Supplementary Information for

Spatially concentrated heat emissions from data centers lead to localized urban warming and heat exposure

Hao Hu[1], Arman Shehabi[2], Cenlin He[3], Kelly Sanders[1], Jiachen Zhang[1,*]

[1] Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA 90089, USA

[2] Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720, USA

[3] NSF National Center for Atmospheric Research (NCAR), Boulder, CO 80301, USA

Corresponding E-mail: jiachen.zhang@usc.edu

Supplementary Texts

Text S1. WRF-SLUCM Details

Meteorological simulations are performed with version 4.6.1 of the Advanced Research WRF model. The modelling framework used two one-way nested domains defined on a Lambert conformal projection, as shown in **Supplementary Fig. S13**. The outer domain (d01) has a horizontal resolution of 12 km and covers the western United States providing boundary conditions for the inner domain. The inner domain (d02) covers the entire state of California at a horizontal resolution of 4 km and is used as the primary analysis domain for assessing the impacts of data-center anthropogenic heat (AH) emissions. The model integration time step is 24 s for the outer domain and a parent-to-child time step ratio of 3:1 to reduce simulation cost. The model uses 45 terrain-following Eta levels in the vertical direction, extending from the surface to 100 hPa. The initial conditions of d01 and d02, as well as the boundary condition of d01, are based on the National Centers for Environmental Prediction (NCEP) Global Data Assimilation System final analysis data at a 0.25° resolution and a temporal resolution of 6 hr^{1}. No nudging techniques are applied in our simulations.

Among all the tested multiple model configurations, we select a physics suite that provided the best baseline performance. The baseline model setup uses the following physics schemes: the Morrison 2-moment for microphysics scheme[2], the Kain–Fritsch cumulus parameterization scheme[3], the RRTMG Shortwave and Longwave for radiation scheme[4], the Asymmetric Convection Model 2 Scheme for planetary boundary layer[5], the Revised MM5 Scheme for surface layer physics scheme[6], and the Unified Noah land surface model for land surface model scheme[7]. The Kain–Fritsch cumulus scheme is activated for the d01 and disabled for the d02 to allow for explicit resolving of convection.

We use the global 100-m urban Local Climate Zone (LCZ) land use classification dataset[8,9]. Urban fraction data is from NLCD2011[10], which refers to the proportion of each grid cell covered by impervious urban surfaces (i.e., roofs, walls, paved roads). For the inner domain 02, we employ the single-layer urban canopy model (SLUCM) to represent land–atmosphere exchanges of water, momentum, and energy over impervious urban surfaces [11,12]. Data-center anthropogenic heat (AH) emissions are incorporated within this urban-canopy framework. We modify the SLUCM source code to ingest the gridded spatially-varying anthropogenic heat flux (AHF) dataset described above and add AHF directly to the canopy-air energy balance calculations. Previous studies have implemented AH in urban canopy models through several pathways[13]. Here, we adopt the approach corresponding to "method 3" in Li et al.[13], in which AH is introduced as a direct source term in the canyon-air temperature energy balance. The modified source code is available in our GitHub repository [14]. Urban morphological parameters (UMPs) are assigned using the default LCZ-based urban canopy

parameters [9]. This choice is supported by a previous Southern California WRF–SLUCM case study [15], which showed that this set of UMPs leads to reasonable model performance and exert a relatively limited influence on 2-m air temperature in the region. Under ±50% perturbations of UMPs, the spatial standard deviation of simulated 2-m air temperature responses remained below 0.2 K[15]. In contrast, these parameters primarily affected the near-surface wind field through changes in aerodynamic roughness, displacement height and momentum exchange.

Model output variables are saved at an hourly interval for analysis. To evaluate the capability of WRF-SLUCM in simulating local near-surface meteorological variables and to compare their performance under different configurations, we utilize 2-m air temperature, relative humidity, and wind speed data from the US EPA's Air Quality System (AQS) database[16]. As we focus on urban climate, we evaluate our model performance against AQS observation sites in urban areas using metrics including mean bias (MB), mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2). The model evaluation results are shown in **Supplementary Fig. S14**.

Supplementary Tables

Supplementary Table S1 | Reference temperature threshold for identifying populations in already hot baseline conditions. Minimum mortality temperatures (MMTs) for selected California locations, compiled from published epidemiological estimates of the daily mean temperature–mortality relationship[17]. Rather than applying these values as location-specific mortality thresholds, we used them to derive an approximate representative temperature threshold for California. This threshold was used only to identify populations already exposed to high baseline daily mean temperatures before assessing additional warming attributable to data-centre heat emissions. The MMT values are reported in degrees Celsius.

County	State	Minimum mortality temperature (°C)
USA Los Angeles	CA	22.8
USA Oakland	CA	20.8
USA Orange County	CA	25.6
USA Riverside-San Bernardino	CA	28.3
USA Sacramento	CA	25.3
USA Salinas	CA	20.6
USA San Diego	CA	23.3
USA San Francisco	CA	20.8
USA San Jose	CA	23.3
USA Ventura County	CA	21.7

Supplementary Table S2 | Crosswalk of EIA fuel-type categories. Fuel-type mapping is used to harmonize balancing-authority-level generation data from EIA-930 with plant-level generation and fuel-consumption records from EIA-923 for the calculation of indirect heat emissions from electricity generation. Aggregated EIA-930 fuel categories were matched to corresponding EIA-923 aggregated fuel-type codes before deriving fuel-specific waste-heat intensities and assigning electricity-generation heat emissions to data-center electricity demand. Only fuel categories used in this study are shown.

EIA 930 Fuel Type[a]	EIA 923 Fuel Type[b]
NG	NG, OOG, MLG, HYC
COL	COL, PC
OIL	DFO, WOO
NUC	NUC
OTH	WWW, ORW, HPS, OTH

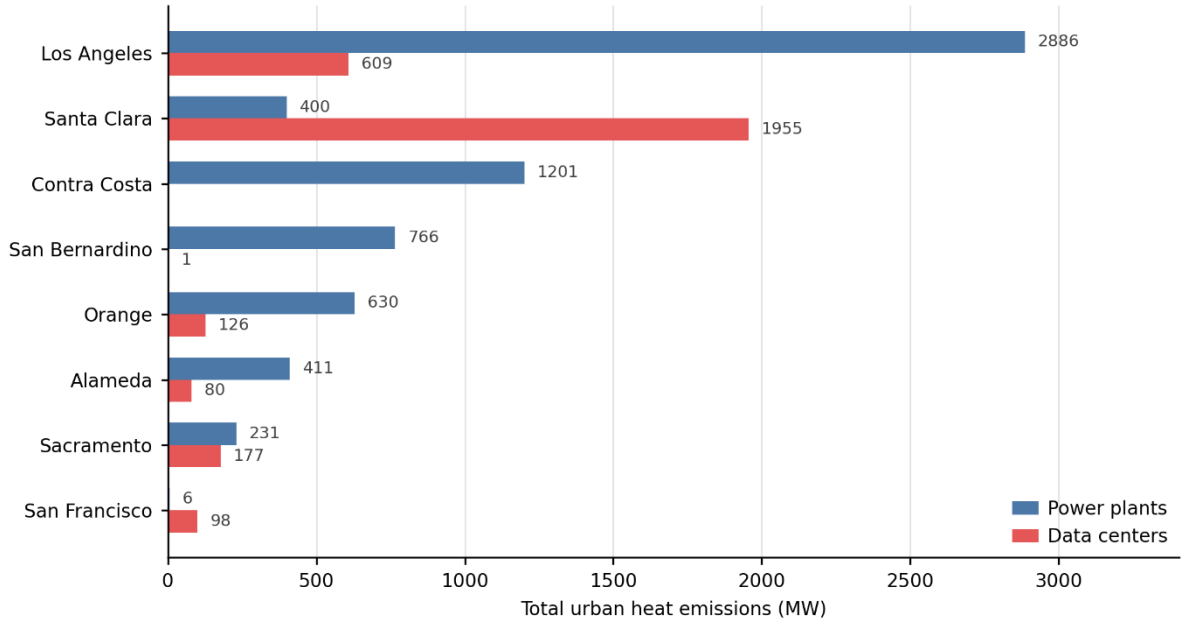
[a] NG, natural gas; COL, coal; OIL, oil/petroleum liquids; NUC, nuclear; OTH, other.

[b] NG, Natural Gas; OOG, Other Gases; MLG, Biogenic Municipal Solid Waste and Landfill Gas; HYC, Hydroelectric Conventional; COL, Coal; PC, Petroleum Coke; DFO, Distillate Petroleum; WOO, Waste Oil; NUC, Nuclear; WWW, Wood and Wood Waste; ORW, Other Renewables; HPS, Hydroelectric Pumped Storage; OTH, Other, including nonbiogenic municipal solid waste.

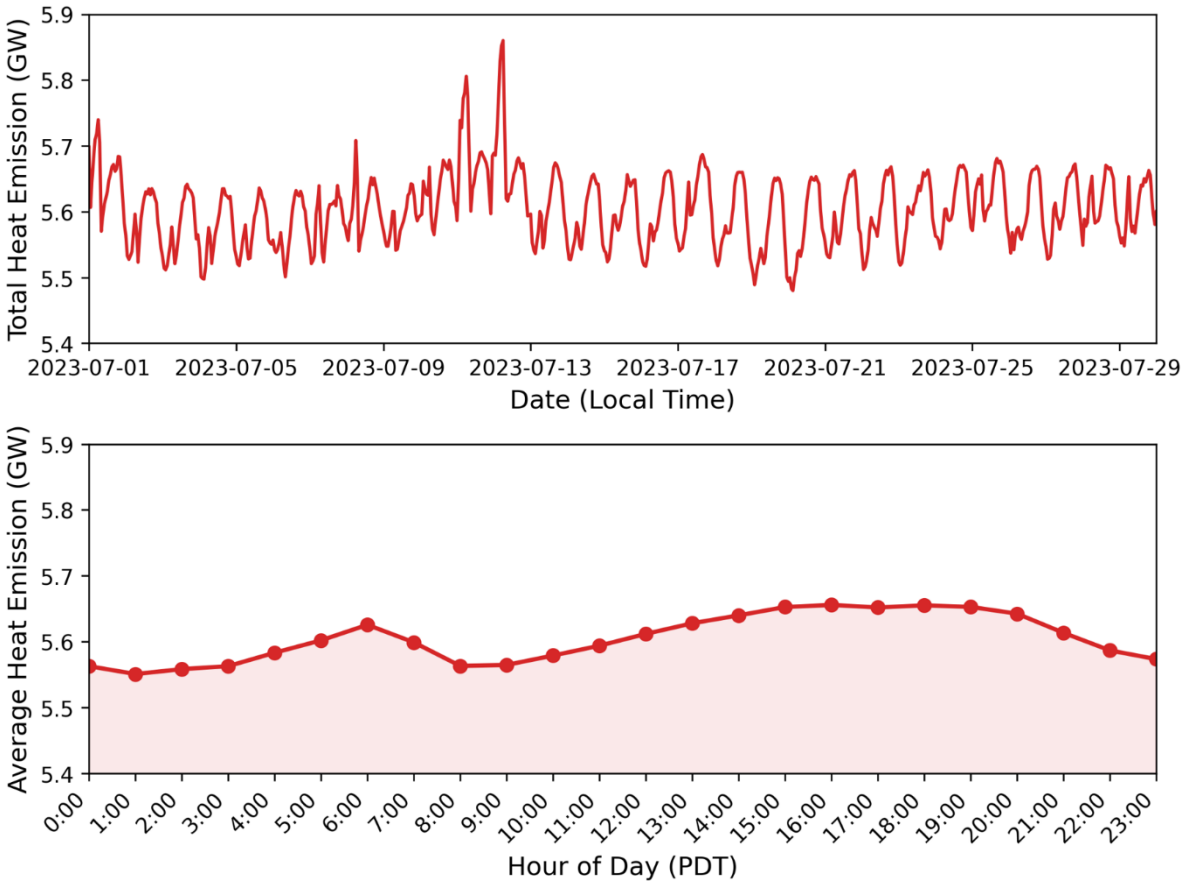
Supplementary Table S3 | Hourly scaling factors used to represent alternative temporal profiles of data-center direct AH emissions. Hourly scaling factors applied to distribute daily data-center heat emissions across a 24-h period. The flat profile assumes temporally constant heat emissions throughout the day, with a scaling factor of 1.00 for all hours. The hourly curve used for the diurnal profile was derived from a public building energy-use dataset [18,19]. Both profiles are normalized around the daily mean and were used to test the sensitivity of simulated warming responses to assumptions about the temporal distribution of heat emissions.

Hour	1	2	3	4	5	6	7	8	9	10	11	12
Flat	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Diurnal	0.82	0.78	0.76	0.72	0.70	0.70	0.76	0.88	0.92	0.98	1.03	1.10
Hour	13	14	15	16	17	18	19	20	21	22	23	24
Flat	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Diurnal	1.15	1.20	1.24	1.28	1.29	1.26	1.19	1.12	1.06	1.00	0.95	0.86

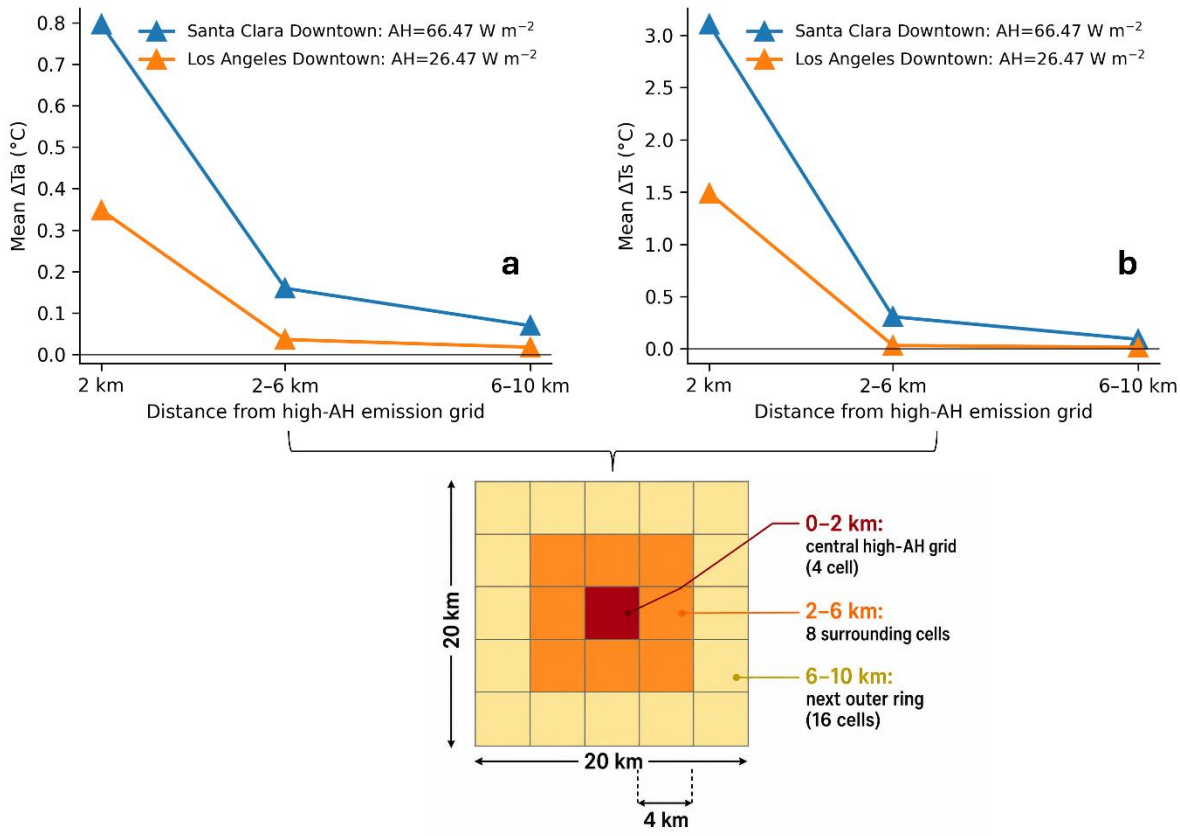
Supplementary Figures



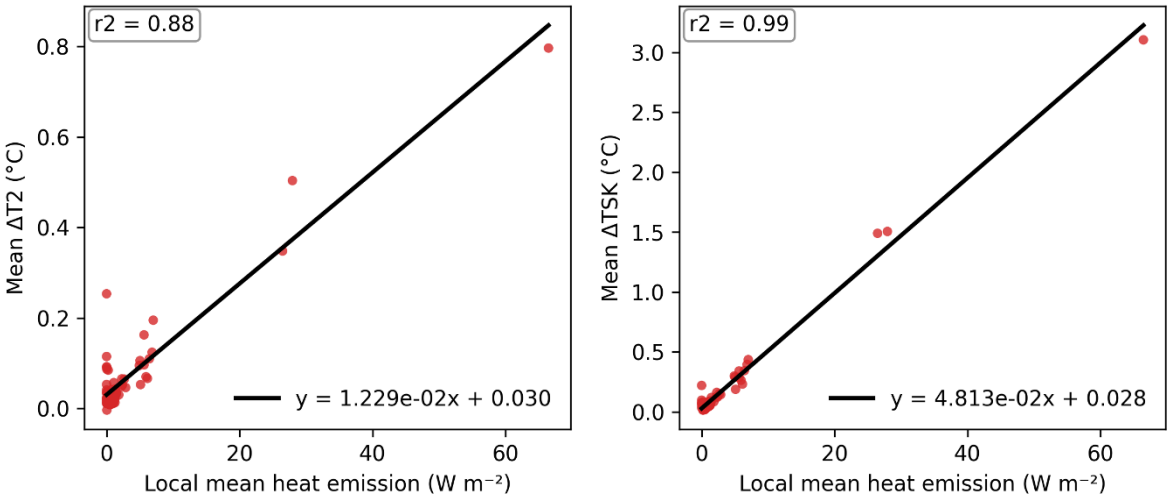
Supplementary Fig. S1 | Comparison of urban heat emissions from power plants and data centers across selected California counties. Heat emissions represent the sum of all facilities within each county, including either data centers or power plants. Only emissions located within urban areas are included, with urban areas identified using LCZ land-use classes[8,9].



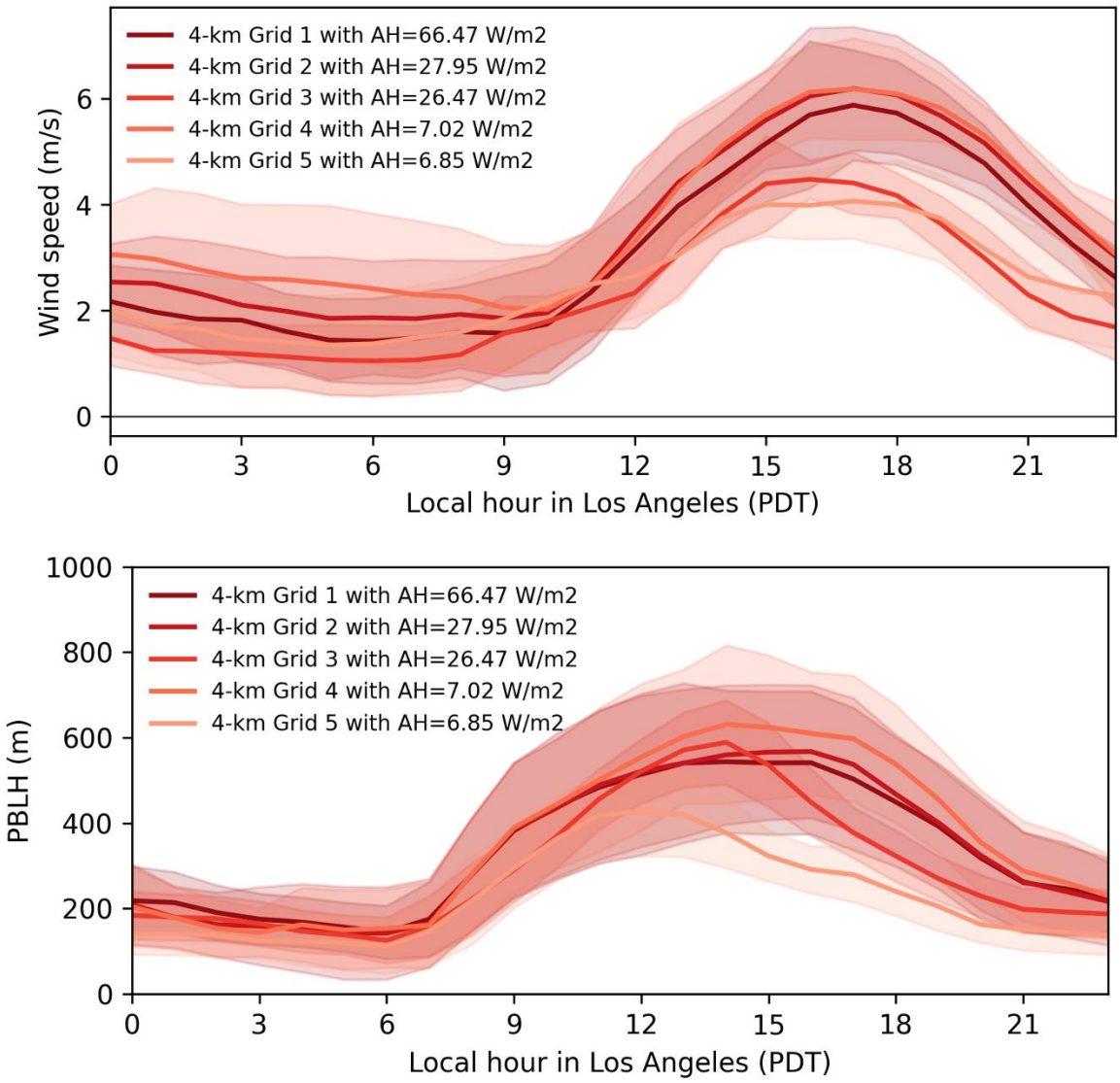
Supplementary Fig. S2 | Temporal persistence of data-center anthropogenic heat emissions. Hourly total anthropogenic heat emissions from California data centers during July 2023 show limited short-term variability, remaining close to 5.5–5.7 GW throughout the month. The composite diurnal cycle further indicates weak hour-to-hour variation, consistent with the continuous operation of data-center IT equipment, cooling systems and supporting infrastructure.



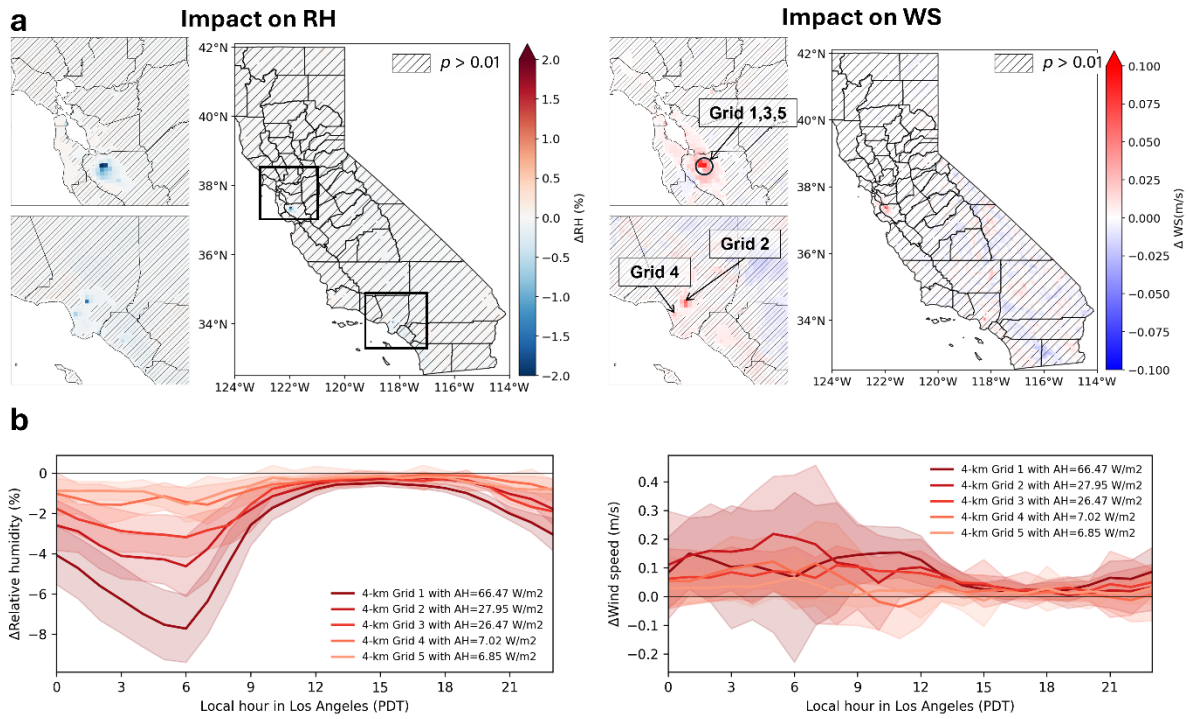
Supplementary Fig. S3 | Distance-dependent decay of data center-induced warming around high-emission grids in July 2023. Mean changes in 2-m air temperature (ΔT_a; a) and surface skin temperature (ΔT_s; b) are shown as a function of distance from the two highest AH emission grids. For each high-AH grid, the 0–2 km bin denotes the central 4-km grid cell itself, the 2–6 km bin denotes the eight immediately surrounding cells, and the 6–10 km bin denotes the next outer ring of sixteen cells, as illustrated schematically in b. Grid 1 is located in downtown Santa Clara, whereas Grid 2 is located in downtown Los Angeles. The locations of these two grids are shown in **Fig. 2a.**



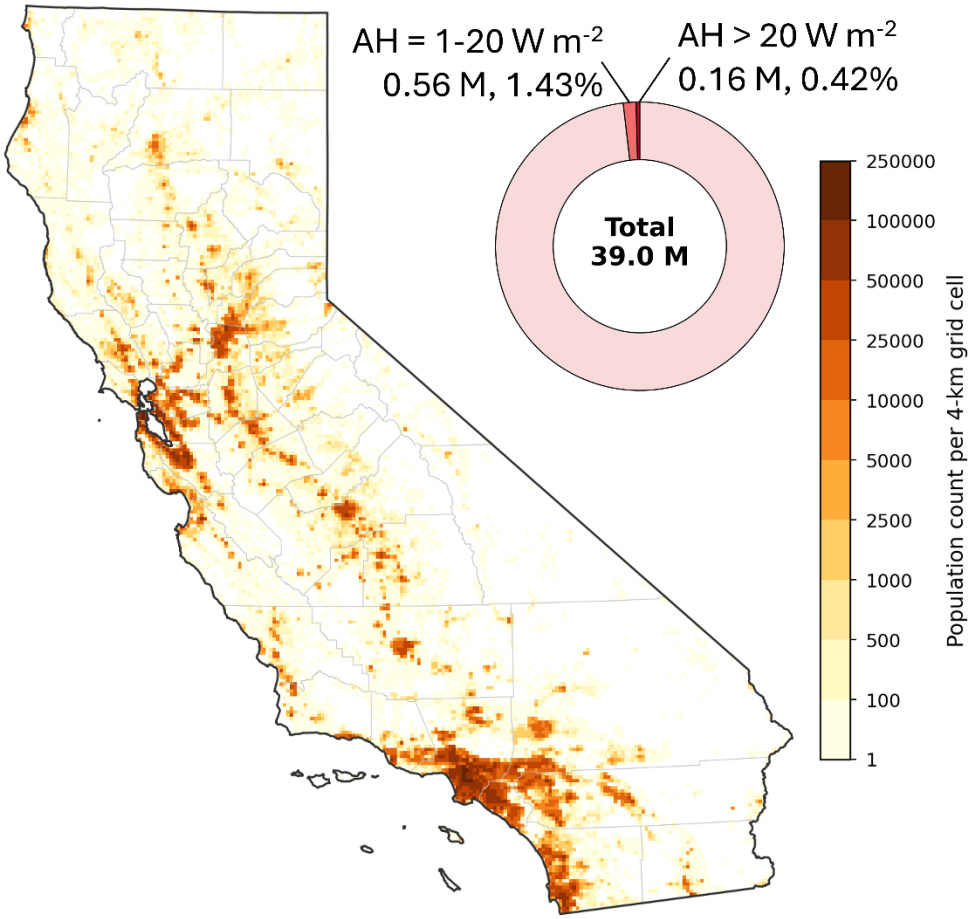
Supplementary Fig. S4 | Relationship between local anthropogenic heat flux and simulated warming. Mean changes in surface air temperature (ΔT_a; left) and surface skin temperature (ΔT_s; right) are shown as a function of local mean anthropogenic heat (AH) across data center-affected 4-km grid cells. Red points represent individual grid cells, and black lines show ordinary least-squares linear fits.



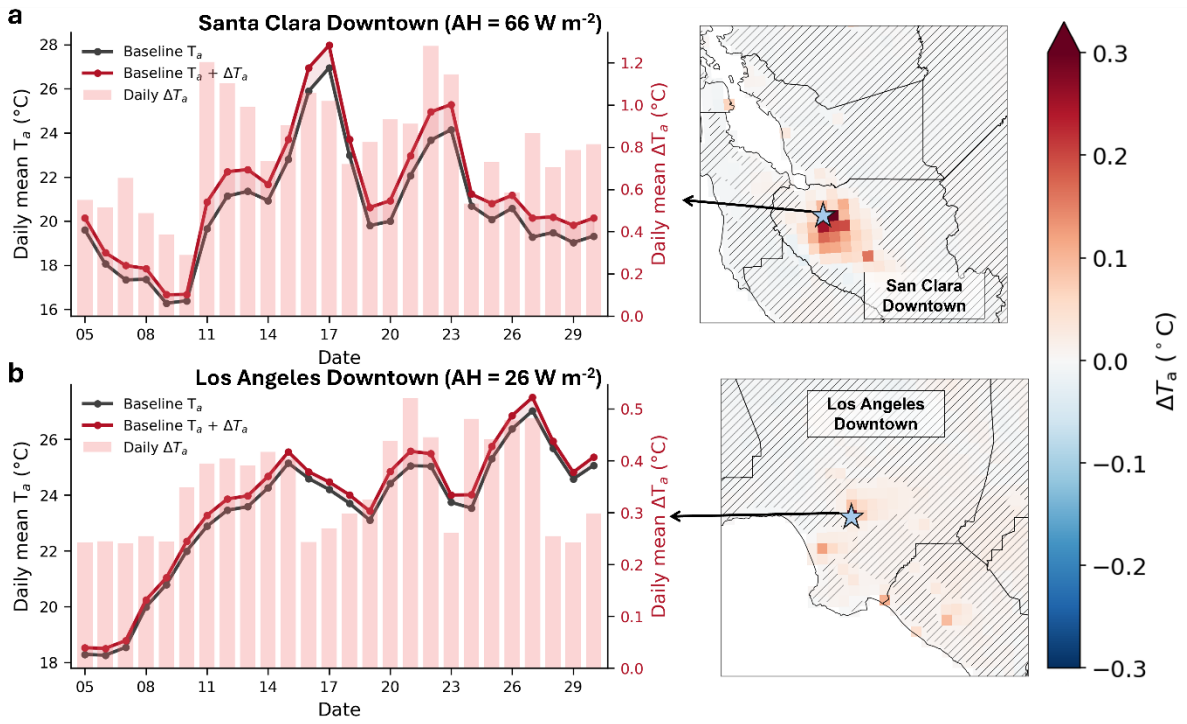
Supplementary Fig. S5 | Diurnal meteorological conditions over grids with high data center AH emissions. Mean diurnal cycles of wind speed (top) and planetary boundary-layer height (PBLH; bottom) are shown for the five 4-km grids with the highest data center AH emissions. Solid lines indicate the ensemble-mean hourly values, and shaded envelopes indicate variability across the analysis period. The locations of the five grids are shown in **Fig. 2**.



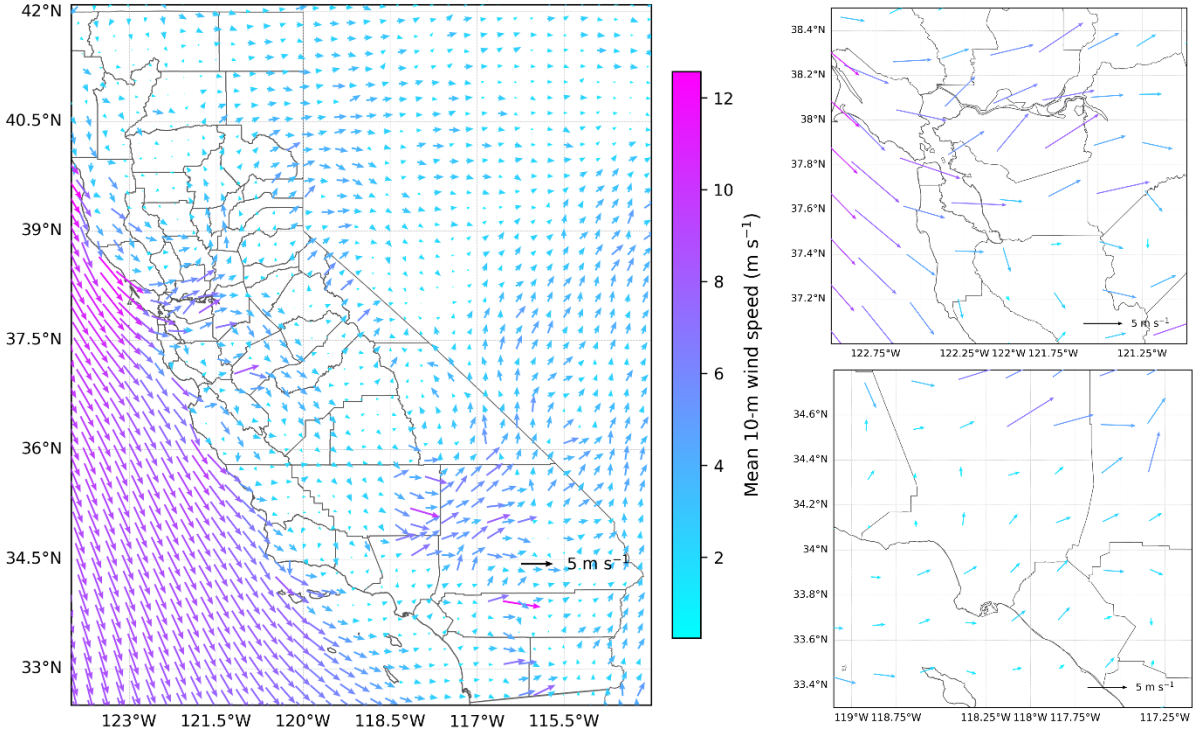
Supplementary Fig. S6 | Spatial distribution of meteorological responses and diurnal cycles over high-AH data-center grids. a Spatial pattern of changes in relative humidity (ΔRH) and wind speed (ΔWS) induced by data-center AH emissions. Hatched areas indicate grid cells where changes are not statistically significant ($p > 0.01$). **b** Panels show mean diurnal cycles of ΔRH and ΔWS for the five highest-AH grids. Solid lines represent mean hourly responses, and shaded envelopes indicate temporal variability. The locations of the five grids are shown in **Fig. 2**.



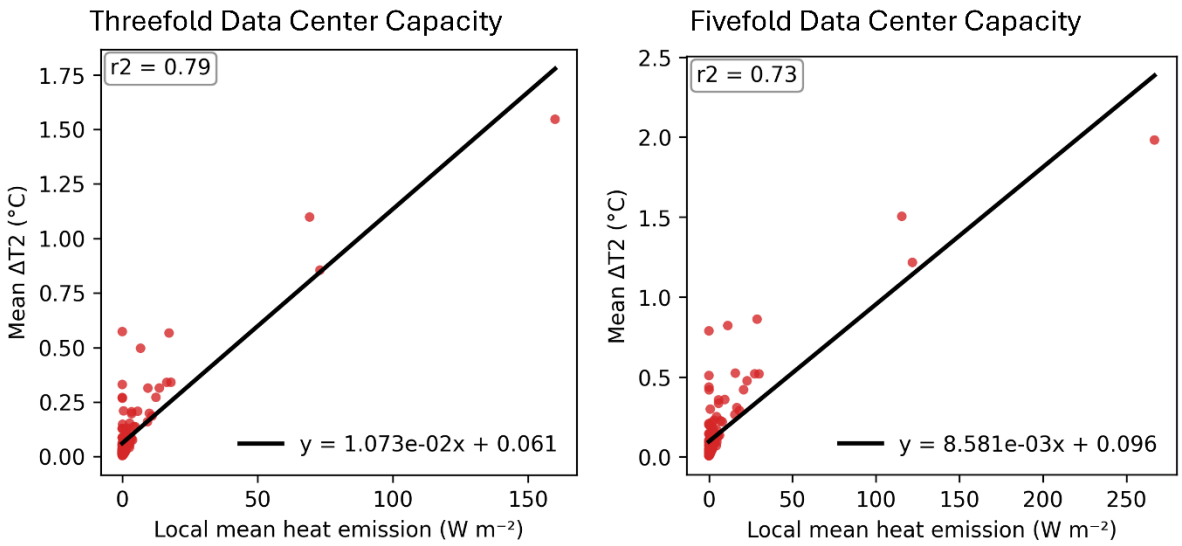
Supplementary Fig. S7 | Population distribution across California. Map shows the population count in each 4-km grid cell across California. The inset donut chart summarizes the statewide population living in grid cells with different data center AH levels. Of California's 39.0 million residents, approximately 0.56 million people (1.43%) live in grid cells with AH between 1 and 20 W m^{-2}, whereas 0.16 million people (0.42%) live in grid cells with AH exceeding 20 W m^{-2}. Population is resampled from Depsky et al. (2022)[20].



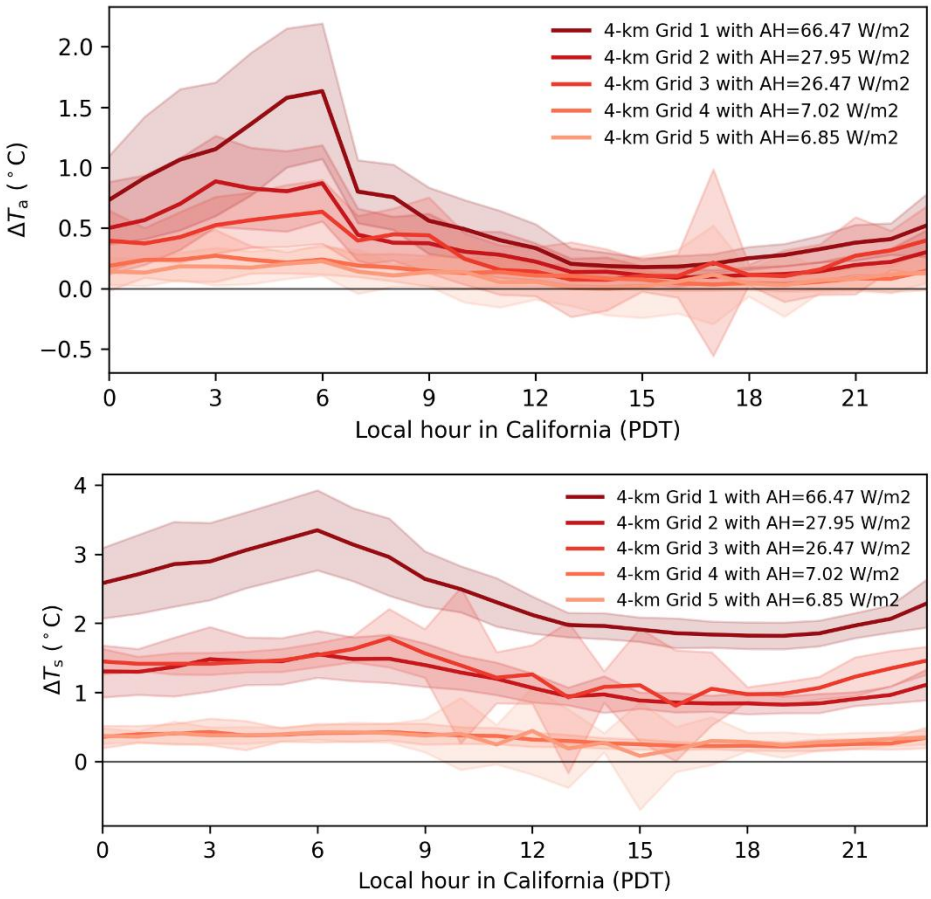
Supplementary Fig. S8 | Daily evolution of data center-induced near-surface warming in downtown Santa Clara and downtown Los Angeles. Left panels show daily mean near-surface air temperature (T_a) under the baseline simulation and the data center AH scenario, with bars indicating the daily mean warming response (ΔT_a). Right panels show the corresponding July-mean spatial pattern of (ΔT_a), with stars marking the selected grid cells used for the daily time-series analysis. Hatched areas denote grid cells where the simulated response is not statistically significant at the ($p \leq 0.01$) level.



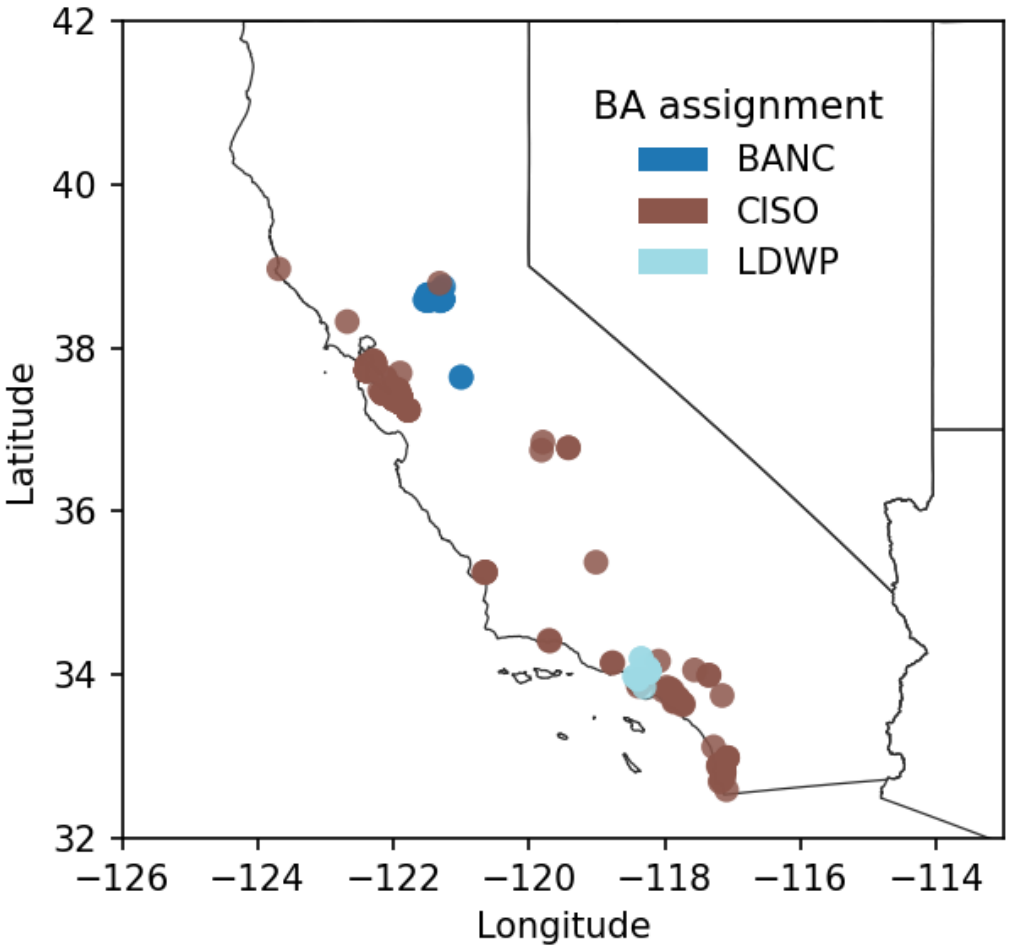
Supplementary Fig. S9 | Mean near-surface wind fields over California and selected data-center hotspot regions in July 2023. Spatial distribution of simulated mean 10-m wind speed and direction over California, with arrows indicating wind direction and arrow color denoting wind speed. The left panel shows the statewide wind field, while the right panels provide zoomed views of the Santa Clara region and the Los Angeles basin, where major data-center AH emission clusters are located. Reference vectors indicate a wind speed of $5\ \mathrm{m\ s^{-1}}$.



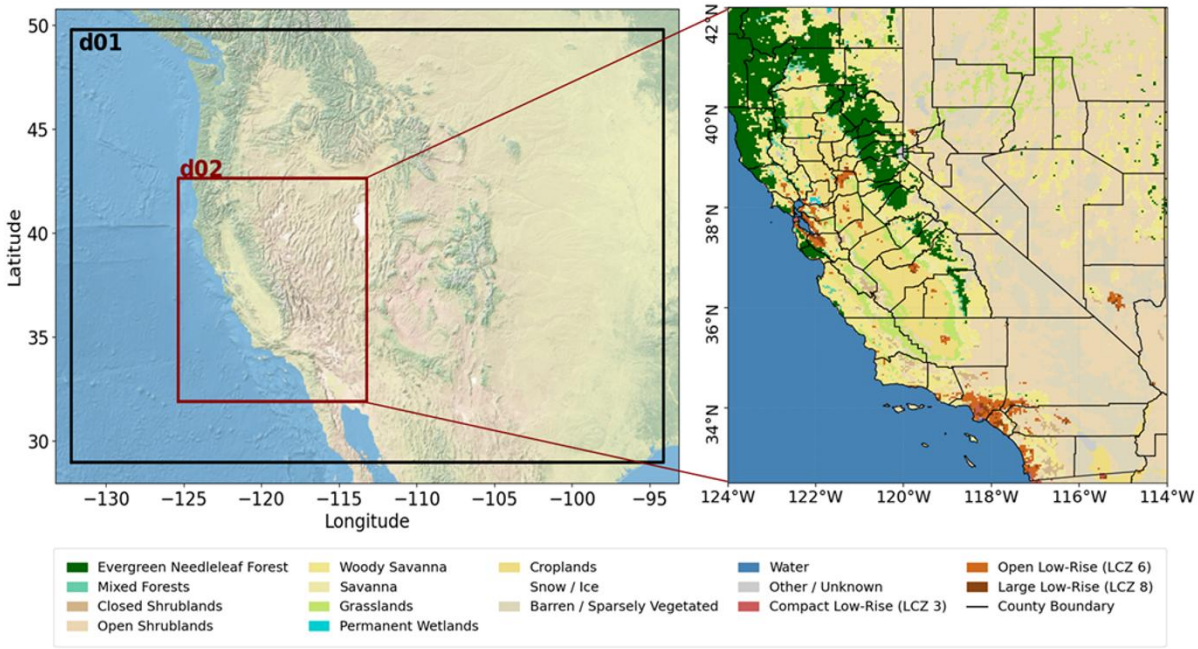
Supplementary Fig. S10 | Scaling relationship between AH emissions and 2-m air-temperature response. Monthly mean ΔT_a is plotted against monthly mean anthropogenic heat emission intensity (AH) across 4-km grid cells under the threefold-capacity scenario (left) and fivefold-capacity scenario (right). Red points denote individual grid cells, and black lines show ordinary least-squares linear fits.



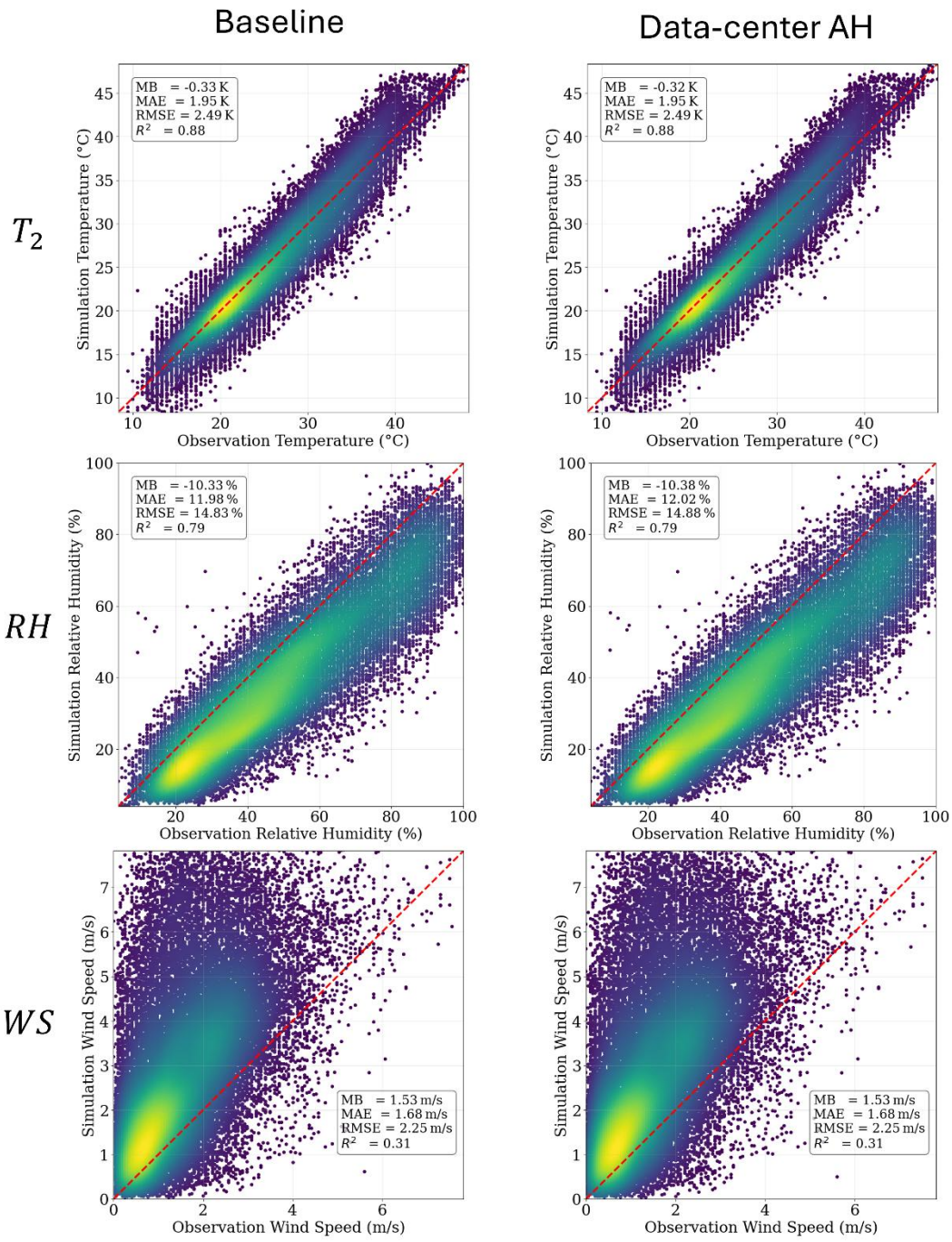
Supplementary Fig. S11 | Diurnal cycles of data-centre-induced warming when anthropogenic heat emissions follow a conventional office-building diurnal load profile. Diurnal curves show hourly mean warming at five selected 4-km high-AH grid cells, with shaded envelopes indicating inter-day variability. The resulting warming magnitudes and diurnal-response patterns remain similar to those in the main simulations, supporting the robustness of the findings to plausible temporal-load assumptions.



Supplementary Fig. S12 | Inferred balancing-authority assignment for California data centers. Data-center facilities are mapped and colored by their inferred balancing authority (BA) region, including BANC, CISO and LDWP. Because facility-level BA identifiers are not available for all data centers, BA membership was approximated using a spatial majority-vote method based on nearby electricity-generation infrastructure.



Supplementary Fig. S13 | WRF model domains and land-use representation for the California simulations. Nested WRF simulation domains used in this study, with the outer domain (d01) covering the western United States with 12 km resolution and the inner high-resolution domain (d02) centred on California with 4 km resolution. The right panel shows the land-use and land-cover classification within d02, including natural vegetation, croplands, water bodies and urban land-use categories represented in the single-layer urban canopy model. County boundaries are overlaid in grey to provide geographic reference. The inner domain is used for analysing data-center heat emissions and their effects on near-surface meteorology across California.



Supplementary Fig. S14 | Model evaluation of near-surface meteorological variables. Scatter-density comparisons between observations and WRF simulations for 2-m air temperature (T_2), 2-m relative humidity (RH) and 10-m wind speed (WS). The left and right columns show the baseline and data-center anthropogenic heat simulations, respectively. Red dashed lines denote the 1:1 relationship; mean bias (MB), mean absolute error (MAE), root-mean-square error (RMSE) and R^2 are shown in each panel. The positive wind-speed bias is partly expected because WRF 10-m winds represent winds above the roughness length and zero-plane displacement height and may also reflect overly strong boundary-layer mixing and underestimated effective urban roughness in urban areas[15,21–23].

References

1. Commerce, N. C. for E. P. W. S. S. D. of. NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids. UCAR/NCAR - Research Data Archive https://doi.org/10.5065/D65Q4T4Z (2015).

2. Morrison, H., Thompson, G. & Tatarskii, V. Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes. https://doi.org/10.1175/2008MWR2556.1 (2009) doi:10.1175/2008MWR2556.1.

3. Berg, L. K., Gustafson, W. I., Kassianov, E. I. & Deng, L. Evaluation of a Modified Scheme for Shallow Convection: Implementation of CuP and Case Studies. https://doi.org/10.1175/MWR-D-12-00136.1 (2013) doi:10.1175/MWR-D-12-00136.1.

4. Iacono, M. J. *et al.* Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres* **113**, (2008).

5. Pleim, J. E. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *Journal of Applied Meteorology and Climatology* **46**, 1383–1395 (2007).

6. Jiménez, P. A. *et al.* A Revised Scheme for the WRF Surface Layer Formulation. *Monthly Weather Review* **140**, 898–918 (2012).

7. Tewari, M. Implementation and verification of the unified Noah land surface model in the WRF model. (2004).

8. Demuzere, M., He, C., Martilli, A. & Zonato, A. *Technical Documentation for the Hybrid 100-m Global Land Cover Dataset with Local Climate Zones for WRF*. (2023). doi:10.5281/zenodo.7670791.

9. Stewart, I. D. & Oke, T. R. Local Climate Zones for Urban Temperature Studies. https://doi.org/10.1175/BAMS-D-11-00019.1 (2012) doi:10.1175/BAMS-D-11-00019.1.

10. Homer, C. *et al.* Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a Decade of Land Cover Change Information. *Photogrammetric Engineering & Remote Sensing* **81**, 345–354 (2015).

11. 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 Meteorology* **101**, 329–358 (2001).

12. Chen, F. *et al.* The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *International Journal of Climatology* **31**, 273–288 (2011).

13. Li, D. *et al.* Structural Uncertainty in the Sensitivity of Urban Temperatures to Anthropogenic Heat Flux. *J Adv Model Earth Syst* **16**, e2024MS004431 (2024).

14. Hu, H. ACE-lab-USC/WRF-SLUCM-Modified. (2026).

15. Hu, H. *et al.* Quantifying Urban Morphology-Induced Uncertainty in Urban Meteorology and Heat Stress Simulations in Southern California. *Journal of Geophysical Research: Atmospheres* **131**, e2025JD045318 (2026).

16. US EPA. Air Quality System Data Mart. (2012).

17. Gasparrini, A. *et al.* Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet* **386**, 369–375 (2015).

18. Melo, F. C., Carrilho da Graça, G. & Oliveira Panão, M. J. N. A review of annual, monthly, and hourly electricity use in buildings. *Energy and Buildings* **293**, 113201 (2023).

19. Miller, C. *et al.* The ASHRAE Great Energy Predictor III competition: Overview and results. *Science and Technology for the Built Environment* **26**, 1427–1447 (2020).

20. Depsky, N. J., Cushing, L. & Morello-Frosch, R. High-resolution gridded estimates of population sociodemographics from the 2020 census in California. *PLoS One* **17**, e0270746 (2022).

21. He, X. *et al.* High-resolution dataset of urban canopy parameters for Beijing and its application to the integrated WRF/Urban modelling system. *Journal of Cleaner Production* **208**, 373–383 (2019).

22. Chen, D. *et al.* Sensitivity analysis of planetary boundary layer parameterization on meteorological simulations in the Yangtze river delta region, China. https://doi.org/10.1039/D4EA00038B (2024) doi:10.1039/D4EA00038B.

23. Cheng, F.-Y. *et al.* Impact of Effective Roughness Length on Mesoscale Meteorological Simulations over Heterogeneous Land Surfaces in Taiwan. *Atmosphere* **10**, 805 (2019).