^{23–29}. Here, we further evaluated the performance of the CLMU against the 4-km gridded observationally based PRISM dataset (<http://www.prism.oregonstate.edu/>) and the 20-km WRF simulations coupled with the single-layer urban canopy model (SLUCM)⁵⁷. The WRF-modelled results were taken from a recent study over the CONUS¹¹. The WRF simulation itself has been extensively validated against observations¹¹. To ensure consistency in this evaluation, we forced the CLMU with the same boundary conditions as used in ref. ¹¹ (ERA-Interim reanalysis), at a resolution of 0.9° latitude \times 1.25° longitude. The same ten US metropolitan areas as in ref. ¹¹, across a range of regional background climate zones, were selected for assessment. The grid cells in the CESM domain where the selected cities reside were compared with the corresponding PRISM and WRF grid cells. The results show that the CLMU reproduces both the observed and WRF-modelled daily distributions of T_{max} , T_{min} and T_a reasonably well for JJA 2000–2009 across the ten cities (Extended Data Fig. 1). The bias in the median of the CLMU compared with the WRF data is improved with respect to the observations for four, four and five cities for T_a , T_{min} and T_{max} , respectively. The variability of the daily distributions modelled by the CLMU is slightly smaller than that modelled by the WRF. The reason for this lies in the different nature of the CLMU and WRF-SLUCM. As an urban scheme in an ESM, the CLMU has to limit its ability to capture some mesoscale physical processes and rely on parameterizations, whereas the mesoscale WRF model is able to resolve certain detailed physics. Therefore, it is expected that some diurnal and daily variabilities could be under-represented in the CLMU simulation compared with the WRF-SLUCM. The validation here demonstrates the ability of the CLMU forced by present-day meteorology to reproduce the observed urban climatology, even at a spatial resolution coarser than that of the WRF simulation.

To assess the performance of the CLMU in a changing climate, we further evaluated the CESM urban warming projection produced by fully coupled simulation against the WRF downscaling simulations over the CONUS. For consistency with our CESM configuration, we used the WRF-SLUCM simulations with present-day urban land representation in both present-day and projected climate simulations. The WRF-modelled urban warming was derived from the difference between the simulation forced by ERA-Interim reanalysis (2000–2009) and the one forced by a CESM-projected climate under RCP 8.5 (2090–2099)¹¹. We area-weighted aggregated the WRF 20-km grid cells (Extended Data Fig. 2a) to the CESM resolution (Extended Data Fig. 2c) and compared their urban warming signals. The results demonstrate good agreement between the CESM-projected and WRF-SLUCM-downscaled urban warming under climate change in terms of both magnitude and spatial variation (Extended Data Fig. 2). The mean absolute difference between the CESM and WRF-SLUCM is 0.47 K, with 10th and 90th quantiles of -0.71 and 0.73 K, respectively. The validation further confirms the fidelity of the CESM to project the impacts of climate change on local urban climatology with similar accuracy to mesoscale dynamical downscaling despite the coarser spatial resolution.

Validation of the urban climate emulator. We further evaluated the statistical robustness of the emulator. We cross-validated our emulator against five other CESM ensemble members that were excluded from the emulator training under RCP 8.5 and RCP 4.5 (<http://www.cgd.ucar.edu/projects/chsp/brace-output-projections.html>). The results demonstrated the excellent accuracy of the emulator in predicting urban climate variables. The out-of-sample global monthly average root-mean-square errors (RMSEs) across ensemble members were 0.2 K, 0.3 K, 0.2 K and 1% for T_a , T_{max} , T_{min} and RH, respectively, under both scenarios (Extended Data Fig. 3). The ‘error’ here is defined as the difference between monthly values modelled by the CESM ensemble member and those modelled by the emulator using atmospheric forcings from the same ensemble member. These numbers are markedly smaller than the urban temperature differences between ensemble members ($\sim 1.5\text{--}2.0$ K, $P < 10^{-6}$), confirming the credibility of the emulator. We also conducted a tenfold cross-validation by randomly selecting 60 years of training data to train the emulator and the remaining 35 years were used for validation. The maximum tenfold cross-validated RMSEs were less than 0.5 K, 0.6 K, 0.5 K and 3% for T_a , T_{max} , T_{min} and RH, respectively, further demonstrating the robustness of our emulator method.

Multi-model urban climate projection. The CESM architecture described above provides the fundamental basis for our emulator to be applied to other ESMs.

The structure of the CLMU-CAM interactions in the CESM is, to some extent, analogous to dynamical downscaling studies using GCMs or reanalysis data, except that dynamical downscaling includes fine-scale (subgrid to ESMs but resolved by RCMs) horizontal interactions such as advection and turbulent transport. The atmospheric forcing data of the type provided by the CESM-CAM can be consistently extracted from other CMIP5 ESMs. These background forcings that drive the CLMU in our emulator framework are similar to the boundary conditions that force the WRF. Like the GCM/reanalysis-produced boundary conditions in WRF downscaling studies, the atmospheric forcings from the CMIP5 models are representative of a non-urban background atmosphere. One difference, as described above, is the representation of small-scale horizontal exchanges in the WRF.

We applied the emulator across all available CMIP5 ESMs to generate multi-model global urban projections over 2006–2100 under RCP 4.5 and RCP 8.5. We tried to use all available ESM RCP runs in CMIP5. Those from the exact same models but only at lower spatial resolutions were excluded. We ended up with a total of 26 ESMs for RCP 8.5 and 25 ESMs for RCP 4.5 (MRI-ESM1 does not have RCP 4.5 runs; Supplementary Table 1). We selected the first ensemble member from each of these ESMs. Because our emulator is geolocation-specific and thus requires the grid cells from other ESMs to align with the CESM grid, we regridded the needed atmospheric forcing fields from these ESMs to the CESM grid. The regridded atmospheric forcings were used as inputs to emulate 26 (25) global urban projections of monthly T_a , T_{max} , T_{min} and RH. The final multi-model analyses were based on these emulated projections plus the original CESM projections.

The fundamental rationale of our methodology was to use the atmospheric forcings from various ESMs to drive the CLMU statistically. The variability and uncertainty in large-scale dynamics, feedbacks and climatology among CMIP5 models are presumably preserved in those atmospheric forcing fields. The variability in the final urban climate projections reflects the impacts of large-scale climatology on local-scale urban climates.

Implications of the emulator framework. The urban climate emulator framework is essentially a statistical/machine-learning technique at its core, but trained on a physical downscaling urban model (CLMU) embedded in a fully coupled ESM (CESM). The physics represented in the CESM, especially the dynamic interactions between urban land, lower atmosphere and large-scale climate change, are therefore preserved in the emulator. This new framework falls into the physics-guided machine-learning (PGML) paradigm, because the derived statistical model is informed by the physics simulated in the ESM. We highlight here three key novel features of this emulator method.

1. The emulator framework is a globally consistent, robust and coherent method for investigating multi-model urban climate simulations across the globe. It is based on global-scale simulations that follow a well-defined and coordinated modelling protocol (CMIP), is calibrated on a common set of input variables (that is predictors) and operates at a common spatial resolution. The consistency and coherence of the method are critical for comparing urban climates and identifying risk hotspots across regions, especially on a large scale (for example, planetary scale). The lack of a common set of predictors and the use of different spatial domains and statistical families in existing statistical downscaling techniques have made direct comparison of the downscaled climates, and the downscaling methods themselves, difficult to achieve over broad scales. This has presented a problem for their applications to climate change impact assessments and climate uncertainty characterization^{58,59}.
2. Traditional statistical downscaling techniques seek to establish empirical relationships between local climate and the large-scale atmospheric state using weather observations and reanalysis/GCM output, whereas this emulator method is in essence a statistical ‘solver’ of the system of equations in the CLMU as it takes all the input variables that the CLMU needs in the coupled CESM to solve the equations. The only difference is that the emulator ‘solves’ the output variables statistically instead of numerically. The emulator’s functions are not to reproduce observed climates empirically, but rather to reproduce the dynamically modelled urban climates by the CLMU. Because of their empirical nature, traditional statistical downscaling is constrained by (1) the length and reliability of the observed historical data, (2) the choice of predictors, which remains the subject of debate and is often incomplete due to observational unavailability or model features, and (3) the omission of climate system feedbacks⁵⁸. There have been arguments that some of these challenges might be addressed by identifying more appropriate large-scale predictors and developing statistical methods that better represent the physical mechanisms⁶⁰. Our emulator method, trained on a fully coupled climate model and ingesting a complete set of variables required for dynamic simulations of the CLMU, captures the physical mechanisms and climate system feedbacks represented in the model.
3. One major limitation of traditional statistical downscaling is its accuracy of extrapolation^{58,60–62}. A fundamental assumption of the application of statistical downscaling is that the derived predictor/predictand relationship based on local observations remains valid outside of the training region and/or in a changing future climate. This assumption, however, has been shown to be questionable in observed records^{58,63–65}. This limitation is overcome in the emulator method, which is a location-dependent statistical model at the

global scale, and is trained on the entire temporal range of the fully coupled simulation (2006–2100) under climate change scenarios. The emulator, by design, is generalizable both spatially (across the globe) and temporally (with climate change feedbacks included).

Multi-model uncertainty and robustness. We used the SNR as a measure of model robustness to quantify how various models agree on the urban projections. The SNR is defined as the ratio of the multi-model ensemble mean to the inter-model variability:

$$\text{SNR}_{ij} = \frac{\mu_{ij}}{\sigma_{ij}} \quad (2)$$

where μ denotes the multi-model ensemble means, σ denotes the inter-model variability, defined as the standard deviation of multi-model values, and the ij pair denotes the index of grid cells that have an urban land unit in the CLM. The SNR can also be considered as an indicator of multi-model spread; the reciprocal of SNR is essentially a multi-model variance normalized by the mean. A smaller SNR indicates higher model spread or smaller signal change (projected change in the multi-model ensemble mean). We quantified the SNR for each grid cell that has an urban land unit to assess the inter-model agreement. We have used $\text{SNR} > 2.5$ and $\text{SNR} > 1$ to indicate high inter-model agreement in the projection of temperature and humidity, respectively. The models generally show less agreement in the modelling of humidity than temperature⁶⁰, therefore we selected a smaller SNR threshold for humidity. These thresholds chosen for the SNR are, to some degree, subjective and for illustration purposes only. Different thresholds could be chosen for different applications of urban climate change to indicate higher or lower tolerance in model disagreement.

Calculation of human-perceived heat stress. There are many methodologies for assessing human-perceived heat stress^{67–69}. We evaluated two commonly used heat stress indicators: the wet-bulb temperature (T_{wb}) and heat index (HI). Both measures take both temperature and humidity into account. The former has the added advantage of serving as an indicator of the evaporative cooling efficiency. The T_{wb} was computed by the formulation developed by Stull⁷⁰. We used a method to calculate the HI appropriate for the United States, that is, the National Weather Service definition. The HI was calculated using the Rothfusz regression with adjustments under certain conditions (<https://www.wpc.ncep.noaa.gov/html/heatindequation.shtml>):

$$\begin{aligned} \text{HI} = & -8.7847 + 1.6114T_a - 0.012308T_a^2 \\ & + \text{RH}(2.3385 - 0.14612T_a + 2.2117 \times 10^{-3}T_a^2) \\ & + \text{RH}^2(-0.016425 + 7.2546 \times 10^{-4}T_a - 3.582 \times 10^{-6}T_a^2) \end{aligned}$$

Both T_{wb} and HI were calculated from the emulated monthly means of urban T_a and RH.

Data availability

All CMIP5 data used in this study are available at the CMIP5 archive via <https://esgfnode.llnl.gov/projects/cmip5/> and the Climate Data Gateway at NCAR via https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm4.CESM_CAM5_BGC_LE.html for RCP 8.5 and https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm4.CESM_CAM5_BGC_ME.html for RCP 4.5. The output data from the emulator are available in the public repository—‘Illinois Data Bank’—via https://doi.org/10.13012/B21DB-4585244_V1. Source data are provided with this paper.

Code availability

The R (<https://www.R-project.org/>) and NCL (The NCAR Command Language, <https://doi.org/10.5065/D6WD3XH5>) codes of the urban climate emulator are available on the NCAR Cheyenne cluster (<https://doi.org/10.5065/D6RX99HX>) and on Github (https://github.com/zhao-research-lab/urban_climate_emulator; <https://doi.org/10.5281/zenodo.3893401>). The CESM (used to perform the simulations) source code releases are available through the public GitHub repository (<https://github.com/ESCOMP/CESM>).

References

51. Oleson, K. et al. in *Technical Description of Version 4.0 of the Community Land Model (CLM)* 257 (National Center for Atmospheric Research, 2010).
52. Jackson, T. L., Feddema, J. J., Oleson, K. W., Bonan, G. B. & Bauer, J. T. Parameterization of urban characteristics for global climate modeling. *Ann. Assoc. Am. Geogr.* **100**, 848–865 (2010).
53. Zhang, J. C., Zhang, K., Liu, J. F. & Ban-Weiss, G. Revisiting the climate impacts of cool roofs around the globe using an Earth system model. *Environ. Res. Lett.* **11**, 084014 (2016).
54. Hu, A. et al. Impact of solar panels on global climate. *Nat. Clim. Change* **6**, 290–294 (2016).
55. Sanderson, B. M., Oleson, K. W., Strand, W. G., Lehner, F. & O’Neill, B. C. A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. *Climatic Change* **146**, 303–318 (2018).
56. Kay, J. E. et al. The Community Earth System Model (CESM) large ensemble project: a community resource for studying climate change in the presence of internal climate variability. *Bull. Am. Meteorol. Soc.* **96**, 1333–1349 (2015).
57. Kusaka, H., Kondo, H., Kikegawa, Y. & Kimura, F. A simple single-layer urban canopy model for atmospheric models: comparison with multi-layer and slab models. *Boundary Layer Meteorol.* **101**, 329–358 (2001).
58. Fowler, H. J., Blenkinsop, S. & Tebaldi, C. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* **27**, 1547–1578 (2007).
59. Vaittinada Ayar, P. et al. Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations. *Clim. Dyn.* **46**, 1301–1329 (2016).
60. Tang, J. et al. Statistical downscaling and dynamical downscaling of regional climate in China: present climate evaluations and future climate projections. *J. Geophys. Res. Atmos.* **121**, 2110–2129 (2016).
61. Murphy, J. Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *Int. J. Climatol.* **20**, 489–501 (2000).
62. Spak, S., Holloway, T., Lynn, B. & Goldberg, R. A comparison of statistical and dynamical downscaling for surface temperature in North America. *J. Geophys. Res. Atmos.* **112**, D08101 (2007).
63. Fowler, H. J. & Kilsby, C. G. Precipitation and the North Atlantic Oscillation: a study of climatic variability in northern England. *Int. J. Climatol.* **22**, 843–866 (2002).
64. Wilby, R. L. Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices. *Clim. Res.* **10**, 163–178 (1998).
65. Slonosky, V. C., Jones, P. D. & Davies, T. D. Atmospheric circulation and surface temperature in Europe from the 18th century to 1995. *Int. J. Climatol.* **21**, 63–75 (2001).
66. Fischer, E. M. & Knutti, R. Robust projections of combined humidity and temperature extremes. *Nat. Clim. Change* **3**, 126–130 (2013).
67. Epstein, Y. & Moran, D. S. Thermal comfort and the heat stress indices. *Ind. Health* **44**, 388–398 (2006).
68. Buzan, J. R., Oleson, K. & Huber, M. Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5. *Geosci. Model Dev.* **8**, 151–170 (2015).
69. Martilli, A., Krayenhoff, E. S. & Nazarian, N. Is the urban heat island intensity relevant for heat mitigation studies? *Urban Clim.* **31**, 100541 (2020).
70. Stull, R. Wet-bulb temperature from relative humidity and air temperature. *J. Appl. Meteorol. Climatol.* **50**, 2267–2269 (2011).

Acknowledgements

We thank the National Center for Atmospheric Research (NCAR) for supercomputing and data storage resources, including the Cheyenne supercomputer (<https://doi.org/10.5065/D6RX99HX>), which were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. L.Z. and M.O. acknowledge the support from the High Meadows Foundation. K.O. acknowledges support by the US National Science Foundation (NSF) under grant no. AGS-1243095, and by NCAR, which is a major facility sponsored by the NSF under cooperative agreement no. 1852977. E.B.Z. acknowledges support by the Army Research Office under contract no. W911NF2010216 (program manager J. Barzyk), and the NSF under grant no. ICER 1664091 and SRN cooperative agreement no. 1444758. Q.Z. is supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area, Office of Biological and Environmental Research of the U.S. Department of Energy Office of Science.

Author contributions

L.Z. conceptualized the research and designed the emulator modelling framework; L.Z. developed the reduced-order emulator with contributions from K.O., A.B. and C.C.; K.O. and L.Z. conducted the CESM simulations; L.Z. performed the data analysis; K.O., E.B.Z. and M.O. contributed ideas to the data analysis; L.Z., K.O. and E.S.K. performed the model validation; L.Z. drafted the manuscript, with discussions and contributions from K.O., E.B.Z., E.S.K., M.O. and other co-authors.

Competing interests

The authors declare no competing interests.

Additional information

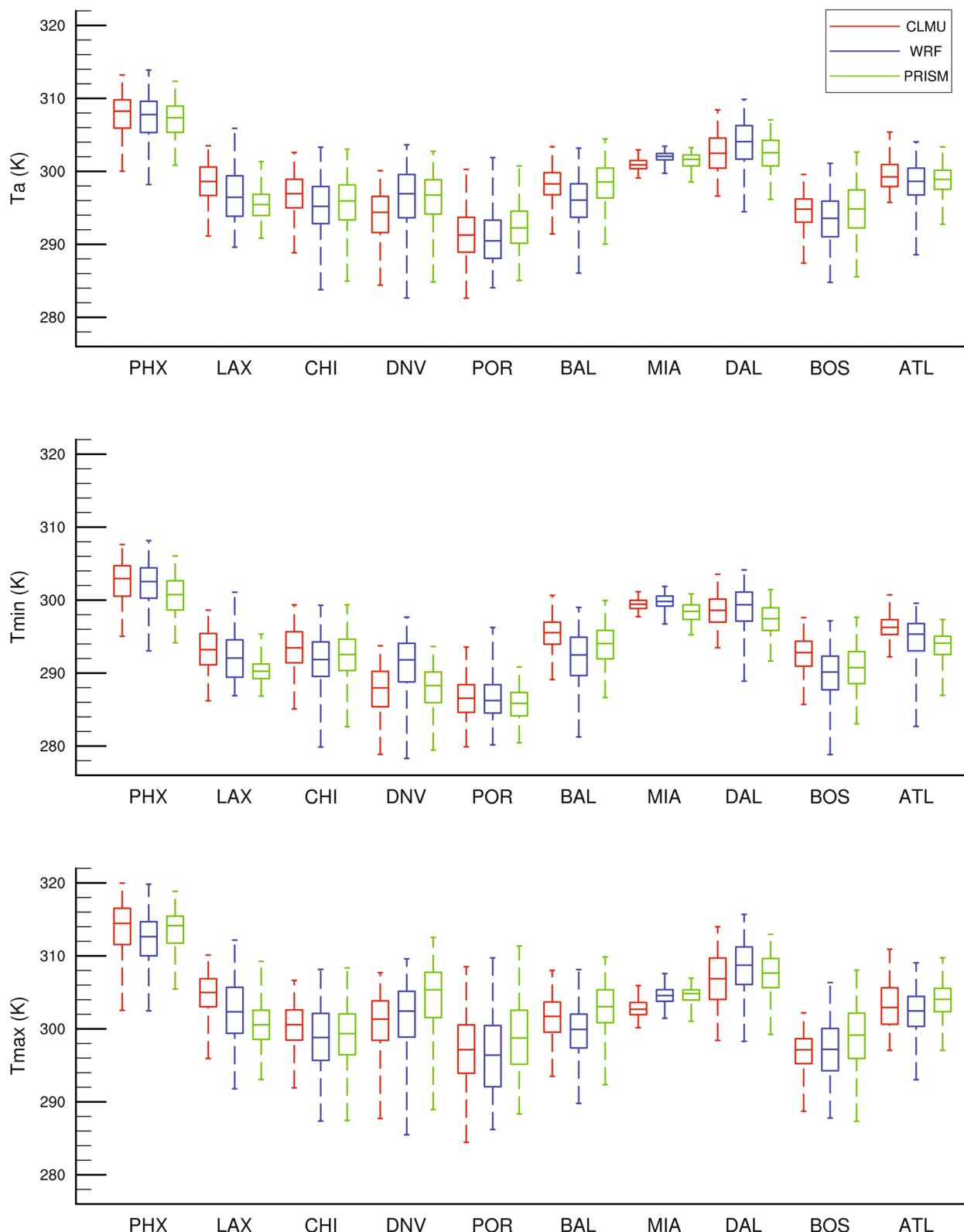
Extended data is available for this paper at <https://doi.org/10.1038/s41558-020-00958-8>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41558-020-00958-8>.

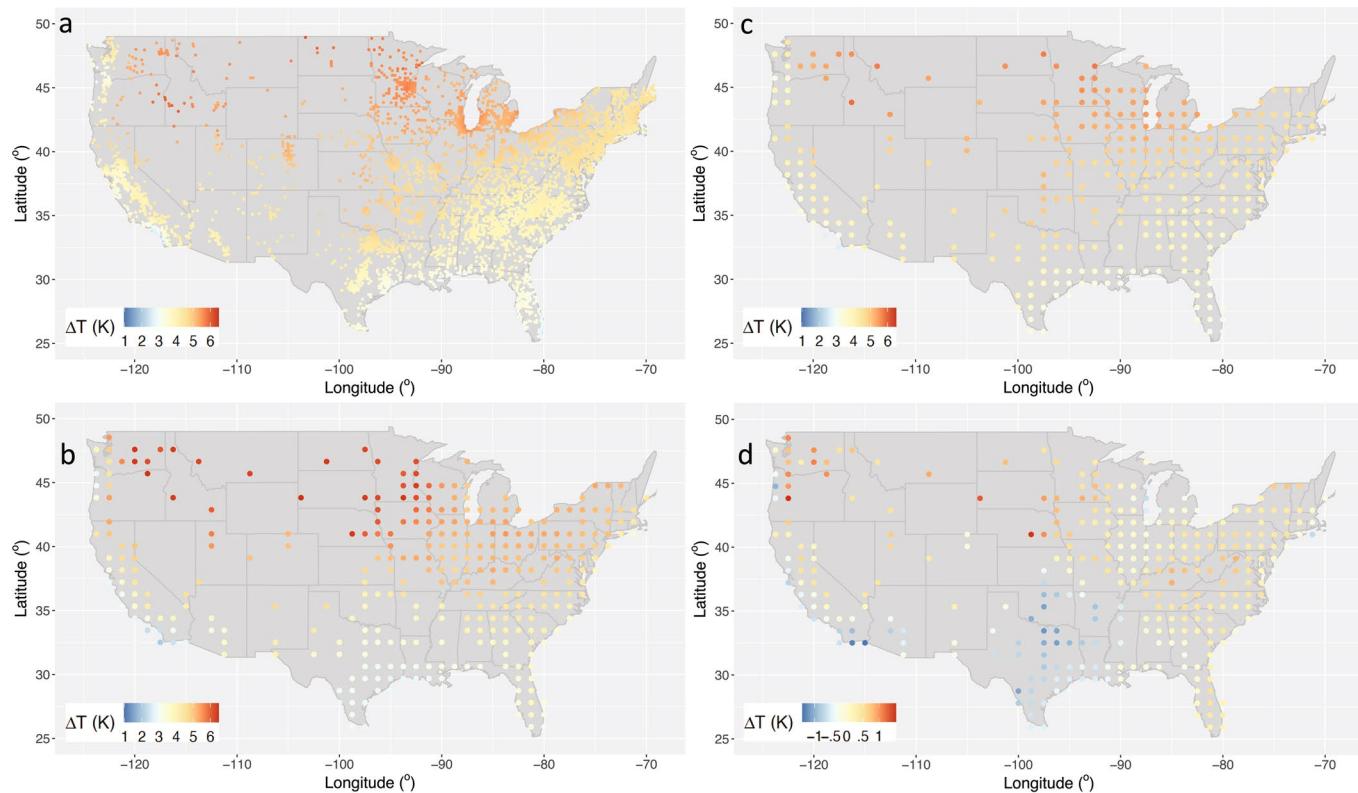
Correspondence and requests for materials should be addressed to L.Z.

Peer review information *Nature Climate Change* thanks Joshua Hacker, Gaby Langendijk and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

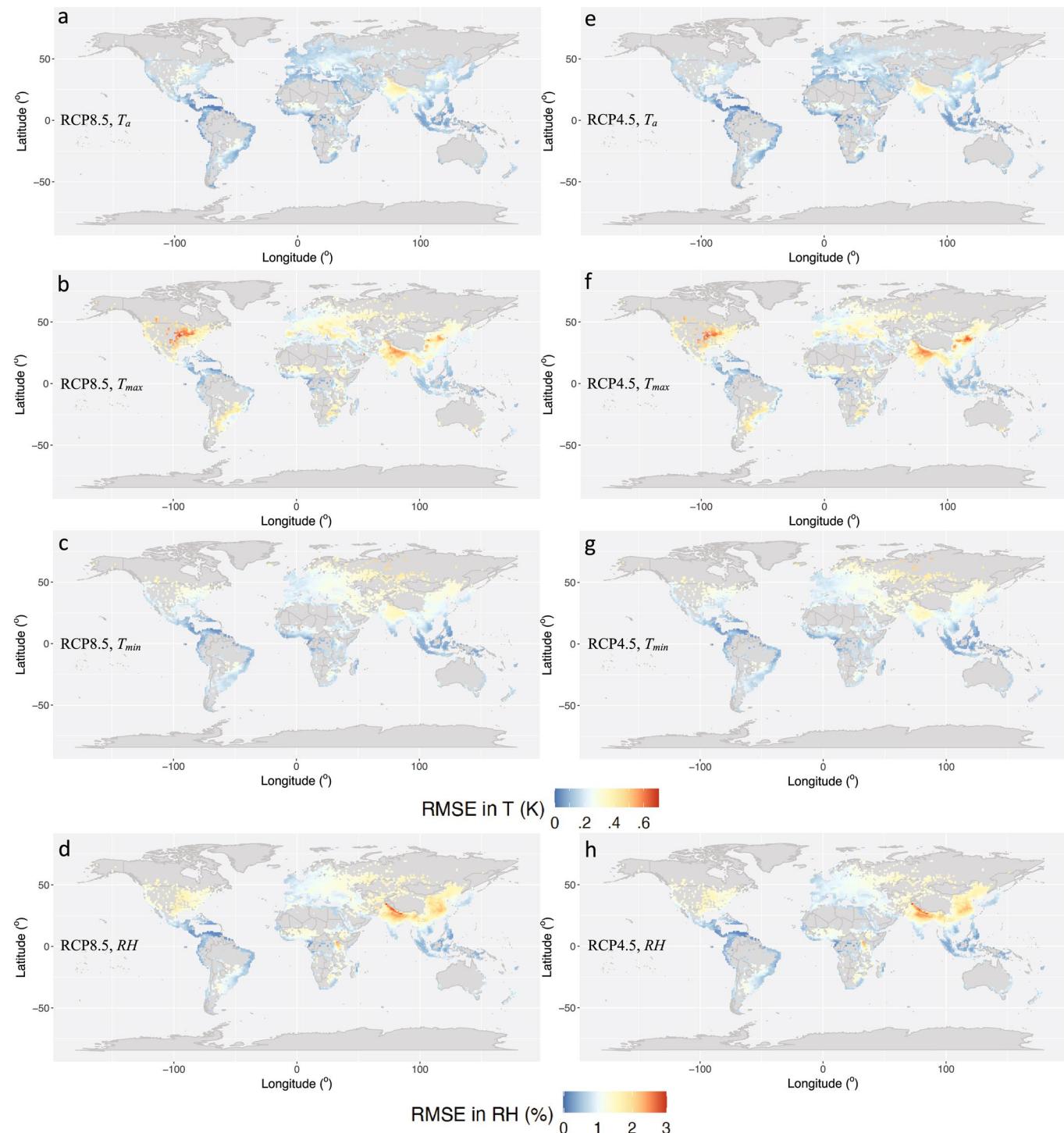
Reprints and permissions information is available at www.nature.com/reprints.



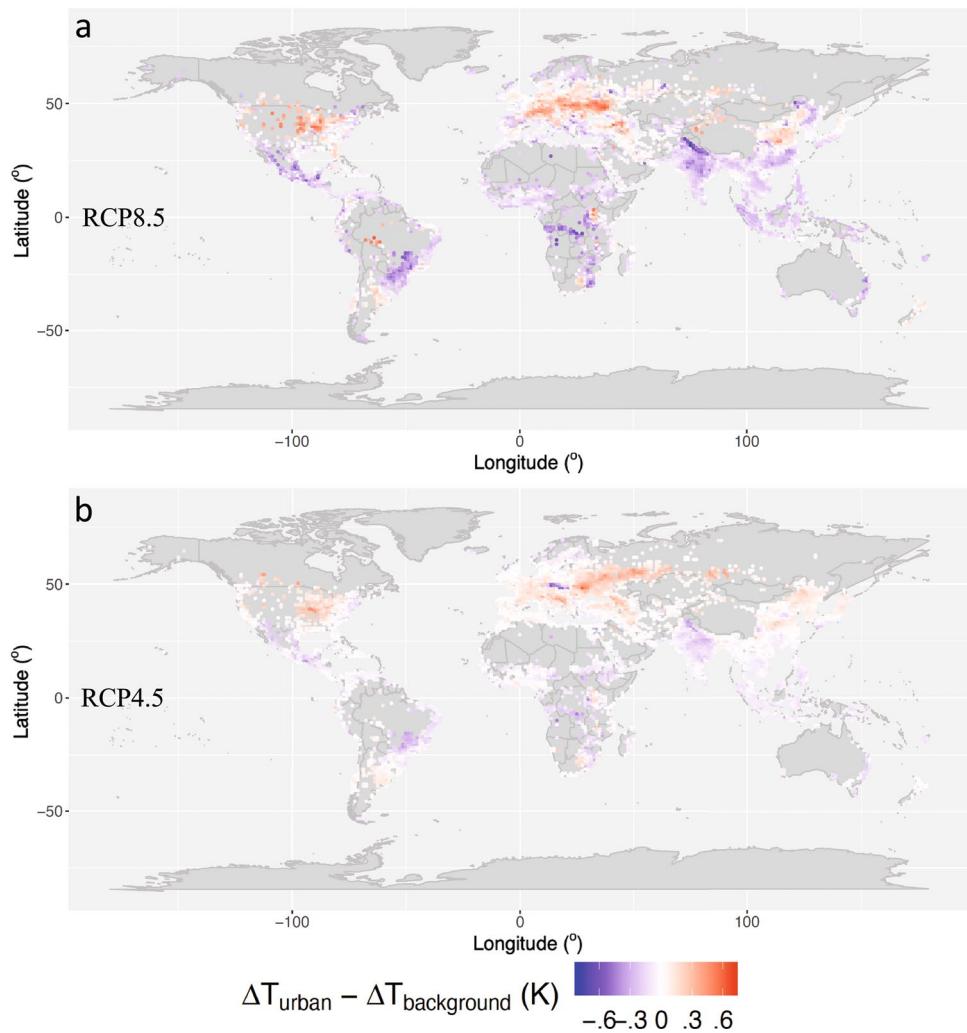
Extended Data Fig. 1 | Comparison of 2000–2009 JJA (June – August) distributions of diurnal average (T_a), minimum (T_{min}) and maximum (T_{max}) 2m urban air temperature between CLMU (red), WRF (blue) and PRISM (green). Center bar represent the median, box edges the 25th and 75th percentiles, and error bars the 1st and 99th percentiles. PHX = Phoenix; LAX = Los Angeles; CHI = Chicago; DNV = Denver; POR = Portland; BAL = Baltimore; MIA = Miami; DAL = Dallas; BOS = Boston; ATL = Atlanta.



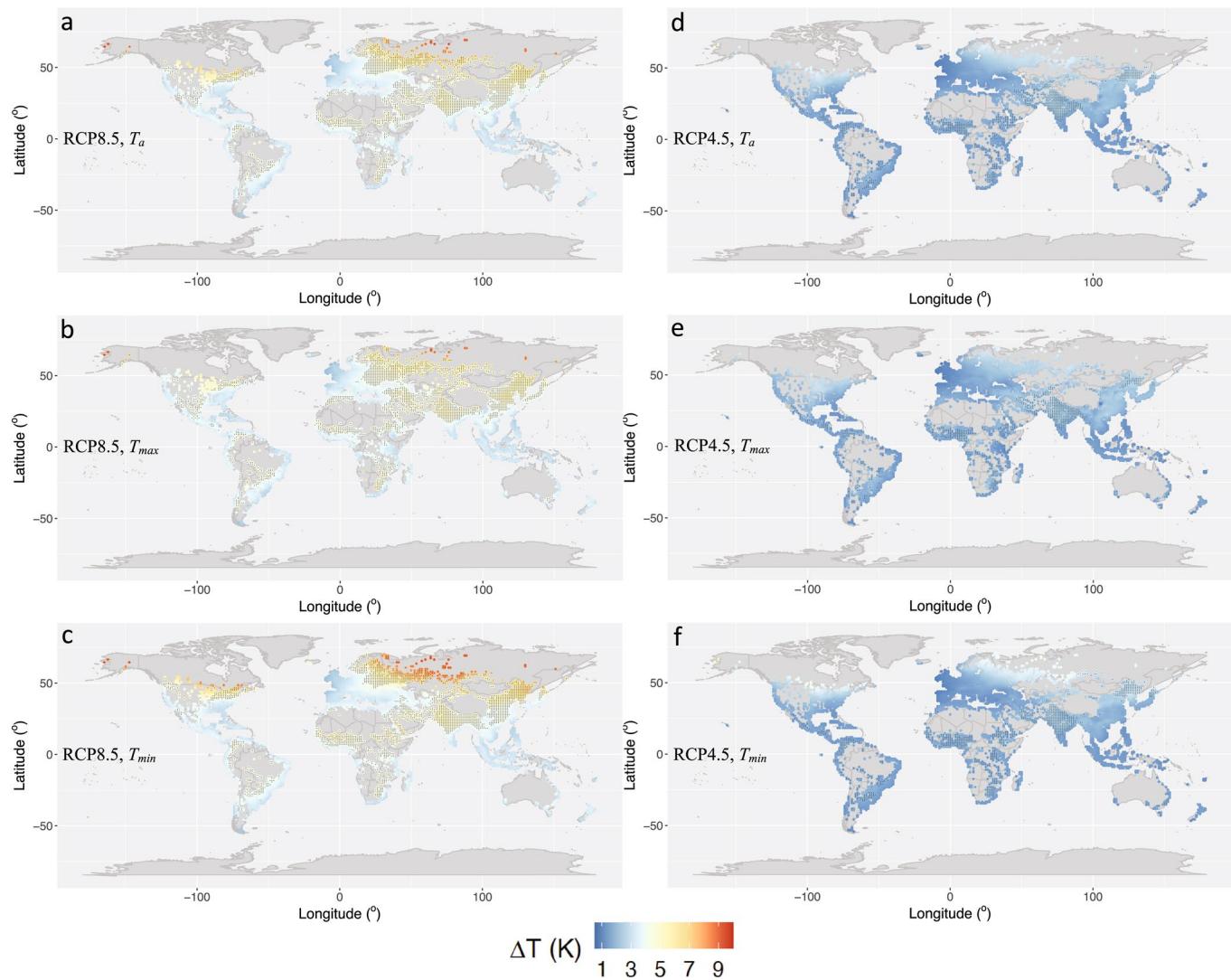
Extended Data Fig. 2 | Comparison of urban warming projection of diurnal average temperature between CESM and WRF-SLUCM simulation for season JJA (June - August) in 2090 - 2099 relative to 2000 - 2009 under RCP 8.5. WRF-SLUCM dynamic downscaling simulation (forced by a CESM-RCP8.5 meteorology) was conducted in ref. ¹¹. **a**, WRF-SLUCM projected urban warming in WRF's original 20 km grids; **b**, CESM projected urban warming in the CESM grids (0.9° latitude × 1.25° longitude); **c**, WRF-SLUCM projected urban warming in the aggregated CESM grids (0.9° latitude × 1.25° longitude); **d**, difference in projected urban warming between CESM and WRF-SLUCM in the CESM grids.



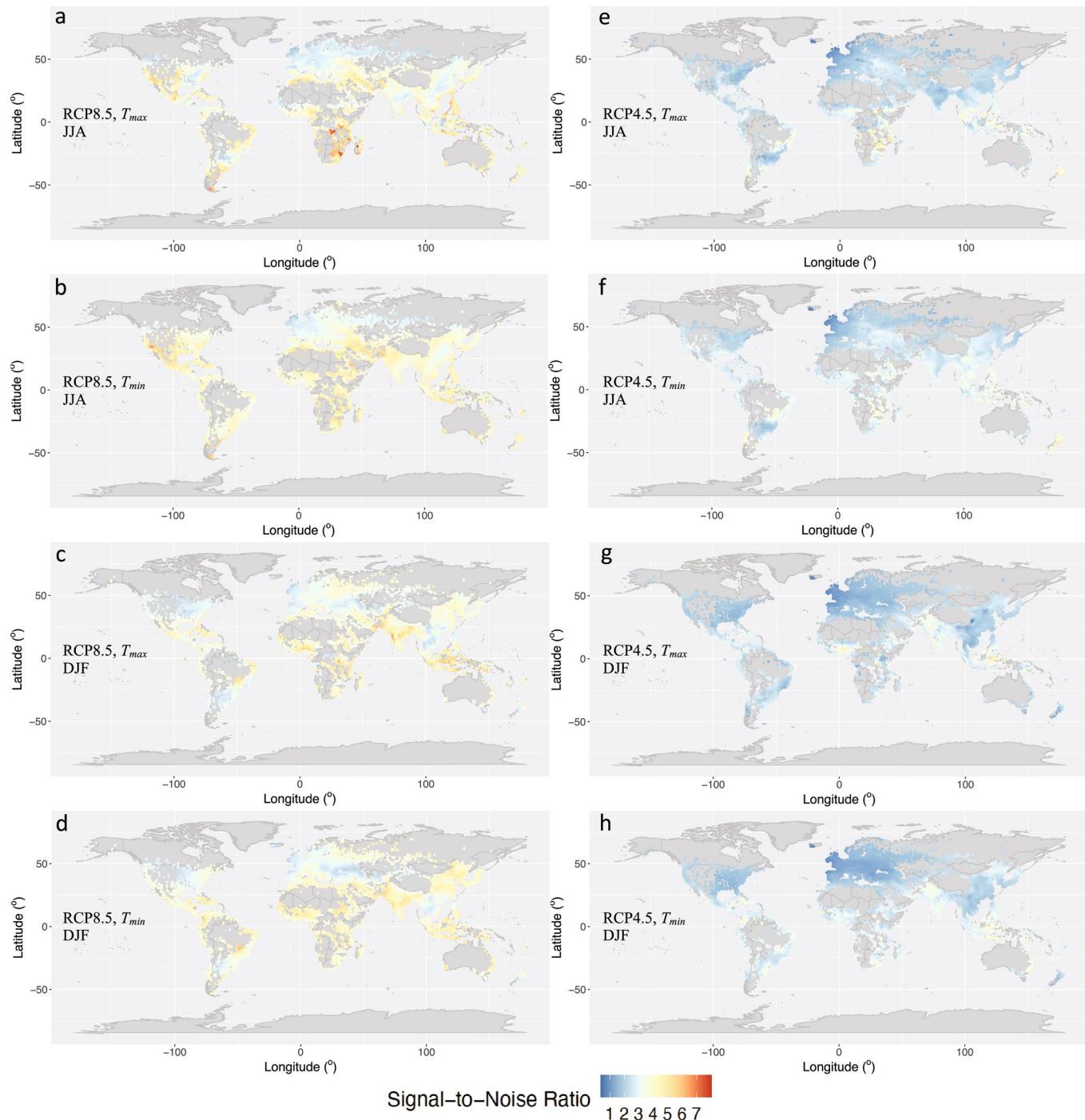
Extended Data Fig. 3 | Average root-mean-square-error (RMSE) in urban temperature and RH validation of the emulator across 5 CESM ensemble member runs (member #2 - #6). The 'error' in RMSE denotes the difference between monthly temperatures or RH dynamically modeled by the CESM ensemble member and the ones modeled by the emulator. The average RMSE was calculated based on the 5 CESM ensemble member runs from 2006 to 2100 (that is 95 years). a-d: RCP 8.5; e-h: RCP 4.5. a and e: RMSE in diurnal average temperature T_a ; b and f: RMSE in diurnal maximum temperature T_{max} ; c and g: RMSE in diurnal minimum temperature T_{min} ; d and h: RMSE in urban relative humidity RH .



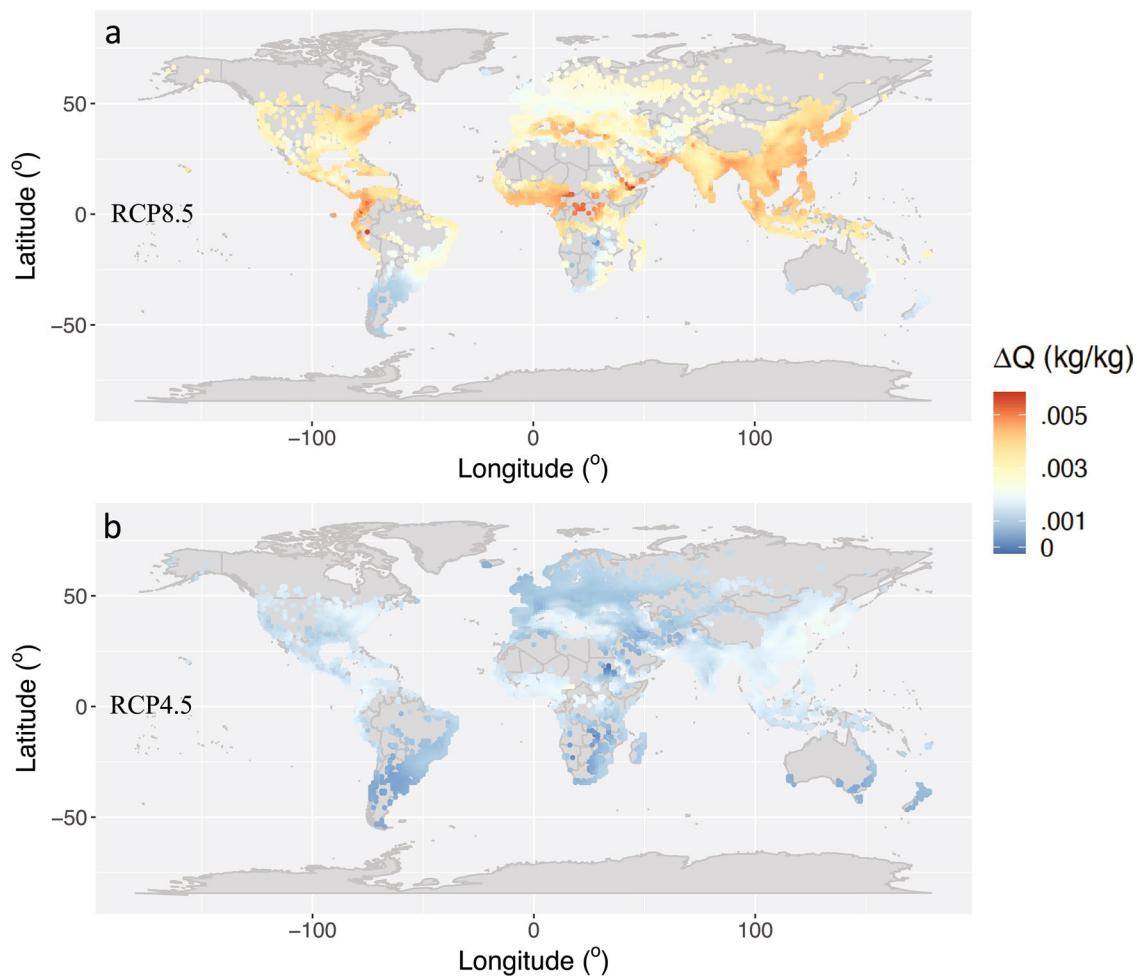
Extended Data Fig. 4 | Comparison of CMIP5 multi-model mean urban warming (ΔT_{urban}) and background regional warming ($\Delta T_{\text{background}}$) for season JJA (June – August) in 2091 – 2100 relative to 2006 – 2015. Δ refers to changes between end of the century and beginning of the century, that is (2091 to 2100) – (2006 to 2015); **a**, RCP 8.5; **b**, RCP 4.5.



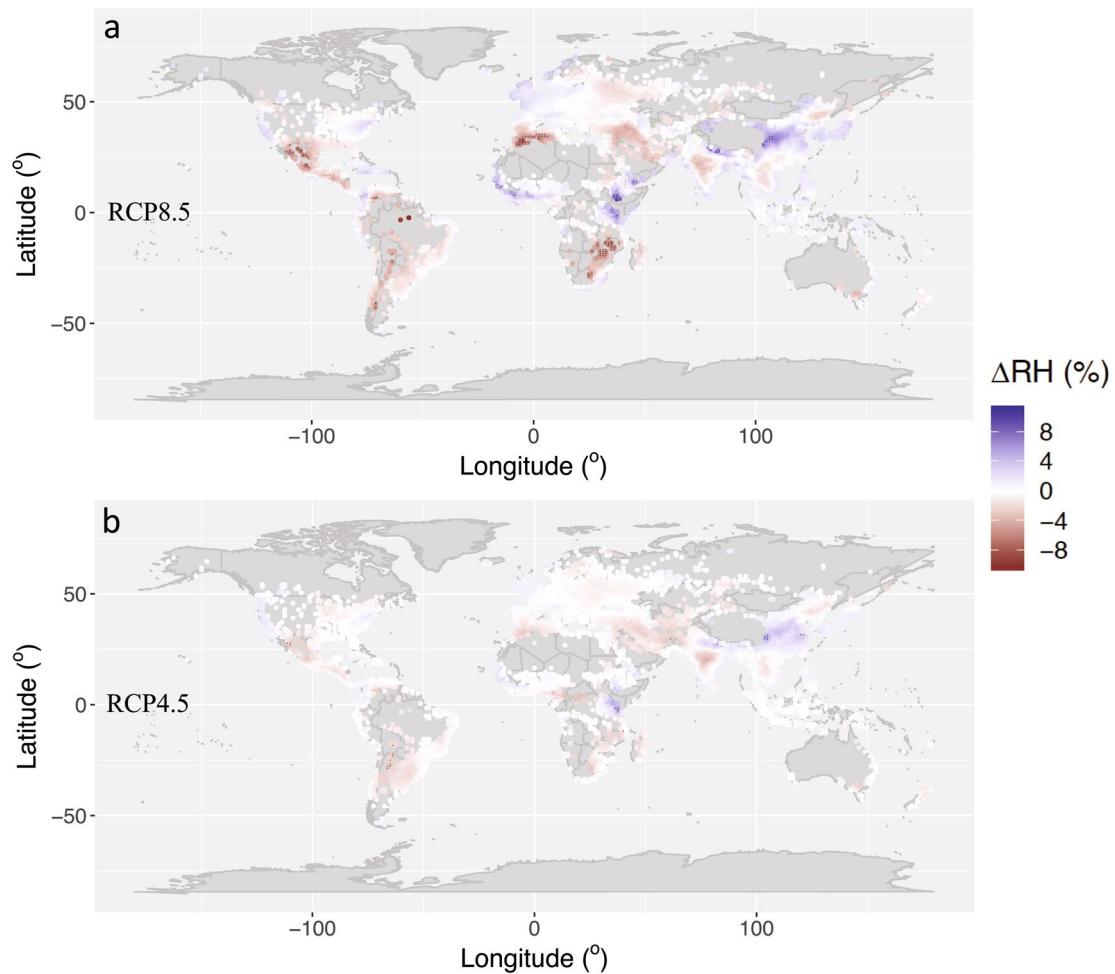
Extended Data Fig. 5 | Multi-model mean urban warming for season DJF (December - February) in 2091 - 2100 relative to 2006 - 2015. a-c: RCP 8.5; d-f: RCP 4.5. a and d: urban warming in diurnal average temperature T_a ; b and e: in diurnal maximum temperature T_{max} ; c and f: in diurnal minimum temperature T_{min} . Stippling indicates substantial change ($\Delta T \geq 4$ K under RCP 8.5 or $\Delta T \geq 1.5$ K under RCP 4.5) with high inter-model robustness (SNR > 2.5).



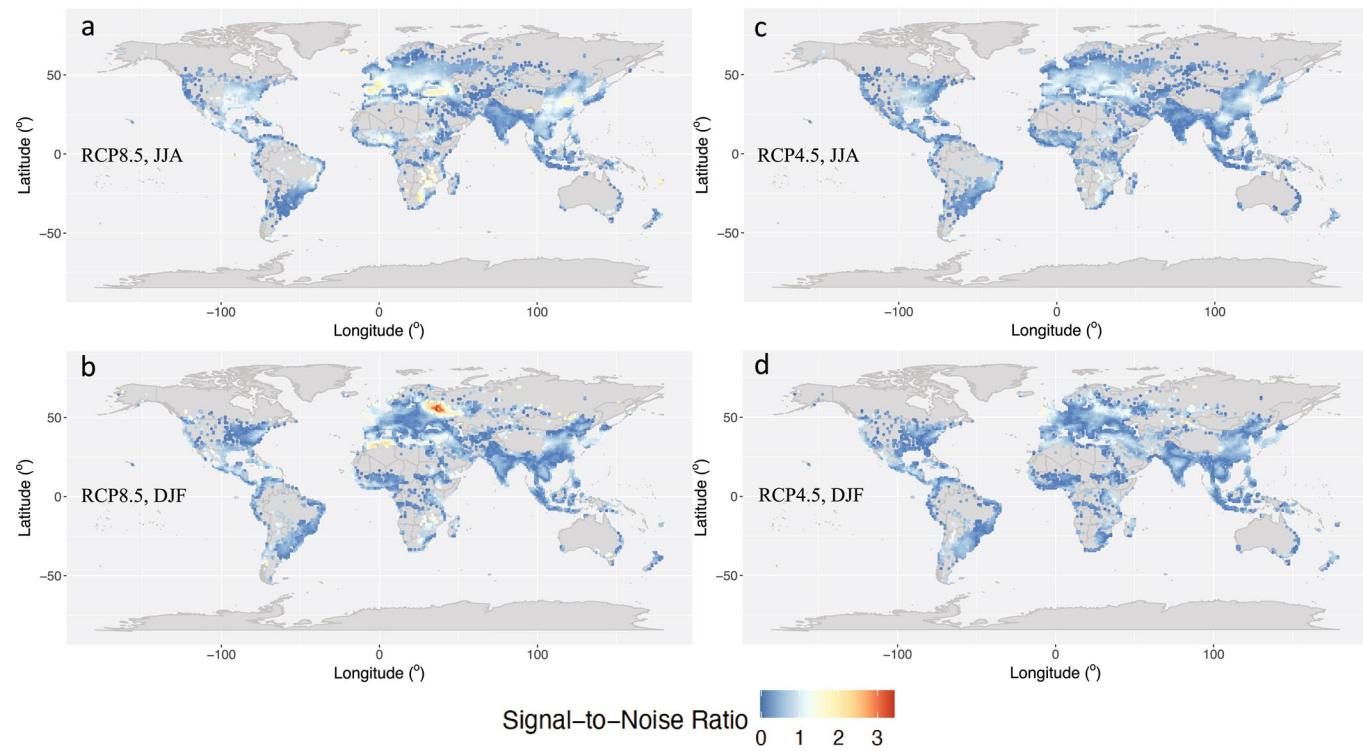
Extended Data Fig. 6 | Inter-model robustness of urban temperature projections measured by signal-to-noise ratio (SNR). **a-d**, RCP 8.5; **e-h**, RCP 4.5. **a, b, e and f**: season JJA (June – August); **c, d, g and h**: season DJF (December – February); **a, c, e and g**: projection in diurnal maximum temperature T_{max} ; **b, d, f and h**: projection in diurnal minimum temperature T_{min} .



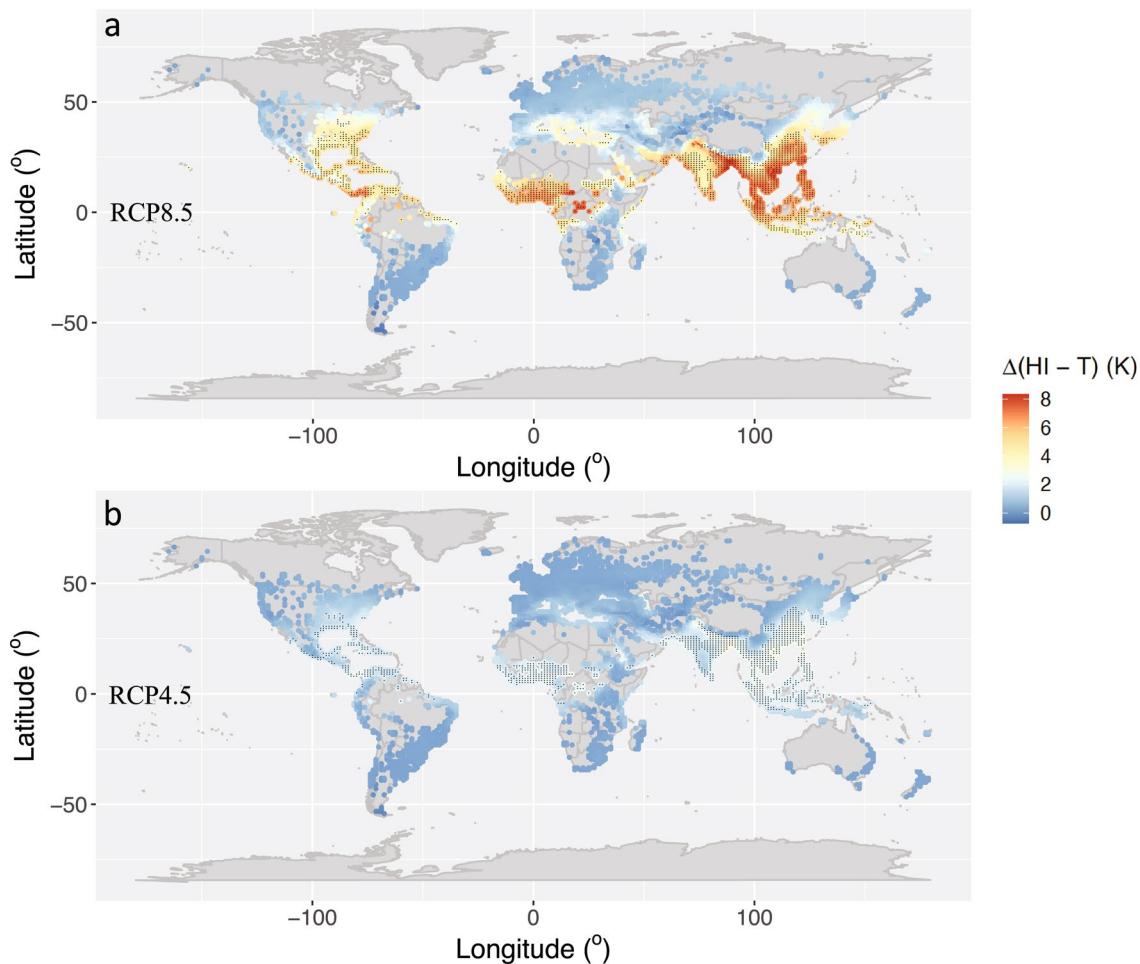
Extended Data Fig. 7 | Multi-model mean urban specific humidity (Q) change for season JJA (June - August) in 2091 - 2100 relative to 2006 - 2015.
a, RCP 8.5; **b**, RCP 4.5.



Extended Data Fig. 8 | Multi-model mean urban relative humidity (RH) change for season DJF (December - February) in 2091 - 2100 relative to 2006 - 2015. **a**, RCP 8.5; **b**, RCP 4.5.



Extended Data Fig. 9 | Inter-model robustness of urban RH projections measured by signal-to-noise ratio (SNR). **a,b:** RCP 8.5; **c,d:** RCP 4.5. **a** and **c**: season JJA (June – August); **b** and **d**: season DJF (December – February).



Extended Data Fig. 10 | Multi-model mean of urban $\Delta(HI - T_0)$ for season JJA (June - August) in 2091 - 2100 relative to 2006 - 2015. **a:** RCP 8.5; **b:** RCP 4.5. Stippling indicates substantial change ($\Delta(HI - T_0) > 3$ K under RCP8.5 or $\Delta(HI - T_0) > 1.5$ K under RCP4.5) with high inter-model robustness (SNR>2.5).