Quantifying Urban Heat Island Trends in the Appalachian Region Using Machine Learning and Geographically Weighted Regression

Shashank Karki
Virginia Tech
USA
shashankkarki@vt.edu

Nitant Rai
Virginia Tech
USA
nitant3@vt.edu

Sonia Sharma
Virginia Tech
USA
soniasharma07@vt.edu

Huy Pham
Virginia Tech
USA
phamanhuy@vt.edu

Abstract

Urban Heat Islands (UHI)—the phenomenon where urban areas experience significantly higher temperatures than surrounding rural regions—represent a critical environmental and public health challenge, especially in rapidly urbanizing areas. The Appalachian region, despite its unique ecological, economic, and historical significance, remains notably understudied in large-scale UHI modeling efforts. To address this research gap, this study develops a novel, high-resolution, spatially enriched dataset spanning 13 Appalachian states, integrating satellite remote sensing data, topographic variables, infrastructural characteristics, and demographic attributes. Employing this comprehensive dataset, the study evaluates two prominent machine learning (ML) methods—Gradient Boosting (GB) and Random Forest (RF)—alongside Geographically Weighted Regression (GWR), a spatially explicit non-ML approach, to quantify and interpret UHI intensity. Comparative analyses reveal that RF slightly outperforms GB in predictive accuracy (RMSE: RF = 2.579, GB = 2.594; R^2: RF = 0.5531, GB = 0.5475). While GB and RF effectively capture broad-scale trends, their global modeling frameworks obscure critical local spatial variations. Conversely, GWR, despite moderate predictive performance (RMSE = 2.29, R^2 = 0.65), significantly enhances interpretability by highlighting spatially heterogeneous relationships between UHII and predictors such as vegetation, population density, and land surface temperature. This research underscores the importance of combining predictive strength with spatial interpretability, demonstrating the value of hybrid modeling frameworks. Such approaches can better inform targeted mitigation and adaptation strategies, crucial for enhancing environmental resilience in complex, diverse regions like Appalachia.

Keywords: Urban Heat Island, Appalachian Region, Machine Learning, Geographically Weighted Regression.

1 Introduction

Urban Heat Islands (UHI)—the phenomenon where urban areas experience significantly higher temperatures than their surrounding rural regions [1]—pose a major environmental challenge, particularly in regions undergoing fast urbanization and land cover changes. UHI contributes to climate change by increasing energy demand for cooling, intensifying heat waves, and pollution. These effects disproportionately impact vulnerable populations, heighten the risk of heat-related illnesses, and disrupt local ecosystems.

The Appalachian region is a critical area of interest due to its unique ecological, economic, and historical significance. Known for its rich biodiversity, the region hosts one of the most diverse temperate forests in the world, supporting a wide range of plant and animal species. Economically, it has long been a major hub for coal mining and timber production, both of which have played a crucial role in shaping the national economy [2]. However, recent suburban expansion has led to land-use and land cover (LULC) changes that exceed the national average [3], contributing to more frequent temperature extremes and intensifying UHI effects [4].

Accurate prediction and quantification of UHI trends in this region are essential for informing urban planners about potential changes and supporting the development of effective mitigation strategies. While various approaches—including integrating geospatial data with machine learning (ML)—have been explored, their application to long-term UHI analysis in the Appalachian region remains limited. To address this research gap, this study develops a novel high-resolution spatial dataset for the region and evaluates two widely used ML models—Gradient Boosting (GB) and Random Forest (RF)—alongside Geographically Weighted Regression (GWR), a spatially explicit regression method. By assessing both predictive performance and spatial interpretability, this research aims to demonstrate the complementary strengths of global ML models and local regression techniques, ultimately providing a hybrid modeling framework suitable for region-specific policy and planning decisions.

2 Related Work

The intensity of urban heat effects varies according to geographic and climatic conditions, with mountainous regions being uniquely affected due to their complex topography and vegetation patterns [5]. UHI is strongly influenced by land cover change, as urban expansion increases heat absorption and elevates land surface temperature (LST) [6]. The replacement of forests and agricultural land with impervious surfaces further intensifies UHI by reducing evapotranspiration and disrupting local microclimates. This effect is especially evident in rapidly urbanizing areas, where green spaces are often converted into transportation infrastructure, amplifying heat accumulation [6].

One such mountainous region significantly affected by these dynamics is the Appalachian region, which extends across 13 states from southern New York to northern Mississippi [7]. However, parts of this region have experienced rapid urbanization, surpassing the national average in land cover conversion [3], increasing impervious surfaces, and intensifying LST [5]. Despite these trends, limited research has applied geospatial and ML techniques to quantify LST variations and understand long-term UHI patterns in the region.

Recent advancements in ML and remote sensing have enhanced the ability to quantify UHI with greater accuracy. For example, artificial neural networks were trained using LULC, LST, and the Urban Thermal Field Variance Index to assess seasonal variations in UHI effects [8]. Deep learning models incorporating green density, urbanization density, and building density have been used to predict regional temperatures in Daegu City [9]. Another study introduced the U-HEAT framework, integrating convolutional neural networks and remote sensing to analyze the spatial and temporal distribution of UHI while assessing community vulnerability to extreme heat [10]. Classical ML algorithms such as random forest were applied to identify key urban characteristics influencing UHI formation, with impervious surface cover and the Normalized Difference Vegetation Index (NDVI) emerging as the most significant factors [11]. Another study used LST, LULC, impervious surfaces, and vegetation indices to train Bagging (to reduce variance) and Random Subspace (to handle feature redundancy) models, which effectively predicted UHI in Jeddah, with statistical analysis (R^2 = 0.763) confirming strong LST-UHI correlation and the cooling effect of water bodies [12].

3 Methodology

This study compares the performance of two classical ML models—Gradient Boosting (GB) and Random Forest (RF)—with Geographically Weighted Regression (GWR), a non-machine learning approach, to evaluate their effectiveness in predicting UHI in the Appalachian region. The dataset consists of spatially referenced data points, which are first extracted using Google Earth Engine (GEE) and then processed in a Geographic Information System (ArcGIS Pro) for data preparation. This workflow supports high-resolution spatial data analysis across the diverse landscape in a large scale of the Appalachian region. Details of the models and the dataset are presented in the following subsections.

3.1 Data Collection

The data collection phase is one of the most time-intensive and critical components of this study, forming the foundation for all subsequent modeling efforts. Due to the lack of a unified, high-resolution, spatially enriched dataset for the Appalachian region, a custom geospatial data pipeline is designed and implemented that integrates satellite, topographic, infrastructural, and demographic variables.

To ensure comprehensive spatial coverage, uniformly spaced grid points at 5 km intervals across the Appalachian region are first generated. The grid is clipped to the Appalachian Regional Commission (ARC) boundary shapefile, capturing the diverse landscapes and administrative divisions spanning 13 states. This resulted in approximately 230,000 grid points across urban, semi-urban, and rural environments.

For each grid point, a suite of remote sensing variables using Landsat 8 imagery from the year 2014 via GEE is extracted. These includes LST, NDVI, Normalized Difference Built-up Index (NDBI), Bare Soil Index, and an Urban Heat Island Index (UHII) index derived from LST differentials. Proximity to water bodies and roads is computed using vector-based distance calculations from respective polygon layers, not directly from satellite images. Similarly, road proximity is calculated using road networks from the U.S. Census Bureau's TIGER/Line shapefiles, allowing for the quantification of infrastructural intensity across space.

Topographic features, including elevation, slope, and aspect, are derived from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM). Demographic and socioeconomic attributes are added via spatial joins with American Community Survey (ACS) 5-year estimates from 2014 at the census block group level. Variables include population density, household count, and median household income. Each grid point is assigned the values of the polygon that it spatially intersects with, ensuring accurate local representation of demographic context.

Collecting demographic data across the Appalachian region requires considerable manual effort. Due to the region comprising multiple states and the absence of a consolidated ACS shapefile for Appalachia, demographic attributes have to be separately downloaded and processed for each individual state before merging and spatially aligning them with the grid points. This systematic geospatial data preparation results in a rich, high-resolution dataset with over 30 features per point, supporting robust and context-sensitive modeling. It also underscores a core challenge in ML for geographic applications: the data engineering phase often dominates the project timeline and complexity, particularly when working in fragmented, data-scarce, or topographically complex regions like Appalachia.

3.2 Data Preprocessing

Two main issues are identified in the raw dataset: missing values and a small number of extreme outliers. The missing values primarily results from cloud cover in the satellite imagery, which obscures ground surfaces and prevents accurate readings by remote sensing instruments. To address this, a moving average filter is applied using a window of 10 neighboring points. This method smooths the data while preserving localized spatial patterns, ensuring the integrity of the affected features.

The second issue involves extreme outlier values present in a small subset of the data. These anomalies stemmed from numerical artifacts introduced during shapefile-based spatial joints, particularly in cases where large administrative boundaries are mistakenly assigned unrealistic values at certain points. For instance, a small grid cell might inherit an exaggerated attribute from an unusually large or misaligned polygon. To mitigate this, the top and bottom 1% of values across affected features are removed. This trimming strategy eliminates the most severe outliers without distorting the broader distribution. After addressing these issues, the final cleaned dataset provides a robust basis for model development.

3.3 Algorithm Selection

3.3.1 Gradient Boosting (GB). Gradient Boosting (GB) is an ensemble method that builds models sequentially, with each new weak learner trained to correct the errors of its predecessor. Decision trees are used as the base learners in GB because they can handle both categorical and numerical data, capture non-linear relationships, and operate effectively without requiring feature scaling. The algorithm starts with an initial prediction, typically the mean of the target values in regression tasks, and then iteratively fits shallow decision trees to the residuals—the differences between the actual and predicted values. Each tree contributes to improving the overall model by minimizing these residuals. The core idea of GB is to reduce model bias by progressively refining predictions to represent the underlying data patterns better.

3.3.2 Random Forest (RF). Random Forest (RF) is a widely used ensemble learning algorithm that constructs multiple decision trees using randomly selected subsets of data and features. Each tree independently predicts the target, and the final prediction is made by averaging their outputs, reducing variance and improving generalization. RF is particularly effective in handling high-dimensional, noisy, and nonlinear datasets—qualities that are common in UHI modeling due to the complex interactions between land cover, topography, and socioeconomic variables. Its parallel training process enhances efficiency, while built-in feature importance metrics offer insights into which factors most strongly influence Urban Heat Island Intensity (UHII), such as NDVI, population density, and proximity to roads.

3.3.3 Geographically Weighted Regression (GWR). Geographically Weighted Regression (GWR) is a spatial modeling technique that extends traditional linear regression by allowing model parameters to vary across geographic space. Unlike standard regression models, which produce a single set of global coefficients, GWR fits a localized model at each observation point by weighting nearby data more heavily than distant points based on geographic proximity.

The method operates on the principle that spatial processes are not uniform and that relationships between predictors and the response variable can change over space. For each location, GWR computes a unique set of regression coefficients by applying a spatial kernel function—typically Gaussian or bisquare—that assigns higher weights to nearby data points. The model is calibrated using these locally weighted subsets, enabling it to reveal spatial heterogeneity in variable influence.

The primary advantage of GWR lies in its ability to provide geographically interpretable results. Instead of generating a single average effect for each feature, it produces spatially varying coefficient surfaces that highlight where and how strongly each variable affects UHII. This makes it a powerful tool for identifying localized drivers of heat intensity and for supporting geographically targeted planning and mitigation strategies.

While GWR does not typically match the predictive performance of advanced ML models, its interpretability and spatial sensitivity make it a valuable complement in hybrid modeling workflows—especially in regions like Appalachia where environmental and socioeconomic conditions are spatially complex and heterogeneous.

3.4 Experimental Setup

The post-processed dataset, containing approximately 230,000 data points, is split in two ways to accommodate the different modeling approaches. For GB and RF, the data is randomly split into 70% for training, 15% for validation, and 15% for testing. To prevent overfitting and enable robust hyperparameter tuning, 3-fold cross-validation is applied to the training set. In parallel, a second split—also into 70% for training, 15% for validation, and 15% for testing—is created without shuffling to preserve spatial continuity. This unshuffled split is used exclusively for GWR, which relies on the spatial dependency of data points and would be invalidated by random sampling.

Since hyperparameter tuning involves evaluating multiple combinations, the cross-validation process is repeated multiple times. For example, in the case of GB, testing 3 learning rates (0.01, 0.05, and 0.1), 3 tree depths (from 2 to 4), and 3 values for the number of trees (100, 200, and 300) results in 27 model configurations—each undergoing 3-fold validation, making the process computationally intensive. Given these constraints and the goal of predicting general UHI trends across a large region (206,000 square miles), a downsampling strategy is adopted. By retaining every 10th data point, the dataset is reduced to approximately 23,000 examples, making the modeling process more tractable without compromising the spatial diversity of the dataset.

3.4.1 Gradient Boosting (GB). For hyperparameter tuning, key parameters influencing the performance of GB with decision trees as base learners include the learning rate, tree depth, and number of trees. The learning rate controls the step size at which the model updates its predictions, helping to balance learning speed and generalization. Three learning rates are implemented: 0.01, 0.05, and 0.1. To align with the core idea of boosting—improving performance by combining weak learners—shallow decision trees are used, with maximum depths of 2, 3, and 4. The number of trees, which determines the total number of boosting rounds, is set to 100, 200, and 300. Combining these parameter combinations with 3-fold cross-validation results in a total of 81 combinations, each trained on a downsampled dataset of approximately 23,000 examples and the best-performing model is selected based on validation mean squared error (MSE).

3.4.2 Random Forest (RF). To optimize the performance of Random Forest (RF), a grid search is conducted over key hyperparameters: number of trees (100, 200, 300), maximum tree depth (5, 10, 15), and feature selection strategies (sqrt, log2). A total of 27 model configurations are evaluated using 3-fold cross-validation. The best-performing model is selected based on validation MSE. Evaluation on the test set yields a strong predictive performance with low error and high R^2. Additionally, learning curves are generated to analyze how prediction error changes with increasing tree count, and feature importance is assessed to interpret which input variables most influenced UHII prediction.

3.4.3 Geographically Weighted Regression (GWR). For GWR, we used the unshuffled spatial split to preserve geographic locality during both model training and validation. GWR requires selecting an optimal kernel bandwidth, which controls how spatially localized each regression becomes. A Gaussian kernel function is employed, and the optimal bandwidth is determined using iterative validation over a range of values.

Unlike ML models, GWR does not involve classical hyperparameters like tree depth or learning rate. Instead, the spatial weighting scheme and kernel bandwidth are the primary controls. Once the optimal bandwidth is selected, the model is trained on the full spatial training set, and predictions are generated for the test set to evaluate localized UHII estimates. This setup allows GWR to uncover spatially varying relationships between predictors (e.g., NDVI, NDBI, elevation) and urban heat intensity—particularly critical in a region as topographically and socioeconomically diverse as Appalachia.

4 Results

4.1 Gradient Boosting (GB)

MSE is used as the primary metric for selecting the optimal hyperparameters in the GB model. Specifically, the best values for tree depth and learning rate are determined based on validation performance. As shown in Figure 1, the combination of maximum depth = 4 and learning rate = 0.1 produces the lowest validation MSE, indicating the best-performing configuration among the options tested. The figure also demonstrates the model's sensitivity to the learning rate: while varying tree depth leads to moderate differences in MSE, changes in learning rate have a more substantial impact on performance.

In addition to tuning depth and learning rate, the total number of trees (i.e., boosting iterations) is evaluated using the training and validation loss curves shown in Figure 2. A sharp decline in both training and validation MSE is observed within the first 30 trees. After this point, the training loss continues to decrease, but the validation loss levels off around 150 trees, suggesting the onset of overfitting. This divergence between training and validation performance is a typical indicator of reduced generalization, and therefore, 150 trees is selected as the optimal number of boosting rounds.

Based on these findings, the final GB model configuration is set to learning rate = 0.1, maximum depth = 4, and number of trees = 150. This model is then evaluated on the test set, with its predictive performance illustrated in Figure 3. The scatter plot of predicted versus actual Urban Heat Island Index (UHII) values shows a strong alignment along the 45-degree reference line, indicating accurate predictions and strong generalization to unseen data.

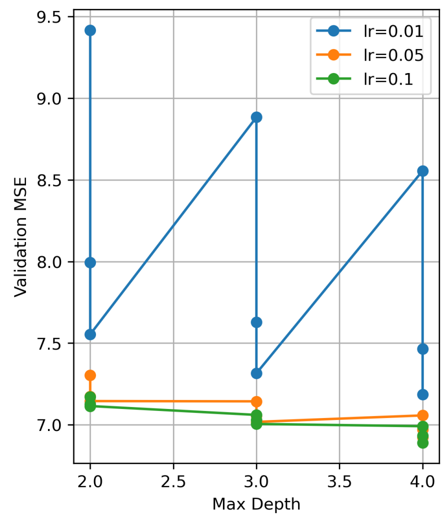


Figure 1: Validation MSE vs max depth for different learning rates (lr) - Gradient Boosting (GB).

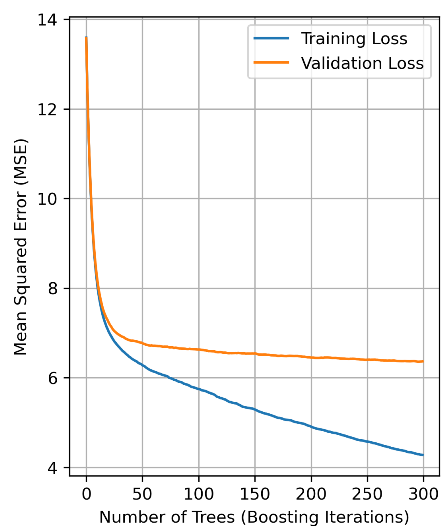


Figure 2: Training and Validation Loss vs Number of Trees - Gradient Boosting (GB).

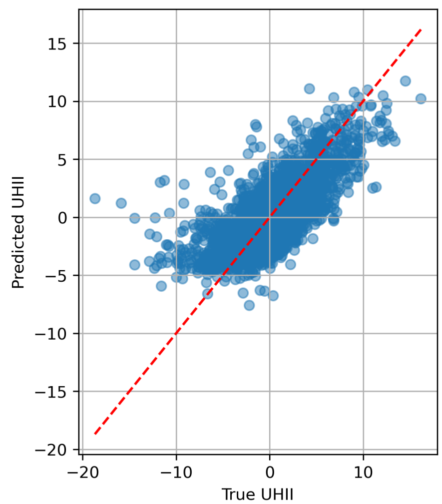


Figure 3: True vs Predicted UHII - Gradient Boosting (GB).

4.2 Random Forest (RF)

RF iss tuned and evaluated using validation MSE as the primary performance metric. Hyperparameter optimization focused on two key parameters: maximum tree depth and the number of features considered at each split (‘max_features‘). As illustrated in Figure 4, validation MSE decreases consistently with increasing tree depth for both `sqrt` and `log2` settings of ‘max_features‘. Among the tested configurations, the lowest validation error is achieved using `max_features = sqrt` and a maximum depth of 15, indicating this combination provided the best predictive performance with minimal overfitting.

In addition to depth and feature selection, the effect of the number of trees (estimators) is examined using the training and validation loss curves shown in Figure 5. The training loss drops sharply within the first 30 trees and gradually plateaus as more trees are added. Similarly, the validation loss shows similar trend to the training loss, suggesting diminishing returns from adding more estimators. The widening gap between training and validation loss beyond this point also indicates mild overfitting. Therefore, a model with approximately 100 trees is selected to balance performance and generalization.

The predictive accuracy of the final RF model is evaluated on the test dataset, as shown in Figure 6. The scatter plot of true versus predicted UHII values shows a strong alignment along the 45-degree reference line, reflecting the model’s ability to generalize well to unseen data. While slight underestimation is observed at higher UHII values, the overall distribution suggests reliable predictive performance. Based on these findings, the final Random Forest configuration is set to a maximum depth of 15, `max_features = sqrt`, and 100 estimators. This configuration delivers robust and accurate UHII predictions with minimal overfitting.

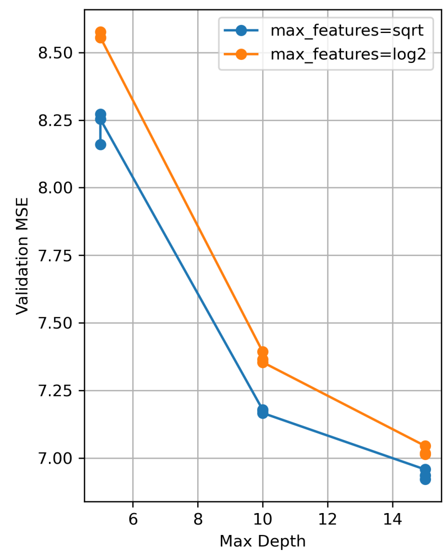


Figure 4: Validation MSE vs max depth for different max features - Random Forest (RF).

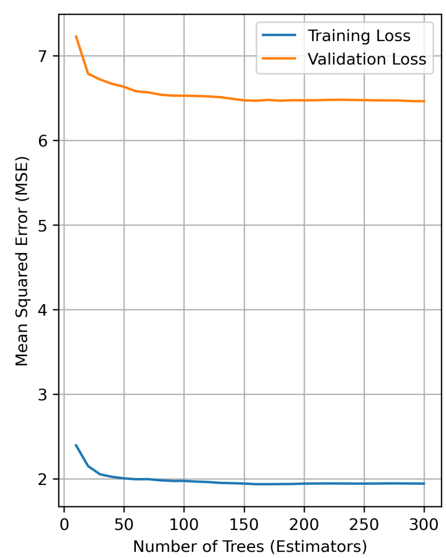


Figure 5: Training and Validation Loss vs Number of Trees - Random Forest (RF).

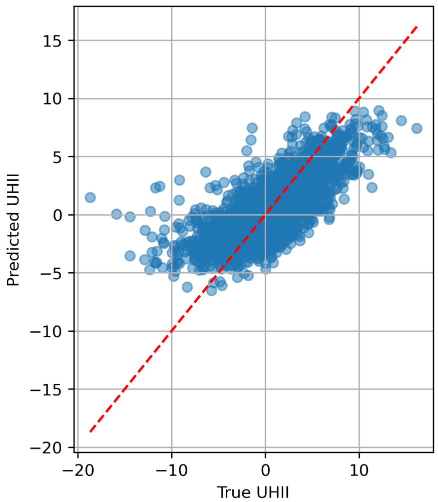


Figure 6: True vs Predicted UHII - Random Forest (RF).

4.3 Geographically Weighted Regression (GWR)

GWR provides essential insight into the spatially varying relationships between environmental and demographic features and UHII. While the model's predictive performance is moderate—with a test RMSE of 2.29 and an R^2 of 0.65—it effectively captures critical spatial nuances across the Appalachian landscape that global models often miss.

To optimize GWR's predictive accuracy and spatial interpretability, bandwidth values ranging from 50 to 300 are systematically evaluated using validation MSE. As shown in Figure 7 (bottom), validation MSE decreases sharply with increasing bandwidth up to approximately 150, then slightly increased again. The optimal bandwidth is identified as 149, minimizing the validation MSE to approximately 7.2. Selecting this optimal bandwidth is essential to balance spatial smoothing and local variability: smaller bandwidths introduced noisy and unstable local estimates, while larger values risk masking important spatial variations. Figure 7 illustrates the model's strong predictive alignment, with predicted UHII values closely clustering around the 1:1 reference line. This close alignment underscores the model's ability to capture local spatial variations and confirms the appropriateness of the selected bandwidth for capturing complex spatial relationships without overfitting.

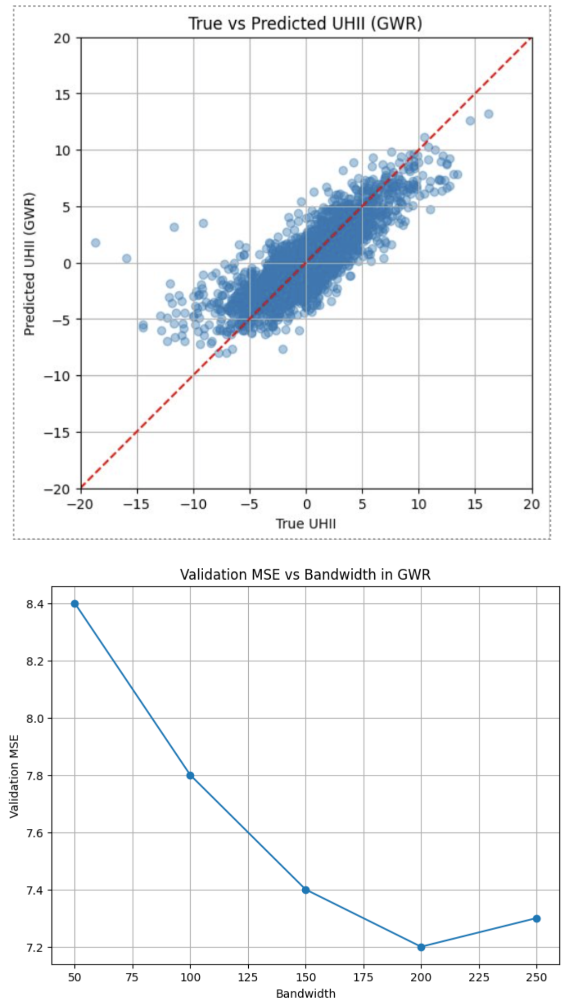


Figure 7: (Top) True vs Predicted UHII from GWR. (Bottom) Validation MSE across bandwidths. The optimal bandwidth is 149, minimizing error to 7.2.

5 Discussion

This study contributes to the growing literature on UHI modeling by addressing a critical geographic gap: the Appalachian region, which has been historically underrepresented in large-scale UHI research. By developing a high-resolution, spatially enriched dataset across 13 Appalachian states, the first regional-scale UHI modeling framework for this ecologically and topographically complex area is provided.

The predictive performances of Gradient Boosting (GB) and Random Forest (RF) are closely comparable, with RF exhibiting slightly better accuracy. Specifically, RF achieved a marginally lower root mean squared error (RMSE) of 2.579 and a slightly higher R^2 of 0.5531 compared to GB, which produced an RMSE of 2.594 and an R^2 of 0.5475. The similarity in performance between these two ensemble approaches suggests that both models effectively capture general trends in Urban Heat Island Intensity (UHII) across the Appalachian region. However, RF's slightly superior performance may be attributed to its intrinsic ability to reduce variance through averaging predictions from independently grown trees, while GB's sequential fitting process can lead to minor overfitting despite careful tuning.

While GB and RF demonstrate strong predictive capability, their global modeling frameworks inherently overlook localized spatial variability, limiting their interpretive value for policy and planning applications. Geographically Weighted Regression (GWR), despite its moderate predictive performance (RMSE = 2.29 and R^2 = 0.65), provides critical additional insight by revealing spatially heterogeneous relationships that are typically hidden in global ML approaches. For example, GWR identified distinct patterns such as the stronger cooling effect of vegetation (NDVI) in forested, high-elevation areas, and more pronounced impacts of Land Surface Temperature (LST) and population density in urbanized valleys. Such spatial insights significantly enhance the interpretability and practical utility of GWR compared to GB and RF alone. Integrating local spatial regression techniques with machine learning offers a more balanced approach—combining strong predictive performance with localized interpretability to better support targeted mitigation and adaptation strategies.

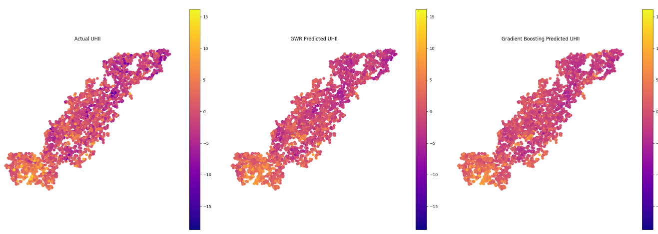
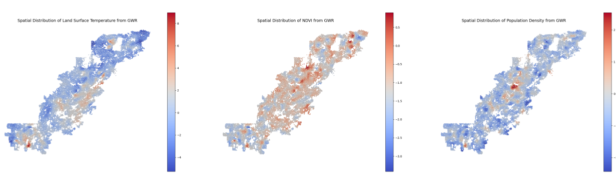


Figure 8: UHII comparison: Actual UHII (left), GWR Predicted UHII (center), and Gradient Boosting Predicted UHII (right).



(a) NDVI **(b) Population Density** **(c) LST**

Figure 9: Spatial distribution of GWR coefficients: (a) NDVI, (b) Population Density, and (c) Land Surface Temperature.

These conclusions are further reinforced through visual analyses of spatial predictions and local feature relationships (Figures 8 and 9). While both GWR and GB capture general UHII spatial trends effectively (Figure 8), GWR distinctly reveals nuanced local variations, particularly in areas of complex terrain or mixed land-use patterns. The spatial coefficient maps from GWR (Figure 9) clearly illustrate these variations: NDVI shows stronger negative associations in forested and higher-elevation zones, confirming vegetation's cooling influence; population density displays localized hotspots of strong positive associations around urban clusters in central Appalachia; and LST coefficients exhibit pronounced positive effects in valleys and lowland areas where heat entrapment is common. Together, these findings underline the unique value of spatially explicit modeling methods for understanding local environmental dynamics and informing targeted urban heat mitigation strategies.

6 Conclusion

This study addressed a critical geographic gap in Urban Heat Island (UHI) research by developing the first comprehensive, regional-scale modeling framework tailored specifically for the ecologically diverse and topographically complex Appalachian region. Leveraging a novel high-resolution, spatially enriched dataset spanning 13 states, this research evaluated the predictive performance and interpretive value of Gradient Boosting (GB), Random Forest (RF), and Geographically Weighted Regression (GWR). The comparative analysis showed that RF and GB provided similar predictive accuracy, with RF performing slightly better (RMSE = 2.579, R^2 = 0.5531) compared to GB (RMSE = 2.594, R^2 = 0.5475). Despite the strong predictive capabilities of these global machine learning models, they inherently overlook localized spatial variations critical for targeted policy and planning decisions. Conversely, GWR offered moderate predictive performance (RMSE = 2.29, R^2 = 0.65) but significantly enhanced interpretability by uncovering spatially heterogeneous relationships between key predictors and UHII.

Ultimately, this research illustrates the considerable value of integrating spatially explicit regression techniques with machine learning approaches, balancing predictive performance with local interpretability. Such hybrid frameworks not only improve scientific understanding of urban heat dynamics but also support practical, targeted decision-making efforts crucial for enhancing environmental resilience in the Appalachian region and similarly complex landscapes elsewhere.

References

[1] Li Yang, Feng Qian, De-Xuan Song, and Ke-Jia Zheng. Research on urban heat-island effect. *Procedia engineering*, 169:11–18, 2016.
[2] Phillip J Obermiller. Thoughts on the importance of identifying appalachians. *Appalachian Journal*, 38(1), 2010.
[3] Robert Oliver and Valerie Thomas. Land conversation in appalachia: Urban growth and change dynamics. 2017.
[4] Evan Kutta and Jason Hubbart. Climatic trends of west virginia: A representative appalachian microcosm. water 11 (6): 1117, 2019.
[5] Chunxia Liu and Yuechen Li. Spatio-temporal features of urban heat island and its relationship with land use/cover in mountainous city: A case study in chongqing. *Sustainability*, 10(6):1943, 2018.
[6] Ronald C Estoque and Yuji Murayama. Monitoring surface urban heat island formation in a tropical mountain city using landsat data (1987–2015). *ISPRS Journal of Photogrammetry and Remote Sensing*, 133:18–29, 2017.
[7] Harold R Wanless, ED McKee, and EJ Crosby. Appalachian region. *Geological Survey Professional Paper*, (853):13, 1975.
[8] Abdulla-Al Kafy, Milan Saha, Zullyadini A Rahaman, Muhammad Tauhidur Rahman, Desheng Liu, Md Abdul Fattah, Abdullah Al Rakib, Ahmad E AlDousari, Sk Nafiz Rahaman, Md Zakaria Hasan, et al. Predicting the impacts of land use/land cover changes on seasonal urban thermal characteristics using machine learning algorithms. *Building and Environment*, 217:109066, 2022.
[9] Yesuel Kim and Youngchul Kim. Deep learning-based analytic framework using comprehensive urbanization index for heat vulnerability assessment in urban areas. *Expert Systems with Applications*, 235:121140, 2024.
[10] Fei Li, Tan Yigitcanlar, Madhav Nepal, Kien Nguyen Thanh, and Fatih Dur. A novel urban heat vulnerability analysis: Integrating machine learning and remote sensing for enhanced insights. *Remote Sensing*, 16(16):3032, 2024.
[11] Sanglim Yoo. Investigating important urban characteristics in the formation of urban heat islands: a machine learning approach. *Journal of Big Data*, 5(1):2, 2018.
[12] Abdullah Addas. Machine learning techniques to map the impact of urban heat island: investigating the city of jeddah. *Land*, 12(6):1159, 2023.