A Multi-Dimensional GeoXAI Framework for Explainable Spatial Modeling of Urban Heat Islands

Qasem Safariallahkheili[*,**]

[*]Lab for Geoinformatics and Geovisualization (g2lab), HafenCity University Hamburg, Henning-Voscherau-Platz 1, 20457 Hamburg, Germany.
[**]Interaction Design Lab, Potsdam University of Applied Sciences, Kiepenheuerallee 5, 14469 Potsdam, Germany.

Abstract

Rapid urbanization intensifies the Urban Heat Island (UHI) effect, posing significant health, energy, and environmental challenges. While machine learning (ML) can predict urban temperature patterns, its limited interpretability restricts its use in decision-making. This cases tudy presents a Geospatial eXplainable Artificial Intelligence (GeoXAI) framework for Berlin that integrates explainable ML with geospatial analysis. Using Landsat-derived Land Surface Temperature and multiple predictors, an XGBoost model achieved high accuracy ($R^2 = 0.854$, $RMSE = 1.61\,°\mathrm{C}$). Model interpretability is enabled through SHAP-based feature attributions, spatial autocorrelation, and uncertainty metrics, delivered via an interactive web-based geovisual analytics platform. The framework provides a structured approach to examining model outputs and their spatial context, with potential relevance for urban analysis and planning.
Keywords: GeoAI, explainable AI, Urban Heat Island

1. Introduction

Rapid urbanization has intensified the Urban Heat Island (UHI) effect, where built-up areas experience elevated temperatures compared to rural surroundings (Sun et al. 2025; Tian et al. 2021). This localized warming exacerbates heat-related health risks, energy demand, and environmental degradation, challenging sustainable urban planning (Gupta et al. 2025). While machine learning (ML) has advanced our capacity to predict urban heat patterns, its black-box nature limits interpretability and trust, which may hinder adoption in policy and spatial decision-making (Hassija et al. 2024). Addressing this gap, this case study develops an explainable GeoAI (GeoXAI) framework to model, interpret, and visualize the drivers of UHI in Berlin, Germany, by integrating advanced explainability mechanisms directly into geospatial ML workflows.

Published in “Proceedings of the 1st International Conference on Geospatial Artificial Intelligence (GeoAI 2026) – Poster Papers”, edited by Haosheng Huang and Nico Van de Weghe, GeoAI 2026, 3-6 June 2026, Ghent, Belgium.

This contribution underwent single-blind peer review based on the extended abstract.

The objective is to create a transparent, reproducible, interactive, and spatially aware GeoXAI pipeline that advances interpretability in urban environmental modeling. Specifically, the study aims to: (i) predict Land Surface Temperature (LST) using interpretable ML models; (ii) compute and integrate explainability metrics—feature importance, spatial autocorrelation, and uncertainty—within a coherent framework; and (iii) provide an interactive and web-based geovisual analytics environment that communicates both predictions and their reasoning. The work contributes to embedding explainability and spatial reasoning in GeoAI to support transparent and data-driven urban climate planning.

2. Methodology

2.1. Study Area and Data

The case study focuses on Berlin, an urban area with diverse morphology and rich open data availability. The target variable, LST, was derived from Landsat 8/9 thermal imagery (Avdan and Jovanovska 2016) for the summer of 2025, processed to 30 m resolution using standard correction and cloud masking. Here, UHI intensity is operationalized using LST as its spatial proxy. Predictor variables were harmonized to a common 30 m grid and include biophysical, morphological, and socio-economic factors (Deilami, Kamruzzaman, and Liu 2018; Sangiorgio, Fiorito, and Santamouris 2020), including *impervious surface area, building height and density, sky view factor, vegetation indices (NDVI, tree cover density), surface albedo, road and POI densities, slope, aspect, elevation, water bodies, and population density.*

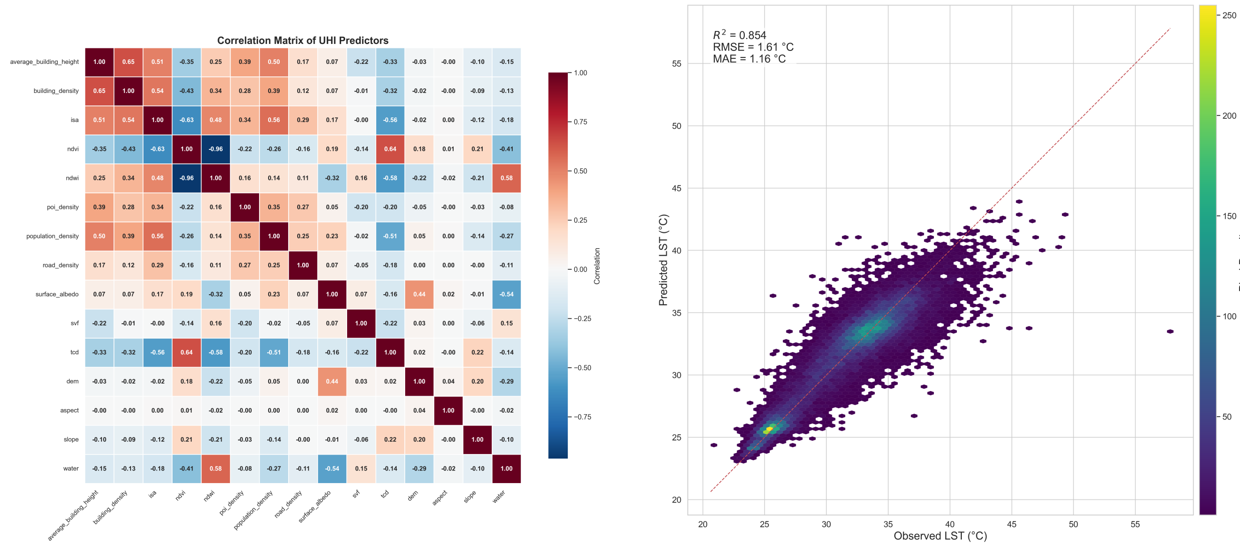
2.2. Modeling Approach

To predict UHI intensity and derive explainability metrics, a machine learning–based framework comprising data sampling, model evaluation, and model selection was employed. Predictor variables were selected based on prior literature and data availability, covering biophysical, morphological, and socio-economic dimensions. A stratified random sample of 100,000 pixels was drawn across the study area to achieve a balanced representation of temperature, and model performance was evaluated on held-out test data. Following a comparative analysis, Extreme Gradient Boosting (XGBoost) and Random Forest (RF) were evaluated as regression models. XGBoost slightly outperformed RF, achieving an R^2 of 0.854 and an RMSE of 1.61°C, compared to the RF performance ($R^2 = 0.842$, RMSE $= 1.74$°C). Due to its superior predictive accuracy and computational efficiency in generating SHAP (SHapley Additive exPlanations) values (Lundberg and Lee 2017), XGBoost was selected as the core engine for the subsequent GeoXAI analysis (Figure 1). The trained model was then applied to the dataset to predict UHI intensity and derive the explainability metrics.

2.3. Explainability Framework

The GeoXAI workflow integrates four complementary explainability dimensions:

- **Feature values:** raw environmental characteristics of each spatial unit.



(a) Correlation matrix of predictors

(b) Observed vs. predicted LST

Figure 1: Model diagnostics for the GeoXAI UHI analysis in Berlin using the XGBoost model

- **Feature importance (SHAP values):** local contributions of each predictor, revealing warming or cooling influences.
- **Spatial autocorrelation:** using Local Moran's I to assess spatial coherence of SHAP attributions and model residuals.
- **Prediction uncertainty:** ensemble-based estimates of model confidence across heterogeneous surface conditions.

Each dimension addresses a specific interpretive question—*what is the predicted temperature*, *how does the model arrive at this prediction based on local feature values*, *how does a location relate to its spatial surroundings or similar settings*, and *how confident is the model*—forming a progressive *WH-based sensemaking* structure that links raw prediction to actionable explanation.

3. Implementation and Results

3.1. Geovisual Analytics Implementation

The workflow was implemented within a tile-based architecture using *PostGIS* and vector tiles, supporting interactive multi-scale visualization of more than 2.6 million spatial records. To facilitate exploration, the interface includes an interactive dashboard that allows users to switch between different explainability dimensions and visualization modes. Multiple geovisual techniques were integrated:

- **Univariate choropleth maps** of predicted LST and individual explainability dimensions to provide an initial spatial overview.

- **Bivariate maps** linking feature magnitude with SHAP contributions to show key drivers of LST variation.

- **Spatial autocorrelation layers** illustrating statistically significant Local Moran's I clusters.

- **Uncertainty visualization** using textured overlays to indicate prediction reliability.

- **Ternary plots** combining prediction, feature value, and SHAP attribution in a single compositional space (Figure 2).

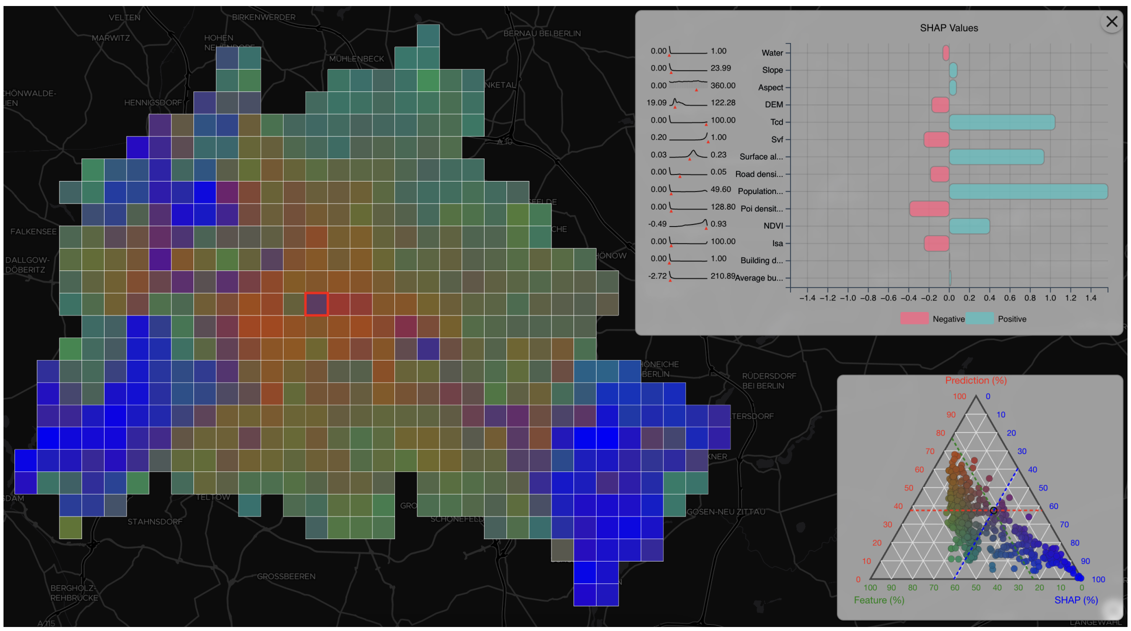


Figure 2: Ternary plot co-visualizing predicted LST, Surface albedo, and its SHAP contribution on the web application.

This visual integration allows users to explore not only where hotspots occur, but also why the model attributes temperature changes to specific drivers and where these relationships hold spatially.

3.2. Preliminary Results

The preliminary outcome of this case study is a web-based GeoXAI application that visualizes explainability dimensions of the UHI model through a GIS-integrated interface. The platform allows users to explore spatial layers representing predicted Land Surface Temperature, SHAP-based feature attributions, uncertainty, and spatial autocorrelation interactively. Through dynamic maps, users can toggle between features, adjust classification thresholds, and visualize univariate, bivariate, or ternary relationships between model outputs and explanatory factors, supporting exploration across multiple spatial scales.

The ternary color encoding links model prediction (red), feature magnitude (green), and SHAP contribution (blue), allowing composite interpretation of "what–why" relationships in UHI dynamics. The user interface is designed for both experts and policymakers, fostering transparency, participatory analysis, and reproducible evaluation of AI-driven spatial decisions.

This prototype demonstrates how explainability dimensions can be operationalized within a GIS environment, transforming abstract ML outputs into interpretable, spatially coherent, and decision-relevant insights.

4. Conclusion and Future Work

This case study integrates established ML models (XGBoost) and explainability metrics (SHAP) into a web-based geovisual analytics platform, enabling spatially explicit interpretation of UHI drivers. The approach is limited by the reliance on a single temporal snapshot and the availability of input data, which may affect generalizability.

The GeoXAI framework extends explainable AI into geospatial modeling by coupling spatial structure, interpretability, and visualization. Methodologically, it (i) introduces a multi-dimensional design integrating SHAP, uncertainty, and spatial autocorrelation; (ii) provides a scalable, reproducible visualization environment to support human sensemaking; and (iii) demonstrates the potential of open, interactive, and transparent GeoAI workflows using open-source tools.

Future work will explore advanced spatially explicit models, conduct detailed analyses of Local Moran's I and uncertainty visualizations, and evaluate the platform through structured user studies to assess interpretability and user trust. Additional directions include investigating temporal dynamics, cross-city transferability, and applications to other domains such as flood risk or air quality. Overall, this approach supports a *transparent, spatially grounded GeoAI*, moving from prediction to understanding and enabling responsible AI for sustainable urban planning.

References

Avdan, Ugur, and Gordana Jovanovska. 2016. "Algorithm for automated mapping of land surface temperature using LANDSAT 8 satellite data." *Journal of sensors* 2016 (1): 1480307.

Deilami, Kaveh, Md Kamruzzaman, and Yan Liu. 2018. "Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures." *International journal of applied earth observation and geoinformation* 67:30–42.

Gupta, Aman, Bhaskar De, Sutapa Das, and Mahua Mukherjee. 2025. "Thermal hazards in urban spaces: A review of climate-resilient planning and design to reduce the heat stress." *Urban Climate* 59:102296.

Hassija, Vikas, Vinay Chamola, Atmesh Mahapatra, Abhinandan Singal, Divyansh Goel, Kaizhu Huang, Simone Scardapane, Indro Spinelli, Mufti Mahmud, and Amir Hussain. 2024. "Interpreting black-box models: a review on explainable artificial intelligence." *Cognitive Computation* 16 (1): 45–74.

Lundberg, Scott M, and Su-In Lee. 2017. "A unified approach to interpreting model predictions." *Advances in neural information processing systems* 30.

Sangiorgio, Valentino, Francesco Fiorito, and Mattheos Santamouris. 2020. “Development of a holistic urban heat island evaluation methodology.” *Scientific reports* 10 (1): 17913.

Sun, Jian, Zezhuang Liu, Fan Xia, Yilu Gu, Xiaofeng Gao, Sha Lu, Yang Xu, Feidan Meng, Qian Zhang, and Tao Zhou. 2025. “Uncovering the Impacts of 2D and 3D Urbanization on Urban Heat Islands in 384 Chinese Cities.” *Environmental Science & Technology* 59 (14): 7106–7116.

Tian, Peng, Jialin Li, Luodan Cao, Ruiliang Pu, Zhongyi Wang, Haitao Zhang, Huilin Chen, and Hongbo Gong. 2021. “Assessing spatiotemporal characteristics of urban heat islands from the perspective of an urban expansion and green infrastructure.” *Sustainable Cities and Society* 74:103208.