Proceedings of the 4th International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2026.38271

Machine Learning and GIS Approaches to Spatial Prediction of Urban Air Temperature

Baiyang Li

College of Social Science, Michigan State University, East Lansing, USA
libaiyan@msu.edu

Abstract. Air temperature is measured directly at weather stations, with small spatial coverage. Weather observations, geographic information systems (GIS), and machine learning can be used to estimate the temperatures in these areas based on the conditions around them. This paper uses literature synthesis and case study analysis to review the data and methods used in spatial prediction of urban air temperature. It investigates the aspects of meteorological observation, spatial and environmental predictors of the city, GIS-based feature extraction, machine learning models, and temperature mapping. Three case studies in Zurich, Oslo, and Szeged demonstrate the application of high-resolution mapping, the predictive value of urban features, and seasonal and day/night variations in model results. The reviewed studies indicate that machine learning can represent complex relationships between air temperature and meteorological and spatial variables. GIS can bring support in this process by integrating data, extracting features, and mapping predictions. The choice of model and the resolution of the mapping depend on the research question being asked and the data available, and must match the station distribution, quality of input data, spatial scale, and local urban conditions. Future efforts should enhance observation coverage and data completeness, increase independent site validation, and test model transferability between cities.

Keywords: urban air temperature, spatial prediction, machine learning, geographic information systems, urban thermal environment

1. Introduction

Weather stations are not necessarily widely distributed enough to capture air-temperature distribution within a city. Measurements apply to specific locations and differ across neighboring communities with different building shapes, landscaping and waterfront. Estimating temperature over these regions necessitates a method that is appropriate to the observations and environmental data available.

Hassani et al. evaluated the performance of ordinary kriging, XGBoost and the Weather Research and Forecasting model in Warsaw, taking into account the available data and computing resource limitations [1]. Geographic information systems (GIS) can be used as a tool to correlate temperature data with the environment of each data point for machine learning purposes. Then building features, vegetation cover and terrain can be added as predictors. Yu et al. investigated the relationship

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.38271

between urban morphology and outdoor air temperature using random forests [2]. Jia et al. then applied this type of analysis to high-resolution mapping and the predictive role of built-environment attributes [3]. These applications must be evaluated in terms of how their predictions are tested. Wang et al. examined 45 studies and reported that there is no obvious link between spatial prediction error and grid size. They also found that there was a need for better reporting of data quality control and model evaluation [4]. Resolution, therefore, is only one of the factors that may be used to assess the usefulness of a published map. Observation coverage and validation procedures continue to play a key role in that judgment.

This paper reviews the relationship between these data needs and the application of GIS and machine learning in urban air temperature prediction. High-resolution mapping, analysis of spatial predictors and predictions under different temporal conditions are compared via literature synthesis and three published cases. A discussion is presented of the way observations lead to model inputs and mapped estimates, and how the evidence for predictions outside the sample space relates to the model.

2. Data and predictors

2.1. Meteorological observation data

Air temperature measurements are made to train and assess prediction models. The data from these measurements are in addition to the data collected from the fixed stations, which track changes at individual sites. Ru et al. employed automatic weather stations at central Taiyuan locations: Forest Park and Fenhe Park. The air temperature, relative humidity, wind speed and vapor pressure were measured every 10 minutes [5]. Another source of observations is mobile surveys. Zhou et al. recorded conditions along road routes in Wuhan and Zhengzhou using cars and bicycles fitted with temperature and humidity loggers and GPS [6]. Such surveys are more numerous but taken at different times along a route. Other meteorological variables like humidity, wind speed, solar radiation, and atmospheric pressure can also be used as predictors. Their effectiveness for model accuracy should be evaluated, rather than taken for granted. Another consideration is temporal vs spatial coverage. While street and neighborhood temperatures can be measured frequently, there is still a need for estimating those outside of the observation network.

2.2. Urban spatial and environmental predictors

Observation sites are described in terms of building density, building height, and impervious surface cover. These variables can be calculated using the spatial data prior to modeling the temperature. To create an urban digital elevation model, Yu et al. integrated building vector data with a digital elevation model (DEM). Within a 100 m grid, they computed mean and maximum heights, building density and green space ratio. Their study consisted of correlating the urban morphology and land surface temperature (LST) extracted from Landsat 8 [7]. This study shows an example of a GIS feature extraction, but this temperature relates to the surface, not the ambient air. More descriptors are needed for vegetation and water. The normalized difference vegetation index (NDVI) and vegetation cover represent vegetation; water coverage and distance to water indicate water features in a near vicinity. Ru et al. found that the mean air temperature in central Taiyuan was higher than that in Forest Park and Fenhe Park throughout June–August 2023. They attributed the vegetation-related cooling to transpiration and shading of their canopies, as well as to the heat capacity and evaporation of water [5]. Geographic setting is important when choosing predictors. In Oslo, Venter

DOI: 10.54254/2755-2721/2026.38271

et al. included elevation and distance to the coast to represent differences between observation locations [8]. These environmental variables are provided by remote sensing over broad regions. Landsat, Sentinel and MODIS data support information on land cover, vegetation and, depending on the sensor, surface temperature. LST is a different quantity to the near surface temperature, but can serve as a predictor. Guo et al. used MODIS LST, local climate zone data and ERA5 meteorological data to develop an air temperature estimation model based on random forest [9].

3. Methods for spatial prediction of urban air temperature

3.1. Spatial feature extraction using GIS

GIS connects the temperature data with the environmental information. The model has spatial predictors in the form of buildings, vegetation, water bodies, and terrain. Coordinate systems, spatial resolution and analysis units must be synchronized before merging these datasets. Observation times should also correspond to meteorological variables that vary with time. One study mapped the predictors on a 100 m grid and correlated the urban morphology and meteorological data with observation point and time [3].

These can be extracted using buffer analysis, raster statistics, and distance calculations. The local built environment can be characterized by building coverage and paved-surface cover within a certain distance from each station [2]. Distance to nearest water body is also a predictor [1]. These are the operations that map features into numerical inputs to the model.

The appropriate neighborhood size can differ among predictors. Li and Sharma extracted LiDAR-derived morphological features within a 15 m radius and land-use features within a 200 m radius in Chicago [10]. The radius used depends on the environmental feature being measured and the study setting. Using a single radius for each predictor could mask these differences.

3.2. Machine learning models for air temperature prediction

The observed air temperature is the response variable, and meteorological variables and GIS-derived features are the predictors. This task is performed using random forest (RF), support vector regression (SVR) and XGBoost. One study included urban form, reference-station air temperature, solar radiation, and wind speed in RF and SVR models [2]. The inputs enable the models to correlate the local temperature difference with weather and environment.

The prediction target varies between studies. An XGBoost application predicted the difference in temperature between a local site and a reference station, with the use of building and green space features as predictors [3]. In another study, Hassani et al. applied crowdsourced air temperature readings as the response variable. As predictors, their XGBoost model contained land surface temperature, vegetation indices, distance to water and ERA5-Land meteorological variables [1].

Roca-Barceló et al. estimated the daily mean air temperature within a 500 m grid in São Paulo. The study utilized weather-station data, remote sensing and re-analysis data. It contrasted RF with multiple linear regression using the same predictors. For tenfold cross-validation grouped by station, the root mean square error (RMSE) was 0.80 °C for RF and 1.02 °C for linear regression [11]. These values are for the data set and validation procedure used in the above study.

The coefficient of determination, root mean square error, and mean absolute error are typical metrics for model evaluation. These metrics are interpreted with the validation design. Li and Sharma used 29 of 30 stations for each run in their Chicago model. They withheld the complete record from the remaining station for validation [10]. This procedure tested predictions at a location

DOI: 10.54254/2755-2721/2026.38271

excluded from training. This spatial transfer should be evaluated in the case of mapping unobserved areas. It is also advisable to compare models with consistent target predictions, input parameters, and validation data, if applicable.

3.3. Spatial prediction and air temperature mapping

Mapping requires predictor values at locations where air temperature was not measured. The areas are typically shown in a regular grid. The spatial scale and definitions of the predictions must match those used for training. The meteorological inputs should be for the same time as that mapped. The trained model then predicts the air temperature of each cell.

The estimates are saved as an air temperature raster in GIS. The spatial resolution and temperature measure should be indicated for each map. The hourly temperature and the daily mean are two aspects of the thermal environment. The period of observation and the weather of that time must then be included as a part of the mapped results.

Good accuracy of prediction is not indicated by a fine grid alone. Regions where there are fewer observations could be under-represented in the training data. If there are missing or weak predictors, this adds more uncertainty. Validation results, station coverage, and availability of inputs should be used to evaluate the reliability of the map.

4. Typical applications of spatial prediction of urban air temperature

4.1. High-resolution urban air temperature mapping

Zumwald et al. mapped air temperature in Zurich, Switzerland, during the heatwave of June 25–30, 2019. Their data set included data from 634 citizen weather stations, 14 reference weather stations, and 43 environmental agency sensors. These observations were correlated with a quantile regression forest model that included reference station land use, elevation, building volume, NDVI, and meteorological variables. Several neighborhood scales were used for calculating spatial features, and the model was applied to a 10 m × 10 m grid [12]. The observations were from June 25–29 for training and June 30 for testing. The RMSE on that test day, averaged over all the sensor types, was 1.69 °C. In each of 10 repetitions 10% of the stations were held out randomly, yielding a mean RMSE of 1.86 °C [12].

The two error values refer to different test cases: error for prediction on another day, error for prediction of stations not used in training. Both help assess what the temperature maps can support. Additionally, Zumwald et al. calculated prediction uncertainty and marked areas where the model was extrapolating from the training data [12]. These outputs offer details of the reliability in addition to the detailed temperature patterns. Another restriction is the weather range of the study; tests were conducted during clear, hot weather, and this would need to be tested separately during rainy or cold weather.

4.2. Analysis of urban spatial factors and air temperature differences

Venter et al. analyzed quality-controlled observations from 1,310 private weather stations over Oslo in 2018. Random forest models were constructed using these records in conjunction with Landsat, Sentinel, LiDAR and topographic data. Spatial predictors were summarized within buffers around the stations. The researchers measured the increment in prediction error due to randomly permuting the values of a predictor in the validation data to evaluate variable importance [8]. LST, elevation

DOI: 10.54254/2755-2721/2026.38271

and distance to the coast were the most important variables in the aggregated importance results. The index-based built-up index (IBI), the standard deviation of building height and NDVI were also significant predicting variables. Two separate correlation analyses revealed an increase in annual mean air temperature as a function of LST and IBI and a decrease in annual mean air temperature as a function of NDVI, elevation, and distance-to-coast [8].

These rankings should be viewed in the context of the geographic location. The vegetation and built-environment features were substantially complemented by terrain and coastal position in Oslo. Moreover, the performance also depended on the extraction scale: buffer radii of 100–500 m provided more accurate estimates of air temperature annual mean in the present study's comparisons [8]. These results can be used to indicate useful predictors and applicable scales in this context. Variable importance is related to the contributions of a predictor to a fitted model, and correlation is related to association between variables. The results do not separately indicate the effect that a proposed change in vegetation, building or land cover would have on air temperature.

4.3. Urban air temperature prediction under different temporal conditions

Guo et al. examined seasonal and day/night differences in Szeged, Hungary. They matched air temperature observations from 24 weather stations with MODIS LST data for 56 clear-sky dates during 2014–2019, adding local climate zone (LCZ) information and ERA5 meteorological variables. Random forest analyses covered eight combinations of season and day/night conditions. The training data for the years 2014 to 2017 and the test data for 2018 to 2019 were used [9]. When test observations were matched with satellite overpass times, summer RMSE was 2.1 °C during the day and 2.2 °C at night. In the fall, the numbers were 4.5 °C and 3.5 °C. LST was the most important variable in both daytime and nighttime analyses in the reported annual mean variable importance results. LCZ's performance was the worst during the day and second worst at night. During the daytime, meteorological variables are more important, and specific humidity is the most important of these variables [9].

The results show that prediction errors and the contributions of individual predictors depend on the temporal conditions being examined. The reported accuracy becomes different when a different validation design is used. The mean daytime and nighttime RMSE values of 0.5 °C and 0.9 °C came from tenfold cross-validation within the training period [9]. They report performance using a different assessment method for the later years. Separate results help to avoid an overly optimistic view of the model's time performance. The scope of the study is further defined because it was carried out in clear skies and mapped at 1 km resolution; the performance outside these conditions was not determined by these tests.

5. Conclusion

The reviewed studies demonstrate the application of GIS and machine learning in the mapping of urban air temperature beyond observation sites. GIS offers spatial predictors from the environment. These predictors are coupled with meteorological data to estimate air temperature using machine learning. Estimates are only reliable given the coverage of observations, input data, and the design of the estimation process. These three cases show examples of the limits of this approach. For Zurich, the 10 m grid was used to generate detailed temperature maps. However, the uncertainty of predictions and the spatial extrapolation still need to be evaluated separately. In Oslo, elevation and distance from the sea were key predictors. Their significance is based on the geographical context and cannot be generalized to other cities. The prediction errors in Szeged fluctuated from one season

DOI: 10.54254/2755-2721/2026.38271

to another, and the importance of predictors ranged from day to night. Error estimates also varied between cross-validation within the training window and testing on later years. These results are of different types and must be differentiated. There is evidence available to support the temperature mapping conditions studied in each of these. Other cities, seasons or weather will require further testing before the approach can be reliably applied. Validation at stations and times not used during training is a priority. Further observations are also required where the existing network is not well suited to urban conditions. These steps will contribute to determining the level of reliability for temperature estimates used in the analysis of the urban thermal environment.

References

[1] Hassani, A., Santos, G. S., Schneider, P., & Castell, N. (2024). Interpolation, satellite-based machine learning, or meteorological simulation? A comparison analysis for spatio-temporal mapping of mesoscale urban air temperature. *Environmental Modeling & Assessment, 29*(2), 291–306. https://doi.org/10.1007/s10666-023-09943-9

[2] Yu, Z., Chen, S., Wong, N. H., Ignatius, M., Deng, J., He, Y., & Hii, D. J. C. (2020). Dependence between urban morphology and outdoor air temperature: A tropical campus study using random forests algorithm. *Sustainable Cities and Society, 61*, Article 102200. https://doi.org/10.1016/j.scs.2020.102200

[3] Jia, Q., Xie, J., Zeng, F., Han, D., Li, S., Zhang, M., Zhang, T., & Lu, L. (2026). High-resolution urban air temperature mapping and interpretable multi-scale built-environment analysis for climate adaptation in a severe cold-region city. *Building and Environment, 304*, Article 115134. https://doi.org/10.1016/j.buildenv.2026.115134

[4] Wang, H., Yang, J., Chen, G., Ren, C., & Zhang, J. (2023). Machine learning applications on air temperature prediction in the urban canopy layer: A critical review of 2011–2022. *Urban Climate, 49*, Article 101499. https://doi.org/10.1016/j.uclim.2023.101499

[5] Ru, H., Chen, Q., Jin, X., Zhang, X., Xun, J., & Lu, R. (2026). Regulatory effects of blue–green spaces on urban heat islands. *Journal of Chinese Urban Forestry, 24*(2), 51–59. https://doi.org/10.12169/zgcsly.2025.12.16.0002. (In Chinese)

[6] Zhou, X., Chen, H., Wu, Y., & Xu, H. (2018). Impacts of mobile survey-based urban spatial forms on summer afternoon thermal environments. *Landscape Architecture, 25*(10), 21–26. https://doi.org/10.14085/j.fjyl.2018.10.0021.06. (In Chinese)

[7] Yu, M., Li, J., Zhang, W., Chen, X., & Ren, J. (2022). Correlation analysis between urban three-dimensional morphology and the thermal environment, and platform development. *Geospatial Information, 20*(7), 77–82. (In Chinese)

[8] Venter, Z. S., Brousse, O., Esau, I., & Meier, F. (2020). Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sensing of Environment, 242*, Article 111791. https://doi.org/10.1016/j.rse.2020.111791

[9] Guo, Y., Unger, J., Khabibolla, A., Tian, G., He, R., Li, H., & Gál, T. (2024). Modeling urban air temperature using satellite-derived surface temperature, meteorological data, and local climate zone pattern—a case study in Szeged, Hungary. *Theoretical and Applied Climatology, 155*(5), 3841–3859. https://doi.org/10.1007/s00704-024-04852-7

[10] Li, P., & Sharma, A. (2024). Hyper-local temperature prediction using detailed urban climate informatics. *Journal of Advances in Modeling Earth Systems, 16*(3), Article e2023MS003943. https://doi.org/10.1029/2023MS003943

[11] Roca-Barceló, A., Schneider, R., Pirani, M., Sebastianelli, A., Piel, F. B., Vineis, P., Nardocci, A. C., & Fecht, D. (2026). A satellite-based machine learning approach for estimating high-resolution daily average air temperature in a megacity in Brazil. *Scientific Reports, 16*(1), Article 7459. https://doi.org/10.1038/s41598-026-35689-x

[12] Zumwald, M., Knüsel, B., Bresch, D. N., & Knutti, R. (2021). Mapping urban temperature using crowd-sensing data and machine learning. *Urban Climate, 35*, Article 100739. https://doi.org/10.1016/j.uclim.2020.100739