Hyper-local temperature prediction using detailed urban climate informatics

Peiyuan Li[1] and Ashish Sharma[1]

[1]University of Illinois System

April 16, 2024

Abstract

** The latest version of this paper has been published in the *Journal of Advances in Modeling Earth Systems* (JAMES) from AGU. Please refer to the latest version on JAMES and cite as:

Li, P., & Sharma, A. (2024). Hyper-local temperature prediction using detailed urban climate informatics. *Journal of Advances in Modeling Earth Systems*, 16, e2023MS003943. https://doi.org/10.1029/2023MS003943

Modeling urban microclimate accurately is challenging due to the high surface heterogeneity of urban land cover and the vertical structure of street morphology. Recent years have witnessed significant efforts in numerical modeling and data collection of the urban environment. Nonetheless, it is difficult for the physical-based models to fully utilize the high-resolution data under the constraints of computing resources. The advancement in machine learning techniques offers the computational strength to handle the massive volume of data. In this study, we proposed a machine learning approach to estimate point-scale street-level air temperature from the urban-resolving mesoscale climate model and a suite of hyper-resolution urban informatics, including three-dimensional urban morphology, parcel-level land use inventory, and a dense weather observation network. We implemented this approach in the City of Chicago as a case study. The proposed approach vastly improves the resolution of temperature predictions in cities, which will help the city with walkability, drivability, and heat-related behavioral studies. Moreover, we tested the model's reliability on out-of-sample locations to investigate the application potentials to the other areas. This study also aims to gain insights into next-gen urban climate modeling and guide city observation efforts to build the strength for the holistic understanding of urban microclimate dynamics.

Hyper-local temperature prediction using detailed urban climate informatics

Peiyuan Li[1] and Ashish Sharma[1,2,3]

[1] *Discovery Partners Institute, University of Illinois System, Chicago, IL 60606 USA*

[2] *Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Champaign, IL 61820 USA*

[3] *Environmental Science Division, Argonne National Laboratory, Lemont, IL 60439 USA*

Abstract

Modeling urban microclimate accurately is challenging due to the high surface heterogeneity of urban land cover and the vertical structure of street morphology. Recent years have witnessed significant efforts in numerical modeling and data collection of the urban environment. Nonetheless, it is difficult for the physical-based models to fully utilize the high-resolution data under the constraints of computing resources. The advancement in machine learning techniques offers the computational strength to handle the massive volume of data. In this study, we proposed a machine learning approach to estimate point-scale street-level air temperature from the urban-resolving mesoscale climate model and a suite of hyper-resolution urban informatics, including three-dimensional urban morphology, parcel-level land use inventory, and a dense weather observation network. We implemented this approach in the City of Chicago as a case study. The proposed approach vastly improves the resolution of temperature predictions in cities, which will help the city with walkability, drivability, and heat-related behavioral studies. Moreover, we tested the model's reliability on out-of-sample locations to investigate the application potentials to the other areas. This study also aims to gain insights into next-gen urban climate modeling and guide city observation efforts to build the strength for the holistic understanding of urban microclimate dynamics.

Keyword

Urban informatics, hyper resolution, Weather Research and Forecast model, Gaussian Process Regression, Street-level temperature

Plain Language Summary

Estimating air temperature at street-level is difficult because of the complex environment in cities and the limitations of the current urban numerical models. In recent years, with the rapid development of data collection and analysis techniques, it is possible to fully utilize the hyper-local data harvested from urban areas by advanced machine learning algorithms. This study presents a modeling method to estimate point-scale street-level air temperature from a conventional urban weather model and a suite of hyper-resolution urban informatics. These datasets were collected using state-of-art techniques, such as sub-meter level Light Detection and Ranging technology and wireless weather observation network. Using the model, we estimated the street-level temperature over the City of Chicago. The modeling results have multiple real-world applications, such as providing navigation suggestions to reduce the thermal discomfort of pedestrians as an example. Moreover, given the current data availability, it is possible to expand the use of our model to other areas. The results of this study can also help the development of the next-generation urban climate and weather models and guide the observation efforts in cities. These together can build the strength for the holistic understanding of urban microclimate dynamics.

Key points:

- The study presents a modeling framework to estimate street-level air temperature using a suite of detailed urban climate informatics.
- Model results showed hyper-local urban features have significant impacts on street-level temperature but with a limited influence radius.
- The investigations on model sensitivity imply the existence of the optimum scale in urban modeling and critical locations in observation.

1 Introduction

Cities will be homes for over two-thirds of the global population by 2050 (UN-Habitat, 2019). This rapid urbanization will escalate the vulnerability of urban residents under various environmental stressors, such as urban heat, hazardous air quality, and extreme weather conditions (Revi et al., 2014). To make matters worse, global climate change tends to amplify the frequency of weather anomalies (Perkins-Kirkpatrick and Lewis, 2020) and induce additional uncertainties to urban environmental issues (Chen and Zhai, 2017; Huang et al., 2019; Kumar, 2021). Mitigating the adverse consequences and increasing preparedness for the future climate have become the most urgent tasks for the development of modern cities. In response, cities are deploying cyberinfrastructures to collect real-time environmental data. This is aiding in establishing a comprehensive set of urban climate informatics, such as dense weather observation networks, high-resolution urban land use and morphological database, traffic, and energy monitoring systems, etc. Such data can subsequently fulfill the increasing demand for accurate and quick predictions of environmental flows in cities and help make communities resilient to climate change (González et al., 2021; Middel et al., 2022).

Modeling the urban environment is challenging primarily due to the highly heterogeneous land use, complex urban fabric, and diverse anthropogenic activities (Oke, 1988; Oke et al., 2017). The past decades have witnessed continuous efforts to improve the representation of the complex urban environment with state-of-the-art urban models across scales. Most urban models can be classified into two categories depending on whether they are coupled with large-scale atmospheric models. The uncoupled models consider the detailed urban surface characteristics but do not dynamically interact with the atmosphere above the urban layer, such as the Town Energy Balance (TEB) model (Masson, 2000), Surface Urban Energy and Water balance Scheme (SUEWS, Järvi et al., 2011), and the Arizona State Single Layer Urban Model (ASLUM, Li and Wang, 2020; Wang et al., 2021). They use the "calculation unit" concept with one unit of the model representing a city with specific characteristics. Since uncoupled models are usually applied to small-scale problems (i.e., neighborhood-to-city scales), the models are often driven by in-situ measurements. The uncoupled models are developed from individual research groups, therefore, have faster iterations regarding the model functionality and have detailed parameterizations of the urban fabric, such as the inclusion of urban trees (Ryu et al., 2015), hydrological and ecological features (Yang and Wang, 2014; Meili et al., 2020), biogenic carbon exchange (Goret et al., 2019; Li and Wang, 2020), and anthropogenic emissions (Järvi et al., 2019).

The coupled model, by definition, considers the two-way interactions between the atmospheric layer and the land surface, therefore, can be applied to larger spatial scales. This modeling approach requires an atmospheric model to provide the meteorological forcings and an urban land surface model to resolve the urban dynamics. The most widely adopted example of coupled modeling framework is the mesoscale Weather Research and Forecast model combined with urban physics (uWRF, Chen et al., 2011). When incorporating advanced parameterization processes such as Local Climate Zones (LCZ, Stewart and Oke, 2012) and the distributed urban canyon parameterization approach (UCP), uWRF can theoretically resolve the urban dynamics up to a few hundred meters. However, compared to the uncoupled models, the urban features used in uWRF are much less sophisticated, with some critical components such as street trees and detailed hydrological processes missing from the street canyon. It is also extremely difficult to find the finest cell size that can reflect the hyper-local urban characteristics without breaking the physical schemes in the model simulation. Even with the most advanced urban parameterization approaches (Shen et al., 2019; Chen et al., 2022), modeling microclimate at the meter level is beyond the resolving power of the current generation of uWRF.

The urban computational fluid dynamics (CFD) models (e.g., Toparlar et al., 2017), on the other hand, consider the detailed urban morphology and simulate the turbulence field at an extremely high resolution (~10m) by solving the simplified Reynolds-averaged Navier-Stokes (RANS) equations or the most computational-expensive Large Eddy Simulation (LES, e.g., Maronga et al., 2020; Suter et al., 2022) and Detached Eddy Simulation (DES) models (Dadioti and Rees, 2017). Though urban CFD models can provide the most detailed flow field to street level, their applications are limited in spatiotemporal coverages due to the restrictions on computational resources. They are not designed for prediction purpose, as a city-scale CFD model usually run slower than real-time.

It is notable that the granularity of urban informatics is one to two orders of magnitude finer than the resolution of the state-of-art predictive urban models. This emerging trend demands new approaches, as an addition to the conventional physical-based numerical methods, to fully utilize the existing urban informatics and effectively harvest the pioneering observation efforts (Middel et al., 2022). Recent studies have explored the prospects of estimating hyper-local air temperature (T_a) via data-driven and statistical approaches. For example, Chen et al. (2019) used multi-variable linear regression to predict T_a based on temperature observation network and land cover information. While Venter et al. (2020) and Zumwald et al. (2021) adopted machine learning (ML) algorithms to hindcast T_a based on the predictors derived from high-resolution remote sensing (RS) imagery, crowd-sourced weather data, and Light Detection and Ranging (LiDAR) measurements. Similarly, Yin et al. (2020) estimated T_a from RS and LiDAR with an ML model trained on data collected from vehicle-borne sensors. Results from these data-driven studies achieved a resolution between 10m to 30m depending on the granularity of RS imageries (e.g., 30m from Landsat 8 or 10m from Sentinel-2). As RS data serves as the source of time-variant predictors, the T_a estimations can be only derived for the past when RS imageries are available, thus, are not prescient. Moreover, without being bound by the physical dynamics, the estimations based on statistical relationships may not explain the spatiotemporal variability of T_a sufficiently, leading to a systematic bias between the estimation and the measurement (Zumwald et al., 2021).

In this study, we propose an innovative approach to estimating street-level T_a from the uWRF model outputs and a group of high-resolution urban informatics over the City of Chicago, including a dense observation network, high-resolution LiDAR point clouds, and a parcel-level land use inventory (Fig. 1). Specifically, we use Gaussian Process Regression (GPR) to identify the relationship between the T_a estimated by uWRF and measured by ground sensors while considering the hyper-local impacts from urban land use types and morphology. The modeling approach offers point-scale air temperature estimation at resampled street locations in Chicago, which can be further integrated into desired resolution. We believe that the approach presented in this study can contribute to the knowledge of important urban problems, such as street walkability (O'Brien et al., 2019), disproportional heat exposure (Chakraborty et al., 2019), heat-related health issues (Heaviside et al., 2017), and behavioral studies (Anderson, 1989; Reeping and Hemenway, 2020).

The rest of the paper is organized as follows. Section 2 describes the details of urban informatics used in this study, as well as the configuration of the uWRF simulation and GPR models. Section 3.1 summarizes the urban features extracted from the urban informatics, followed by the results of street-level temperature estimation in temporal (Section 3.2) and spatial extent (Section 3.3), respectively. In Section 4, we discuss the model uncertainty, implications, and the limitations of this study. The key concluding remarks drawn from the results and discussion are presented in Section 5.

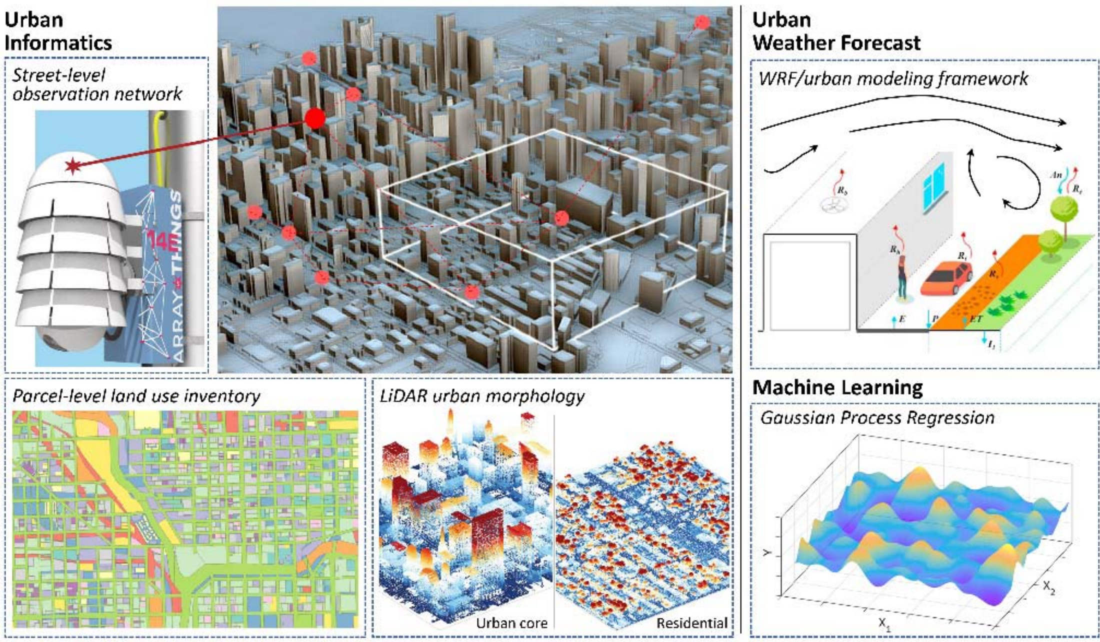

Figure 1. Concept diagrams of the three components: urban informatics, physical-based urban weather prediction model, and machine learning algorithms used in this study.

2 Data and Method

2.1 Study area

The City of Chicago, located by the shore of Lake Michigan (41.88^{o}N, -87.62^{o}W) in Illinois, is the urban core of the third-most populous metropolitan region in the United States (US Census 2020). It has a highly developed downtown area with 125 skyscrapers over 500 ft (152 m) and a radial urban-suburban gradient extending west from the lakeshore. Many studies have investigated Chicago's urban environment in terms of morphology (Patel et al., 2023), extreme heat (Sharma et al., 2016; Sharma et al., 2017), precipitation (Vavrus and Van Dorn, 2010), anthropogenic emissions (Conry et al., 2015), etc. The city authority also led efforts to create resilient and sustainable communities across the region by setting ambitious climate action plans since 2008. Moreover, as part of a US Department of Energy-supported project called Community Research on Climate and Urban Science (CROCUS), several Chicago-area institutions have joined forces to study urban climates and develop community-based solutions.

All the previous and ongoing endeavors have made the Chicago region an ideal testbed for urban climate studies.

2.2 Urban Informatics

2.2.1 Street-level observation network

The Array of Things (AoT) project started in 2018 and was designed to monitor the urban environment of Chicago via a dense observational network. The measurement nodes contain an array of environmental sensors that are mounted on existing urban infrastructures (such as traffic light poles, building walls, bus stations, etc.) at over 100 locations in Chicago city (Fig. 2a). The nodes measure the meteorological variables, air quality, noise level, and traffic at sub-minute intervals. These measurements are wirelessly transmitted to a data center in real-time and compiled into a complete dataset for public access. Most sensors are located 2 to 4 meters above the ground, thus reasonably representing street-level conditions.

During the designed operation period (Jan 2018 to Sept 2021), the AoT network effectively collected air temperature readings from 106 nodes with 3100 measurement hours per node and an overall 11% effective rate. The most effective observations are between Jul 2018 and Sep 2019 (Fig. S1). To ensure the data quality of these low-cost sensors, we compare the readings from AoT nodes to the nearby research-grade weather stations (Fig. 2b). These high-quality weather recordings can be found at the National Weather Service (NWS), Chicago Data Portal, and National Centers for Environmental Information (NECI) database. The comparison throughout the operation period shows that AoT nodes have a mean bias of 1.88 ^{o}C on T_a with 0.02 ^{o}C/^{o}C slope bias due to internal sensor heating. We correct these biases and conduct thorough quality control on the AoT dataset. But due to the very limited number of research-grade weather stations in Chicago (Fig. 2a), it is possible that the calibrated temperature readings from AoT nodes are still associated with uncertainties, bias, and errors. In this paper, we treat the AoT observation network as an extension of the existing weather stations to serve as the best proxy for the "ground truth" of the urban environment. The full set of AoT data can be downloaded with additional information at http://arrayofthings.github.io/.

2.2.2 Chicago land use inventory

The 2018 parcel-level land use inventory (LUI) for the City of Chicago (Fig. 2a) can be found at the Chicago Metropolitan Agency for Planning (CMAP) in vector format (https://www.cmap.illinois.gov/data/land-use/inventory). The LUI classifies the land use into 10 major and 56 minor categories at an extremely high resolution and can be used to delineate the footprint of individual buildings, roads, streets, boundaries of premises, etc. It can also be used to calculate the fractions of impervious surfaces. Compared to the 30-m National Land Cover Database (NLCD) with four urban categories based on development intensity, the parcel-level LUI has a more detailed classification based on the primary use of the urban land. We cross-check the land use from LUI to the 2019 NLCD to ensure the land cover information used in the numerical model (Section 2.3) accurately represents the built environment. The parcel-level LUI is also used to derive urban features for ML model training and prediction (Section 3.1).

2.2.3 Urban morphology

In addition to the above 2D land cover data, we adopt the three-dimensional (3D) urban morphological data from the Illinois Height Modernization project (https://clearinghouse.isgs.illinois.edu/data/elevation/illinois-height-modernization-ilhmp). This dataset uses Light Detection and Ranging (LiDAR) technology and provides the 3D scan of Illinois at 2-ft (0.6m) spatial resolution in a point cloud format. Using this unique dataset, we extract the vertical urban features such as the heights of buildings and vegetation, tree locations, and coverage. Combined with the parcel-level LUI data, the LiDAR dataset provides the most accurate and precise descriptions of the urban morphology at an extremely high resolution.

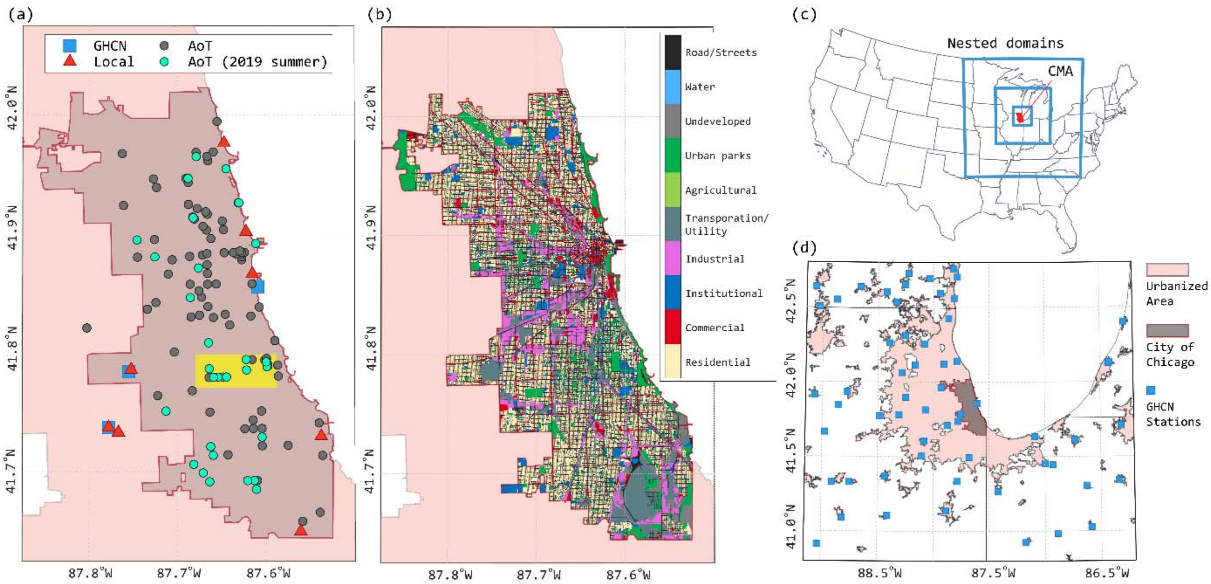

Figure 2. (a) Locations of AoT nodes and other weather stations; **(b)** Land use inventory (LUI) of the City of Chicago from Chicago Metropolitan Agency for Planning; **(c)** Domain configuration of uWRF model; and **(d)** Locations of GHCN weather stations used in the validation of uWRF modeling result over the inner-most domain. CMA: Chicago Metropolitan Area.

2.3 Physical-based urban climate model

In this study, we use the urbanized Weather Research and Forecast model (version 4.3, Skamarock et al., 2019) to provide coarse estimations of hourly 2-m air temperature (T_2) at 1-km resolution. Specifically, we set up three two-way nested domains with the outermost boundary covering the east-north central region of the Midwest US and the innermost domain covering the City of Chicago and its surrounding metropolis (Fig. 2c). The spatial resolutions of the three domains are 9 km, 3 km, and 1 km, respectively. The lateral boundary conditions are from North American Regional Reanalysis (NARR) from the National Center for Environmental Prediction (NCEP, https://rda.ucar.edu/datasets/ds608.0/). Physical schemes of microphysics, convection, radiation, and boundary layer are configured using a well-tested combination for the Chicago region, as described in Sharma et al. (2017). We use the single-layer urban canopy model for impervious urban surface (Chen et al., 2011) and Noah-land surface model (Noah-LSM, Chen and Dudhia, 2001) for natural land and the pervious portion of the urban grids.

In this study, we select the 2019 summer (May 1, 2019 to Aug 31 2019; 123 days) as the simulation period. This period matches the data coverage from the urban informatics described in Section 2.2. In this implementation, the latest NLCD 2019 data (Dewitz and USGS, 2021) is used to derive the land use index and fractions in the uWRF simulation. We use the default three-category single-layer urban canyon parameterizations without any special treatment to keep the simplicity of the model configuration. The modeling results are validated using the air temperature recording from ground weather stations in the Global Historical Climatology Network (GHCN) at 63 locations (Fig. 2d) within the boundary of the inner-most domain. We evaluate the model performance using the root mean squared error (RMSE), mean absolute error (MAE), and mean bias error (MBE), calculated as,

$$\mathrm{RMSE} = \sqrt{\frac{\sum\left(X_{\mathrm{sim}} - X_{\mathrm{obs}}\right)^2}{n}}, \qquad \text{Eq. (1)}$$

$$\mathrm{MAE} = \frac{\sum\left|X_{\mathrm{sim}} - X_{\mathrm{obs}}\right|}{n}, \qquad \text{Eq. (2)}$$

and

$$\mathrm{MBE} = \frac{\sum\left(X_{\mathrm{sim}} - X_{\mathrm{obs}}\right)}{n}, \qquad \text{Eq. (3)}$$

respectively, where X_{sim} is the model simulation; X_{obs} is the observation from ground weather stations at daily or hourly intervals; n is the number of observations. The model daily average RMSE is 2.15 $^{\circ}$C (Fig. 3a), while MAE and MBE are 1.68 $^{\circ}$C and 0.21 $^{\circ}$C, respectively. We also calculate the RMSE for the daily mean temperature from AoT observations in the city (Fig. 3b). Though the RMSE is slightly higher in the urban core (2.52 $^{\circ}$C from AoT) than that in the rest of the study domain (2.15°C from GHCN), the performance of uWRF is acceptable even with the default urban parameterization and can be used as a reliable source of weather prediction over the Chicago Metropolitan Area.

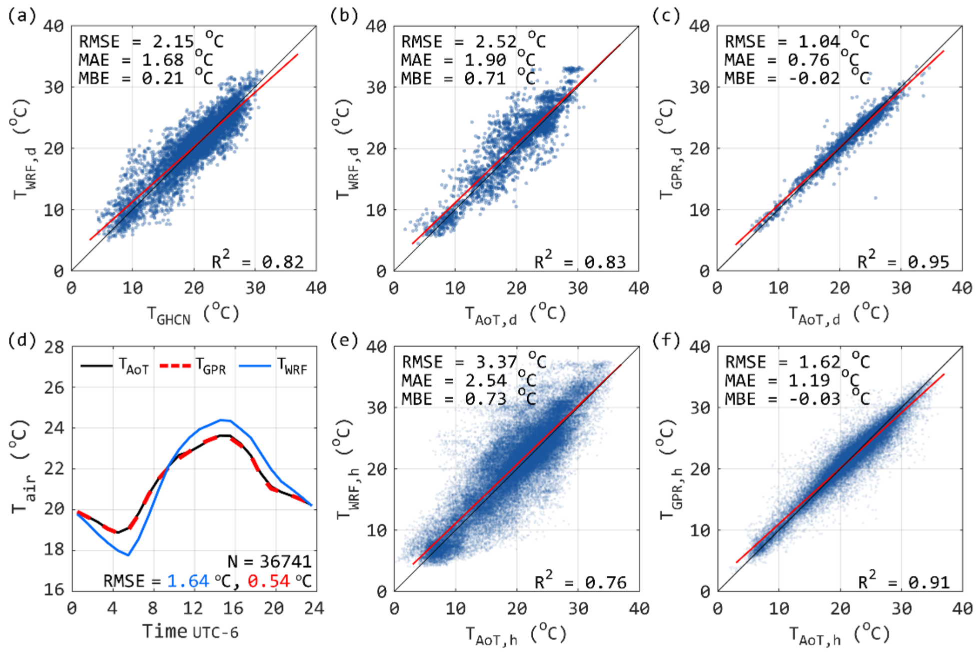

Figure 3. Temperature estimation accuracy in terms of the daily mean air temperature between **(a)** uWRF vs GHCN; **(b)** uWRF vs AoT; **(c)** GPR vs AoT. **(d)** The diurnal variation of air temperature estimated from uWRF (blue line), GPR (red dashed line), and measured by AoT (black line). Hourly temperature estimation accuracy between **(e)** uWRF vs AoT; and **(f)** GPR vs AoT.

2.4 Machine learning model

We adopt Gaussian Process Regression (GPR) to link the 1-km estimation of simulated T_2 to the air temperature (T_a) measured by AoT nodes. GPR is a Bayesian non-parametric model that uses a Gaussian Process (GP) to describe the distribution of the quantity of interest and Bayes' theorem to infer the posterior distribution. Since it is a non-parametric and stochastic model, GPR does not make strong assumptions about the functional form of the relationship between inputs and outputs. Instead, it learns the relationship from the mean and covariance of the dataset and makes predictions using Bayesian inference (Rasmussen and Williams, 2006). GPR has demonstrated exceptional accuracy and robustness in simulating predicted temperatures (Zhang et al., 2021), solar radiation (Lubbe et al., 2020), evaporation (Shabani et al., 2020), and urban environments (Li et al., 2022).

Specific to this study, we train the GPR model with simulated T_2 from WRF and measured T_a from AoT. The model will learn their covariance under the inference from the input urban features. The inputs of the model are selected via a trial-and-error process. Given that the model bias is primarily from the underrepresentation of the complex urban terrain, we formulate the inputs as a group of variables describing the urban morphology (see Section 3.1), such as the building height, impervious surface fraction, canopy height, vegetation coverage, development intensity, etc. After benchmarking the initial model, we test groups of GPR models with different combinations of the variables and pick the model with the best performance. This nominated model is further tuned via the hyperparameter optimization process. The model training and testing are conducted using MATLAB® R2022a. GPR package and library are also available under free software licenses and can be implemented on open-source programming platforms like Python.

3 Results

3.1 Urban features

During the study period, 30 AoT stations have reliable recordings (green dots in Fig. 2a) and can be used in model training and testing. At each available AoT node, 19 features are derived from urban informatics to represent the hyper-local environment (Table 1). It is critical to select a proper spatial range to average and extract the urban features around the nodes. Since most of the AoT sensors are mounted on traffic light poles at intersections, if the averaging scale is too small, the land below will be primarily impervious pavement. At the same time, if the scale is too large, all locations will be similar. The GPR model needs hyper-local characteristics that can interpret the cause of the difference between uWRF prediction and AoT observation at different locations. We test various combinations of urban features as input variables and select the variable list with the best performance (Table 1). For the hyper-local urban features derived from LiDAR data, the optimum averaging radius is 15 m (~50ft), which covers the intersection of neighborhood streets in the US (~11m or 35ft) and the land at the street corner (Fig. S2a). These features provide the impactful factors embedded in the vertical structure of the built environment, such as shading from the building and street trees. While the planar land use information from LUI and NLCD represents the general characteristics of the street blocks centered around the node. Since a typical city block in Chicago is around 100 m by 200 m, we select 200 m as the averaging radius for LUI and NLCD data to cover 3 to 4 street blocks (Fig. S2b). While the model performance will not change significantly with the change of the averaging range on a reasonable scale, we observe an optimum value at approximately 200 meters. If the feature is too small, the model may be less likely to capture its impact on the environment; for instance, parks need to be large enough to cause a cooling effect. On the other hand, the optimum averaging range may imply a minimum scale that needs to be considered to reflect the heterogeneity in urban models. Nevertheless, our analysis indicates that the model performance is more sensitive to vertical urban features from LiDAR to horizontal urban features from LUI/NLCD, indicating a notable influence of vertical urban morphology on the thermal environment.

Our analysis assumes that urban features do not change during the summer months but should be distinct at each location to provide wide coverage of the variable space for model training. Nonetheless, it is impossible to fully represent the diverse land use of Chicago by a limited number of nodes at discrete locations. Figure S3 shows the histograms of the extracted urban features. Most of the AoT nodes in training are in residential areas with a medium to high development intensity. This causes some of the features, like mean tree height (Var. 07) and building fraction (Var. 08), not to have a continuous distribution (Fig. S3), which may affect out-of-sample performance (Section 3.3) and lead to model uncertainties (Section 4.1).

3.2 GPR prediction on time-series

The input variables for the GPR model contain the time of day, the basic meteorological conditions from uWRF, and the urban features extracted at 30 AoT locations. Due to the measurement inconsistencies at some locations, the total usable data volume is 36741 measurement hours after quality control. We normalize the variables to the prescribed ranges, respectively (Table 1), and randomly select 30% (N = 11022) of the normalized data as the training dataset. The rest 70% (N = 25719) are used as the validation dataset to test the model performance. There is no overlapped data in model training and validation. Once trained, the GPR model estimates the street-level temperature at the AoT locations. We then compare the temperature from GPR (T_{GPR}) and the previous uWRF prediction (T_{WRF}) to the "ground truth" measurement from AoT nodes (T_{AoT}). Figure. 3c&f show the scatter plots of the temperature estimations at daily and hourly intervals, respectively. The GPR model improves the estimation accuracy (calculated as RMSE) from 2.52 $^{\circ}$C to 1.04 $^{\circ}$C on the daily mean (Fig. 3b c.f. 3b) and 3.31 $^{\circ}$C to 1.66 $^{\circ}$C at hourly intervals (Fig. 3e c.f. 3f). We also find that the GPR model replicates the mean of the diurnal cycle over the 30 stations with high accuracy (Fig. S4) and improve the model RMSE from 1.64 ^{o}C to 0.50^{o}C (Fig. 3d). From the diurnal cycle, it is notable that uWRF constantly underestimates nighttime (daily minimum) temperature while overestimates daytime (daily maximum) temperature at given urban locations (Fig. 3d). This phenomenon indicates that the parameterizations and physics in uWRF underestimate the thermal inertia of urban land, which resists the change of temperature and largely contribute to urban heat island effect at night (Varquez and Kanda, 2018).

Additionally, we test the model performance when varying the training sample size. We retrain the models with 1% (N = 367) to 70% (N = 18371) of total usable measurement hours and track the change in model RMSE. Since the maximum training ratio is 70%, we reduce the validation dataset to 30% (N = 11022) for all models. Both training and validation data are randomly selected and do not overlap. For different training sample sizes, we train 40 models separately and show the mean RMSE at each sample size in Figure 4a. The model accuracy increases with increasing sample size. For example, even with a minimal training dataset (1%), the model can improve simulated hourly temperature RMSE from around 3.28 ^{o}C (uWRF) to 2.50^{o}C (GPR). The likely rationale for this improvement is that the GPR model captures the general trend and corrects uWRF by counteracting the underestimation in thermal inertia, which is responsible for most bias in uWRF. When increasing the training sample size to 20%, the model RMSE reduces below 2 ^{o}C (Fig. 4a). Model performance continuously increases, but the improvement becomes incremental when the training sample size is greater than 20%. Using more data from diverse land uses, the model can identify other factors contributing to uWRF bias, such as the inaccurate building height and the lack of trees in street canyons. Although, as noted in Section 3.1, the 30 locations in training do not represent all possible combinations of land use mix in Chicago; therefore, additional testing is needed to ensure the accuracy of spatial patterns of air temperature estimations.

Table 1. List of input variables for GPR models.

No.	Variable	Description	Unit	Source	Min.	Max.
Var.01	T_2	2-meter air temperature	^oC	uWRF	0	40
Var.02	Q_2	2-meter air humidity	kg kg^{-1}	uWRF	0	0.02
Var.03	SW	Shortwave radiation	W m^{-2}	uWRF	0	1200
Var.04	U	Wind speed	m s^{-1}	uWRF	0	20
Var.05	H_b	Mean building height	ft	LiDAR, ILHMP	0	80
Var.06	$H_{b,max}$	Maximum building height	ft	LiDAR, ILHMP	0	100
Var.07	H_t	Mean tree height	ft	LiDAR, ILHMP	0	60
Var.08	F_b	Building fraction	-	LiDAR, ILHMP	0	0.6
Var.09	F_v	Vegetation fraction	-	LiDAR, ILHMP	0	0.6
Var.10	F_1	Fraction of residential land	-	LUI, CMAP	0	0.8
Var.11	F_2	Fraction of commercial land	-	LUI, CMAP	0	0.2
Var.12	F_3	Fraction of institutional land	-	LUI, CMAP	0	0.6
Var.13	F_4	Fraction of industrial land	-	LUI, CMAP	0	0.4
Var.14	F_5	Fraction of transportation land	-	LUI, CMAP	0	0.2
Var.15	F_6	Fraction of agricultural land	-	LUI, CMAP	0	0.4
Var.16	F_7	Fraction of urban parks	-	LUI, CMAP	0	0.2
Var.17	F_8	Fraction of undeveloped land	-	LUI, CMAP	0	0.2
Var.18	F_9	Fraction of road/street	-	LUI, CMAP	0	0.4
Var.19	F_{low}	Fraction of low-density urban land	-	NLCD	0	0.8
Var.20	F_{mid}	Fraction of mid-density urban land	-	NLCD	0	0.8
Var.21	F_{high}	Fraction of high-density urban land	-	NLCD	0	1
Var.22	F_{imp}	Fraction of impervious surface	-	NLCD	40	100
Var.23	F_{can}	Fraction of tree canopy	-	NLCD/LiDAR	0	60
Var.24	t	Time of day	-	-	0	24

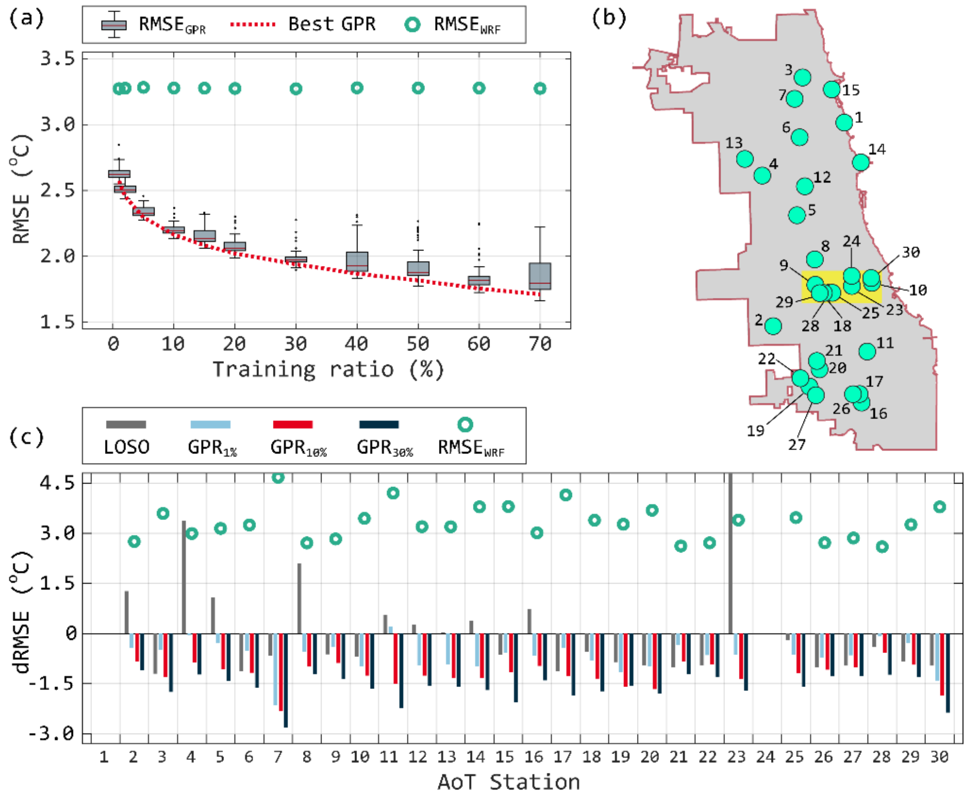

Figure 4. (a) Model performance variation with different training data volume; **(b)** locations and numbering of the AoT nodes used in GPR model training and testing; **(c)** spatial variation of the model performance and the comparison between "leave-one-station-out" models (LOSO) and the complete GPR models with different training data volume (see Section 3.3).

<u>3.3 GPR prediction on spatial patterns</u>

We apply a "leave-one-out" cross-validation (Hastie et al., 2009) to further investigate the model capabilities. Specifically, we train a series of models using data from 29 stations and leaving out one station at a time (Fig. 4b). All measurements from the leave-out station are used for model validation. In this case, we use these "<u>l</u>eave-<u>o</u>ne-<u>s</u>tation-<u>o</u>ut" models (LOSO models hereafter) to test the models' capability in interpreting the impact of land use mix on temperature estimation. Figure 4c shows the improvement of model RMSE, given as the station-wise change of RMSE, calculated by

$$\mathrm{dRMSE} = \mathrm{RMSE}_{\mathrm{GPR}} - \mathrm{RMSE}_{\mathrm{WRF}}, \qquad \text{Eq. (4)}$$

where the subscript "GPR" indicates the model trained on data from 30 locations, while "LOSO" indicates the models trained by the LOSO approach. When dRMSE < 0, the LOSO model is better than uWRF, verse visa. We find LOSO models are generally reliable over the 30 locations but with a few exceptions. For example, the LOSO model without data from Station 23 performs poorly at the leave-out location with a much higher RMSE than uWRF (Fig. 4c). Similar situations apply to a few other stations. However, when comparing LOSO models with GPR models trained by data from all locations, the latter performs better even with a very small training sample size. More specifically, the GPR model with 1% training data is significantly more accurate than the LOSO model at Station 23 (Fig. 4c). The drastic difference between LOSO models and the previous GPR models indicates the importance of including more locations in the model training for better performance and less uncertainty when applying the mode to new locations. Stations with positive dRMSE in Fig. 4c are essential for the integration of the model. Nevertheless, LOSO models generally estimate the hourly temperature with reasonable accuracy. They are sometimes comparable to GPR models with 10% training data, such as Station 3, 6, 17, 21, 22, 26, 27, and 29 (Fig. 4c). The result indicates that the model trained on 30 locations can interpret most of the spatial variations influenced by the land use mix. Therefore, we can use the trained model to generate street-level temperature based on uWRF simulation and urban informatics.

Figure 5 shows a map of the temperature deviations in a selected area to its areal mean. This street district is in south Chicago (yellow patch in Fig. 2a) and has a diverse land use with various development intensities, including residential, commercial, highways, and parks. There are 9 active AoT stations used in model training. We randomly sampled 17500 spots, shown as temperature dots in Fig. 5a, and estimated the temperatures using the GPR model trained on 30 AoT stations. Figure 5b shows landmarks that have recognizable temperature deviations from the areal mean. For example, the temperature near urban parks (P1-P3) is noticeably lower, but the cooling is limited to the street block adjacent to the park. Similarly, streets along the green belt (R1) are cooler than the surroundings due to high vegetation fraction and canopy coverage. These localized cool zones can connect and form a cool corridor if they are close to each other, such as the area between R1 and P1. In contrast, areas with fewer trees, more impervious areas, and tall buildings have hotter temperatures. For instance, the freeway (R2), commercial corner (C1), and hospital (C2) are hotter than their surroundings. Interestingly, the cooling is not significant in the university campus (U1) near the hospital (C2) despite the campus having a very dense canopy. Comparing area U1 to a similar area, U2, the result shows that cooling at U2 is more noticeable. The heights of vegetation and buildings in area U2 are both low. Dense trees and tall buildings may trap the heat released from the buildings. But we find the campus (U1) cooler than the university hospital (C2), which has taller buildings and fewer trees.

Areas H1 and H2 have the same land use classification but different development intensities. Most homes in area H1 are 2-story single-family houses. Street trees there are nearly twice as high as the homes. While area H2 has more apartments and multi-family homes with 3 and more stories. Street trees are at the same level as the buildings. The mean temperature difference of H2 is around 1^{o}C to 1.5 ^{o}C larger than H1, mostly due to dense land use, proximity to the highway (R2) and non-vegetated railway yard south of it.

The above examples illustrate how land cover and urban morphology affect air temperature at a hyper-local scale. The results are consistent with the current knowledge of urban microclimate dynamics. More importantly, the temperature in our illustrative zone can vary by up to 2 ^{o}C between two locations only a few street blocks apart. These spatial variations are difficult to capture in models with coarser resolutions. The GPR model successfully identifies the cool and hot spots within the zone and explains the spatial variation in temperature based on urban characteristics.

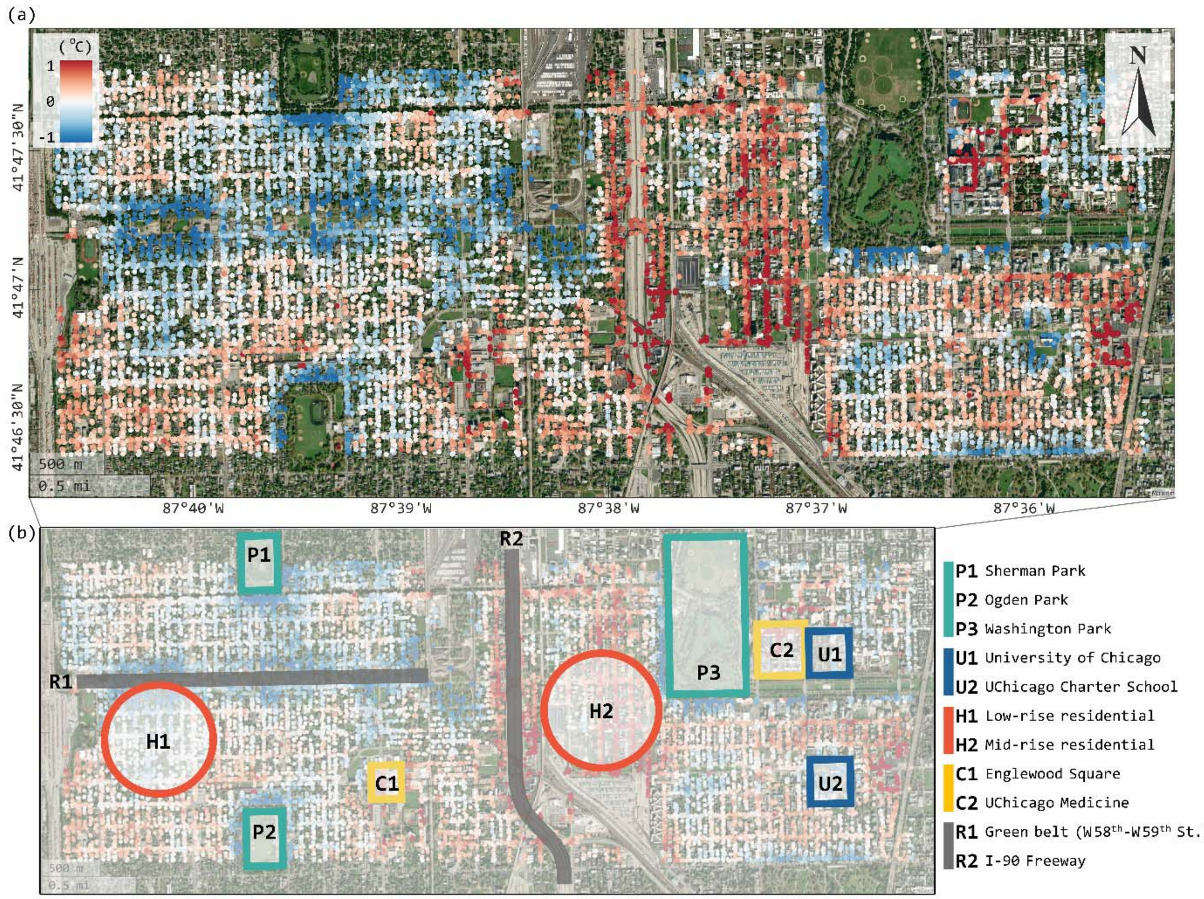

Figure 5. (a) Air temperature estimations from the GPR model on the streets over a selected Chicago neighborhood. The color map shows the temperature deviations from the mean air temperature of this region. **(b)** The representative street blocks and landmarks in this region as examples of distinctive air temperature deviations at hyper-local scales. The location of this neighborhood is shown in the yellow box in Fig. 2a.

4. Discussions

4.1 Model uncertainties

The study aims to develop a modeling framework based on advanced urban informatics, the state-of-art physical-based numerical model, and a machine learning algorithm. As a result, the uncertainties associated with individual components contribute to the overall uncertainty of the modeling framework. The quality and quantity of AoT data can be the primary sources of uncertainty. The data processing phase reveals that most AoT sensors operate intermittently throughout the operating period, resulting in discontinuous measurements. Thus, a small percentage of the data is deemed credible for scientific use after rigorous quality control and calibration. Our results also show that the volume of training data significantly impacts model performance for timeseries predictions (Section 3.2). The LOSO test in Section 3.3 also illustrates that the station coverage (i.e., the number of available stations) is critical for interpreting the spatial patterns and completing the modeling framework. As these in-situ measurements serve as the modeling target and provide the basis for model validation, better data quality and quantity will help users gain confidence when interpreting the results from the ML models.

The uncertainty of uWRF also contributes to the overall variability of model performance and credibility. uWRF provides the basic weather conditions at a larger spatial scale, leaving the subsequent ML model to explain the impacts of environmental factors from the hyper-local features. Therefore, the uWRF is critical in providing the initial estimate based on regional geological features and mesoscale atmospheric dynamics. Uncertainty in uWRF mainly arises from the parameterization of urban surfaces, including the underrepresentation of land cover heterogeneity, the lack of critical land surface processes, and the inaccurate descriptions of urban morphology. As resolving urban hydroclimate dynamics is essential for climate models across the scales (Sharma et al., 2021), we intentionally did not modify uWRF specifically for the Chicago region to showcase the capability of the ML models better. Nevertheless, uWRF can be more accurate and bias-free with a more detailed urban canyon parameterization dataset (Ching et al., 2018), but will run at a slower speed. As a result, ML models are better candidates to explain the discrepancies between uWRF results and the observations.

The uncertainties from both measurements (AoT) and model outputs (uWRF) accumulate when the datasets are used as model inputs. The GPR model is aimed at learning these uncertainties by quantifying the relationship between the hyper-local urban features and uWRF bias. An inherent uncertainty associated with GPR models and most ML algorithms in supervised learning is their ability to predict beyond the training dataset. This issue can be mitigated by including more data, as illustrated in the LOSO test, or by extending the training period (e.g., using three summers instead of one). Alternatively, one may inform the ML model with the results from finer-scale uncoupled urban canopy models, as they typically represent local urban features better and are considered more accurate at a neighborhood scale. As a result of this adaptive learning approach, ML models can better understand the physical dynamics at the street-level and may effectively turn the extrapolation problem (i.e., predicting at new locations without constraints) into an interpolation problem (i.e., constrained by physical laws). Both approaches for uncertainty mitigation require additional field and modeling efforts in the future but are worth pursuing, as these efforts can serve as a springboard for the development of the next generation of urban climate and weather models.

4.2 Implications from GPR models

The urban informatics used in this study is at a sub-meter level and requires spatial averaging to reflect environmental characteristics while retaining spatial variation information. The temperature deviation map (Fig. 5) from GPR illustrates how land cover and urban morphology affect air temperature at a hyper-local scale. The results are consistent with the current knowledge of urban microclimate dynamics. More importantly, the temperature in our illustrative zone can vary by up to 2 ^{o}C between two locations only a few street blocks apart. These spatial variations are difficult to capture in models with coarser resolutions. However, due to the limitations of computational resources and model stability, it is challenging to incorporate exhaustive urban informatics into physical-based models. The recent implementations of urban canopy parameterization and LCZ in uWRF have demonstrated the growing demand for high-resolution urban simulations (Chen et al., 2022). In our study, we find that the averaging range affects model performance in a non-monotonic manner, suggesting an optimum scale for certain urban features when simulating built environments (Section 3.1). Thus, there is a need for further research on this scaling problem to establish expectations for the next generation of urban weather and climate models. This may be accomplished by running parameter optimization on fast surrogate models or adopting the explainable artificial intelligence (XAI) approach. Our study, although in its infancy, explains the scale problem, illustrates the capabilities of analyzing this issue from an ML perspective, and sheds light on future endeavors in this field.

When examining the internal logic of GPR models, we notice the spatial variability of the model performance (Section 3.3), which means that the data at certain locations are critical and cannot be replaced or compensated by data from other locations. This implies that the monitoring locations need to be carefully designed if only a limited number of sensors are available. In recent years, these hyper-local, dense, and real-time sensors have become the most common method to collect data in cities (Alvarez et al., 2019; Ma et al., 2019; Enlund et al., 2022). Many cities around the world have urban observation networks that are used to study urban climate, such as Baltimore (Shi et al., 2021), Twin Cities (Smoliak et al., 2015), Shanghai (Tan et al., 2015), Tainan (Chen et al., 2019), to name a few. In Chicago, a new generation of urban sensors are being deployed under the SAGE project as a successor of AoT weather nodes. There are also air quality sensors available from Microsoft Research that cover a wide range of demographics (Esie et al., 2022). In the future deployment of such monitoring networks or other cyber-infrastructures for urban informatics, the proposed GPR framework can provide the key locations that need to be monitored for the best efficiency.

In addition to the accuracy and informativeness of GPR models, fast computation speed is another merit for the potential end users of this modeling framework. For example, hyper-local weather conditions can significantly affect the walkability and drivability of streets. Real-time weather information at high resolution can assist pedestrians and autonomous vehicles in better-informed decisions when traversing the city. Given the initial success of air temperature estimation, adopting this proposed framework to the other meteorological conditions is applicable once the other types of high-resolution in-situ data are available.

4.3 Limitations

In this case study, we recognize the importance of the unique datasets and the pioneering efforts of the City of Chicago. Though urban informatics at sub-meter resolution and dense observation networks are gaining attention and being deployed in cities around the globe, they are not widely available. A few alternative sources of urban informatics can extend the application range to global cities. For example, height information can be derived from the high-resolution Global Ecosystem Dynamics Investigation (GEDI) LiDAR dataset. As originally designed to retrieve canopy height, it can extract building height information via the waveform profile with a horizontal resolution of 25m. Alternatively, a deep learning model may also be used to retrieve height when combined with point cloud LiDAR data (Kamath et al., 2022), synthetic aperture radar (Sun et al., 2022), or street-view imagery (Al-Habashna, 2020). Canopy height, however, needs to be investigated at a finer resolution due to the small footprint of individual trees and their hyper-local impact on the environment. Most current approaches to quantify canopy height reply on LiDAR data at smaller scales (Lee et al., 2016; Matasci et al., 2018; Heo et al., 2019; Xuan et al., 2023).

A GPR model trained on the Chicago dataset may not be applicable to other regions without additional manipulation. However, the credibility is dependent on the similarity of the target city and the Chicago region in terms of climate, geography, land use, etc. As proposed in a few pioneering studies (Wang et al., 2018; Zhao et al., 2021; Chen et al., 2022), transfer learning can also migrate knowledge between cities. The future development of the model should include deeper investigations of the spatial variability of the model to reduce uncertainties when implementing the model at new locations.

5. Concluding remarks

High-resolution urban informatics provides new opportunities for the advancement of urban weather and climate modeling techniques. In conjunction with the conventional numerical model uWRF and the AoT observation network, we demonstrated the capability of the GPR models to predict temperature timeseries and spatial patterns. The model framework proposed in this study successfully estimated the hyper-local street air temperature in the City of Chicago and with a high degree of accuracy. In the context of data-driven and high-resolution urban models, we investigated the model uncertainties and highlighted the critical importance of data quality and data quantity. The implications derived from the model performance and sensitivity analysis can guide future design and deployment of cyberinfrastructures for cost-efficient urban environment observations. Based on the findings, we also identified the prospects for future iterations of the model based on data availability, modeling capability, and the user community's needs.

While the study is novel, several caveats may prevent it from being a universal approach for a larger collection of cities. As urban informatics advances, this study will be one of the first to invest and harvest the joint efforts of urban research communities in an interdisciplinary manner as a significant contribution to improving the resilience, efficiency, and livability of modern cities.

Acknowledgment

This research is supported by the Walder Foundation and NSF awards #139316 and 2230772. This work is also supported by the U.S. Department of Energy, Office of Science, Biological and

Environmental Research, under contract DE-AC02-06CH11357. We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We also acknowledge NOAA, City of Chicago, and Chicago Metropolitan Agency for Planning for providing the data used in this study.

Conflict of Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Research Statement

All the datasets used in this study are publicly available with open access and allow direct download. The Chicago land use inventory (LUI) can be found at https://www.cmap.illinois.gov/data/land-use/inventory. AoT dataset can be found at https://www.mcs.anl.gov/research/projects/waggle/downloads/datasets/AoT_Chicago.complete.latest.tar. Illinois Height Modernization (ILHMP) LiDAR Data can be found at https://clearinghouse.isgs.illinois.edu/data/elevation/illinois-height-modernization-ilhmp. GHCN dataset can be found at https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily. NLCD dataset can be found at https://www.mrlc.gov/data/nlcd-2019-land-cover-conus.

Reference

Al-Habashna, A. (2020). An open-source system for building-height estimation using street-view images, deep learning, and building footprints. Retrieved from: https://www150.statcan.gc.ca/n1/pub/18-001-x/18-001-x2020002-eng.pdf

Alvarez, M.G., et al. (2019) Integration and exploitation of sensor data in smart cities through event-driven applications. *Sensors* **19** 6 1372. https://doi.org/10.3390/s19061372

Anderson, C.A. (1989) Temperature and aggression: Ubiquitous effects of heat on occurrence of human violence. *Psychological Bulletin* **106** 74-96. http://doi.org/10.1037/0033-2909.106.1.74

Chakraborty, T., et al. (2019) Disproportionately higher exposure to urban heat in lower-income neighborhoods: a multi-city perspective. *Environmental Research Letters* **14** 10 105003. http://doi.org/10.1088/1748-9326/ab3b99

Chen, F. and Dudhia, J. (2001) Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part II: Preliminary model validation. *Monthly Weather Review* **129** 4 587-604. http://doi.org/10.1175/1520-0493(2001)129<0587:Caalsh>2.0.Co;2

Chen, F., et al. (2011) The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *International Journal of Climatology* **31** 2 273-288. http://doi.org/10.1002/joc.2158

Chen, G., et al. (2022) A cross-city federal transfer learning framework: A case study on urban region profileing. *arXiv*. https://doi.org/10.48550/arXiv.2206.00007

Chen, F., et al. (2022) Improved urban finescale forecasting during a heat wave by using high-resolution urban canopy parameters. *Frontiers in Climate* **3** http://doi.org/10.3389/fclim.2021.771441

Chen, Y.-C., et al. (2019) The application of a high-density street-level air temperature observation network (HiSAN): The relationship between air temperature, urban development, and geographic features. *Science of The Total Environment* **685** 710-722. http://doi.org/https://doi.org/10.1016/j.scitotenv.2019.06.066

Chen, Y. and Zhai, P. (2017) Revisiting summertime hot extremes in China during 1961–2015: Overlooked compound extremes and significant changes. *Geophysical Research Letters* **44** 10 5096-5103. http://doi.org/https://doi.org/10.1002/2016GL072281

Ching, J., et al. (2018) WUDAPT: An urban weather, climate, and environmental modeling infrastructure for the anthropocene. *Bulletin of the American Meteorological Society* **99** 9 1907-1924. http://doi.org/https://doi.org/10.1175/BAMS-D-16-0236.1

Conry, P., et al. (2015) Chicago's heat island and climate change: Bridging the scales via dynamical downscaling. *Journal of Applied Meteorology and Climatology* **54** 7 1430-1448. http://doi.org/https://doi.org/10.1175/JAMC-D-14-0241.1

Dadioti, R. and Rees, S. (2017) performance of detached eddy simulation applied to analysis of a university campus wind environment. *Energy Procedia* **134** 366-375. http://doi.org/https://doi.org/10.1016/j.egypro.2017.09.551

Dewitz, J., and U.S. Geological Survey (2021) National Land Cover Database (NLCD) 2019 products (ver. 2.0, June 2021): U.S. Geological Survey data release, https://doi.org/10.5066/P9KZCM54

Enlund, D., et al. (2022) The role of sensors in the production of smart city spaces. *Big Data & Society* **9** 2 20539517221110218. http://doi.org/10.1177/20539517221110218

Esie, P., et al. (2022) Neighborhood composition and air pollution in Chicago: Monitoring inequities with a dense, low-cost sensing network, 2021. *American Journal of Public Health* **112** 12 1765-1773. http://doi.org/10.2105/ajph.2022.307068

González, J.E., et al. (2021) Urban climate and resiliency: A synthesis report of state of the art and future research directions. *Urban Climate* **38** 100858. http://doi.org/https://doi.org/10.1016/j.uclim.2021.100858

Goret, M., et al. (2019) Inclusion of CO2 flux modelling in an urban canopy layer model and an evaluation over an old European city centre. *Atmospheric Environment: X* **3** http://doi.org/10.1016/j.aeaoa.2019.100042

Hastie, T., Tibshirani, R., and Friedman, J. (2009) The elements of statistical learning: Data mining, inference, and prediction. Springer, New York.

Heaviside, C., et al. (2017) The urban heat island: Implications for health in a changing environment. *Current Environmental Health Reports* **4** 3 296-305. http://doi.org/10.1007/s40572-017-0150-3

Heo, H.K., et al. (2019) Estimating the heights and diameters at breast height of trees in an urban park and along a street using mobile LiDAR. *Landscape and Ecological Engineering* **15** 3 253-263. http://doi.org/10.1007/s11355-019-00379-6

Huang, K., et al. (2019) Projecting global urban land expansion and heat island intensification through 2050. *Environmental Research Letters* **14** 11 114037. http://doi.org/10.1088/1748-9326/ab4b71

Järvi, L., et al. (2011) The Surface Urban Energy and Water Balance Scheme (SUEWS): evaluation in Los Angeles and Vancouver. *Journal of Hydrology* **411** 3-4 219-237. http://doi.org/10.1016/j.jhydrol.2011.10.001

Järvi, L., et al. (2019) Spatial modeling of local-scale biogenic and anthropogenic carbon dioxide emissions in Helsinki. *Journal of Geophysical Research: Atmospheres* **124** 15 8363-8384. http://doi.org/https://doi.org/10.1029/2018JD029576

Kamath, H.G., et al. (2022) GLOBUS: GLObal Building heights for urban studies. *arXiv*. https://doi.org/10.48550/arXiv.2205.12224

Kumar, P. (2021) Climate change and cities: Challenges ahead. *Frontiers in Sustainable Cities* **3** http://doi.org/10.3389/frsc.2021.645613

Lee, J.-H., et al. (2016) The feasibility of remotely sensed data to estimate urban tree dimensions and biomass. *Urban Forestry & Urban Greening* **16** 208-220. http://doi.org/https://doi.org/10.1016/j.ufug.2016.02.010

Li, P. and Wang, Z.-H. (2020) Modeling carbon dioxide exchange in a single-layer urban canopy model. *Building and Environment* **184** 107243. http://doi.org/10.1016/j.buildenv.2020.107243

Li, P., et al. (2022) Multi-objective optimization of urban environmental system design using machine learning. *Computers, Environment and Urban Systems* **94** 101796. http://doi.org/10.1016/j.compenvurbsys.2022.101796

Lubbe, F., et al. (2020) Evaluating the potential of Gaussian Process Regression for solar radiation forecasting: A case study. *Energies* **13** 20 5509. https://doi.org/10.3390/en13205509

Ma, M., et al. (2019) Data sets, modeling, and decision making in smart cities: A survey. *ACM Trans. Cyber-Phys. Syst.* **4** 2 Article 14. http://doi.org/10.1145/3355283

Maronga, B., et al. (2020) Overview of the PALM model system 6.0. *Geosci. Model Dev.* **13** 3 1335-1372. http://doi.org/10.5194/gmd-13-1335-2020

Masson, V. (2000) A Physically-based scheme for the urban energy budget in atmospheric models. *Boundary-Layer Meteorology* **94** 3 357-397. http://doi.org/10.1023/A:1002463829265

Matasci, G., et al. (2018) Mapping tree canopies in urban environments using airborne laser scanning (ALS): a Vancouver case study. *Forest Ecosystems* **5** 1 31. http://doi.org/10.1186/s40663-018-0146-y

Meili, N., et al. (2020) An urban ecohydrological model to quantify the effect of vegetation on urban climate and hydrology (UT&C v1.0). *Geoscientific Model Development* **13** 1 335-362. http://doi.org/10.5194/gmd-13-335-2020

Middel, A., et al. (2022) Urban climate informatics: An emerging research field. *Frontiers in Environmental Science* **10** http://doi.org/10.3389/fenvs.2022.867434

O'Brien, G.A., et al. (2019) The heat penalty of walkable neighbourhoods. *International Journal of Biometeorology* **63** 3 429-433. http://doi.org/10.1007/s00484-018-01663-0

Oke, T.R. (1988) The urban energy balance. *Progress in Physical Geography: Earth and Environment* **12** 4 471-508. http://doi.org/10.1177/030913338801200401

Oke, T.R., et al. (2017). Urban Climates. Cambridge, Cambridge University Press.

Patel, P., et al. (2023) Deep learning based urban morphology for city-scale environmental modeling. *PNAS Nexus* http://doi.org/10.1093/pnasnexus/pgad027

Perkins-Kirkpatrick, S.E. and Lewis, S.C. (2020) Increasing trends in regional heatwaves. *Nature Communications* **11** 1 3357. http://doi.org/10.1038/s41467-020-16970-7

Rasmussen, C.E. and Williams, C.K.I. (2006). Gaussian processes for machine learning, The MIT Press.

Reeping, P.M. and Hemenway, D. (2020) The association between weather and the number of daily shootings in Chicago (2012–2016). *Injury Epidemiology* **7** 1 31. http://doi.org/10.1186/s40621-020-00260-3

Revi, A., et al. (2014) Urban areas. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 535-612.

Ryu, Y.-H., et al. (2015) Realistic representation of trees in an urban canopy model. *Boundary-Layer Meteorology* **159** 2 193-220. http://doi.org/10.1007/s10546-015-0120-y

Shabani, S., et al. (2020) Modeling pan evaporation using Gaussian Process Regression, K-Nearest Neighbors, Random Forest and Support Vector Machines: Comparative analysis. *Atmosphere* **11** 1 66. https://doi.org/10.3390/atmos11010066

Sharma, A., et al. (2016) Green and cool roofs to mitigate urban heat island effects in the Chicago metropolitan area: evaluation with a regional climate model. *Environmental Research Letters* **11** 6 064004. http://doi.org/10.1088/1748-9326/11/6/064004

Sharma, A., et al. (2017) Urban meteorological modeling using WRF: A sensitivity study. *International Journal of Climatology* **37** 4 1885-1900. http://doi.org/https://doi.org/10.1002/joc.4819

Sharma, A., et al. (2021) The need for urban-resolving climate modeling across scales. *AGU Advances* **2** 1 http://doi.org/10.1029/2020av000271

Shen, C., et al. (2019) Impacts of high-resolution urban canopy parameters within the wrf model on dynamical and thermal fields over Guangzhou, China. *Journal of Applied Meteorology and Climatology* **58** 5 1155-1176. http://doi.org/https://doi.org/10.1175/JAMC-D-18-0114.1

Shi, R., et al. (2021) Monitoring intra-urban temperature with dense sensor networks: Fixed or mobile? An empirical study in Baltimore, MD. *Urban Climate* **39** 100979. http://doi.org/https://doi.org/10.1016/j.uclim.2021.100979

Skamarock, W.C., et al. (2019) A description of the advanced research WRF version 4. http://doi.org/10.6084/m9.figshare.7369994.v4

Smoliak, B.V., et al. (2015) Dense network observations of the Twin Cities canopy-layer urban heat island. *Journal of Applied Meteorology and Climatology* **54** 9 1899-1917. http://doi.org/https://doi.org/10.1175/JAMC-D-14-0239.1

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

Sun, Y., et al. (2022) Large-scale building height retrieval from single SAR imagery based on bounding box regression networks. *ISPRS Journal of Photogrammetry and Remote Sensing* **184** 79-95. http://doi.org/https://doi.org/10.1016/j.isprsjprs.2021.11.024

Suter, I., et al. (2022) uDALES 1.0: a large-eddy simulation model for urban environments. *Geosci. Model Dev.* **15** 13 5309-5335. http://doi.org/10.5194/gmd-15-5309-2022

Tan, J., et al. (2015) Urban integrated meteorological observations: Practice and experience in Shanghai, China. *Bulletin of the American Meteorological Society* **96** 1 85-102. http://doi.org/https://doi.org/10.1175/BAMS-D-13-00216.1

Toparlar, Y., et al. (2017) A review on the CFD analysis of urban microclimate. *Renewable and Sustainable Energy Reviews* **80** 1613-1640. http://doi.org/10.1016/j.rser.2017.05.248

UN-Habitat (2019) *World Urbanization Prospects: The 2018 Revision*. United Nations. New York. Retrieved from: https://population.un.org/wup/publications/Files/WUP2018-Report.pdf

US Census Bureau. (2020). *Urban and rural*. Retrieved from https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html

Varquez, A.C.G. and Kanda, M. (2018) Global urban climatology: a meta-analysis of air temperature trends (1960–2009). *npj Climate and Atmospheric Science* **1** 1 32. http://doi.org/10.1038/s41612-018-0042-8

Vavrus, S. and Van Dorn, J. (2010) Projected future temperature and precipitation extremes in Chicago. *Journal of Great Lakes Research* **36** 22-32. http://doi.org/https://doi.org/10.1016/j.jglr.2009.09.005

Venter, Z.S., et al. (2020) Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sensing of Environment* **242** 111791. http://doi.org/https://doi.org/10.1016/j.rse.2020.111791

Wang, C., et al. (2021) A single-layer urban canopy model with transmissive radiation exchange between trees and street canyons. *Building and Environment* **191** 107593. http://doi.org/10.1016/j.buildenv.2021.107593

Wang, L., Guo, B., and Yang, Q. (2018) Smart city development with urban transfer learning. *Computer*. **51** 12 32-41. http://doi.org/10.1109/MC.2018.2880015

Xuan, J., et al. (2023) Intelligent estimating the tree height in urban forests based on deep learning combined with a smartphone and a comparison with UAV-LiDAR. *Remote Sensing* **15** 1 97. https://doi.org/10.3390/rs15010097

Yang, J. and Wang, Z.-H. (2014) Physical parameterization and sensitivity of urban hydrological models: Application to green roof systems. *Building and Environment* **75** 250-263. http://doi.org/10.1016/j.buildenv.2014.02.006

Yin, Y., et al. (2020) Urban ambient air temperature estimation using hyperlocal data from smart vehicle-borne sensors. *Computers, Environment and Urban Systems* **84** 101538. http://doi.org/https://doi.org/10.1016/j.compenvurbsys.2020.101538

Zhang, Y., et al. (2021) A Gaussian process regression-based sea surface temperature interpolation algorithm. *Journal of Oceanology and Limnology* **39** 4 1211-1221. http://doi.org/10.1007/s00343-020-0062-1

Zhao, G., et al. (2021) Improving urban flood susceptibility mapping using transfer learning. *Journal of Hydrology* **602** 126777. http://doi.org/https://doi.org/10.1016/j.jhydrol.2021.126777

Zumwald, M., et al. (2021) Mapping urban temperature using crowd-sensing data and machine learning. *Urban Climate* **35** 100739. http://doi.org/https://doi.org/10.1016/j.uclim.2020.100739