

Article

Forecasting Urban Heat Island Intensification in Arkansas, USA, Using the XGBoost Machine Learning

Rasool Vahid ^{1,*}  and Mohamed H. Aly ² ¹ Environmental Dynamics Program, University of Arkansas, Fayetteville, AR 72701, USA² Arkansas Center for Space & Planetary Sciences, University of Arkansas, Fayetteville, AR 72701, USA

* Correspondence: rvahidbe@uark.edu

Abstract

Urban heat islands (UHIs) significantly influence microclimatic conditions, energy consumption, and public health. This research leverages ensemble models and correlation analysis based on Landsat 5–8 satellite data to forecast LST and explore its environmental relationships. This study employed the XGBoost machine learning algorithm to model seasonal LST dynamics in three rapidly urbanizing Arkansas cities, including Fort Smith, Little Rock, and Northwest Arkansas, using Landsat imagery from 2001 to 2021. The results show significant increases in urban heat, particularly in the summer, with Fort Smith seeing an increase in the area classified in higher-temperature bins (35–45 °C) from approximately 33% in 2001 to more than 83% by 2021. Model validation showed high predictive performance ($R^2 = 0.74$ – 0.78 , $RMSE \leq 1.46$ °C), indicating reliable project-based estimation of spatial LST variability for 2026 and 2031. The results revealed a substantial intensification of built-up area expansion, to 9.8% by 2026 and 20.7% by 2031, accompanied by cropland reductions of 13.2% and 25.5%, respectively. This rapid urban growth is projected to elevate summer LSTs above 45 °C across more than 700 km² combined, and winter LSTs to ≥ 25 °C across nearly 125 km² in the region by 2031. The integration of Landsat time series data and machine learning provide valuable insights for urban planners and policymakers, underscoring the critical importance of targeted climate-resilient strategies and sustainable urban development practices.

Keywords: land surface temperature (LST); urbanization impact; XGBoost; machine learning; remote sensing; urban heat island (UHI)

1. Introduction

Land Surface Temperature (LST) is an important biophysical parameter that indicates the thermal condition of the Earth's surface and has a substantial impact on energy exchanges between the land and atmosphere [1]. In urban environments, LST is closely linked to evapotranspiration, surface radiation balance, and near-surface thermal conditions, making it an essential variable for understanding urban microclimates [2,3]. Rapid urbanization modifies land surfaces by expanding impervious materials such as asphalt, concrete, and rooftops, which absorb and re-radiate heat more efficiently than vegetated or permeable surfaces. As a result, urban areas often exhibit elevated LST compared with surrounding rural landscapes, contributing to the urban heat island (UHI) effect [4]. This thermal difference leads to higher ambient temperatures in urban areas than in surrounding rural regions [5]. The amplified UHI effect increases building energy consumption,



Academic Editor: Adel Sepehr

Received: 31 May 2026

Revised: 28 June 2026

Accepted: 6 July 2026

Published: 8 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

degrades air quality, and raises heat-related health risks, particularly for socioeconomically disadvantaged groups [6].

With the rapid advancement of urban development, measuring and predicting LST has become critical for assessing climate resilience and forming mitigation strategies. Urbanization significantly increases surface heat retention due to reduced evapotranspiration, altered albedo, and denser configurations of synthetic materials. Urbanization greatly increases surface heat retention due to reduced evapotranspiration, changed albedo, and dense synthetic material configurations [7].

While broad climatic patterns influence atmospheric temperature, LST is highly sensitive to local land-surface factors such as vegetation cover, water bodies, and urban infrastructure characteristics. [8,9]. These localized thermal variations necessitate high-resolution, spatially explicit modeling to capture the full scope of urban climate transformation.

Researchers are now better able to track and simulate LST in both space and time, thanks to developments in thermal RS and geographic information systems (GIS) [4]. Researchers may derive complex surface condition variables and incorporate them into modeling algorithms using the comprehensive satellite imagery available, including Landsat 5 and 8, and spectral indices like the normalized difference vegetation index (NDVI) and the Normalized Difference Built-up Index (NDBI) [10–12].

Recent developments in machine learning (ML) and remote sensing (RS) methods have greatly improved our capacity to track and forecast urban thermal dynamics, particularly LST variations. [13–15]. To estimate LST, several studies have used different ML algorithms, highlighting both their advantages and disadvantages. One example is the successful application of XGBoost, a sophisticated gradient boosting technique, to create nonlinear relationships between environmental variables and LST. [16,17]. It has shown strong performance in capturing intricate relationships between predictors such as NDVI, surface emissivity, and land-use patterns. Furthermore, ensemble models that combine XGBoost with other techniques, such as Bagging and AdaBoost, have improved forecasting accuracy by effectively capturing nonlinear relationships among environmental parameters [18,19]. In addition, several urban climate studies used the LSTM model because of its superior ability to capture long-term dependencies and complex temporal-spatial dynamics of urban heat islands, outperforming traditional linear regression models. [7,20,21]. Studies that combined the urban thermal field variance index and land use/cover (LULC) simulations with artificial neural network (ANN) and cellular automata were also successful in predicting future urban heating scenarios, demonstrating the impact of urban expansion on thermal environments [22,23]. Although these studies demonstrate the value of ML for urban climate analysis, many applications have focused on large metropolitan areas or on short-term modeling settings. At the same time, mid-sized cities in the southern United States remain comparatively underexplored. [24].

This gap is particularly relevant in Arkansas, where cities such as Fort Smith, Little Rock, and Northwest Arkansas have experienced sustained urban expansion, vegetation loss, and increasing development pressure over the past two decades. These urban centers represent distinct geographic and environmental settings, yet all are undergoing land transformations that may intensify surface warming. Previous land use and land cover analyses have indicated continued growth of impervious surfaces and associated environmental change in these regions. Still, fewer studies have explicitly modeled the spatial variability of LST using multi-temporal remote sensing and ML in this regional context.

Accordingly, this study uses an XGBoost-based ML framework to model seasonal LST dynamics in three rapidly urbanizing Arkansas cities using multi-temporal Landsat imagery from 2001 to 2021. Landsat-derived spectral bands, thermal information, vegetation and built-up indices, and spatial variables were used to examine relationships between land

surface characteristics and LST and to evaluate model performance across contrasting urban environments. The specific objectives of this study are to: (1) characterize spatio-temporal patterns of LST across Fort Smith, Little Rock, and Northwest Arkansas; (2) assess the relationships between LST and Landsat-derived spectral and environmental predictors; and (3) evaluate the performance of XGBoost for modeling urban thermal variability using multi-temporal remote sensing data. By focusing on recent LST dynamics rather than unsupported long-range forecasting, this study contributes a more methodologically consistent framework for urban heat assessment. It offers insights relevant to climate-resilient urban planning and land management.

Crucially, to the best of the authors' knowledge, no prior research in the State of Arkansas has employed coupled LULC and LST predictive modeling, nor has any regional study deployed advanced machine learning architectures to address these localized environmental dynamics. This research directly addresses this regional literature gap by presenting a comprehensive multi-city assessment. The novelty of this work rests on establishing the first predictive machine learning baseline for Arkansas's primary urban networks while evaluating how centralized monocentric development forms (Little Rock and Fort Smith) differ in thermal accumulation from rapidly sprawling polycentric linear corridors (Northwest Arkansas).

2. Study Area

This study examines three major urban areas in Arkansas, USA: Fort Smith, Little Rock, and Northwest Arkansas (NWA) (Figure 1). These locations were chosen because they are the state's three largest urban centers in terms of population and development, and each has experienced significant land expansion over the past two decades. More importantly, including all three areas enables evaluation of the LST forecasting model across three urbanization patterns typical of the southern United States, allowing comparison of how distinct spatial forms and economic functions influence surface thermal conditions.

Fort Smith is located in western Arkansas near the Oklahoma border at approximately 35.39° N and 94.40° W, with an elevation of about 141 m. It is one of the state's principal industrial and commercial centers and has undergone notable urban and suburban development. These changes have altered land surface characteristics, making the area suitable for assessing relationships between urbanization and surface thermal behavior. Unlike Little Rock, Arkansas's capital city, Fort Smith exhibits a low-to-moderate-density industrial grid spatial morphology. Its population growth and development trajectory have been linear and more conservative, resulting in a land-cover composition that blends heavy industrial materials (with low albedo and high thermal mass) with stable agricultural landscapes.

Northwest Arkansas is a rapidly growing metropolitan corridor that includes Fayetteville, Springdale, Rogers, and Bentonville [25]. Located on the Ozark Plateau in the northwestern part of the state, the region generally ranges in elevation from approximately 350 to 420 m. Compared with Fort Smith and Little Rock, NWA has a different topographic and land cover setting, with higher elevation, strong recent population growth, and extensive conversion of vegetated and agricultural land to urban use. These characteristics make it especially relevant for evaluating emerging urban heat patterns in a fast-developing landscape [26].

All three study areas are influenced by a humid subtropical climate, characterized by hot summers, mild winters, and annual precipitation generally ranging from about 1100 to 1250 mm [27]. However, local variations in elevation, urban density, vegetation cover, and development intensity contribute to differences in thermal response across the three regions. These contrasting characteristics make Fort Smith, Little Rock, and Northwest Arkansas appropriate case studies for analyzing LST variability and machine-learning-based urban

heat modeling in Arkansas, demonstrating that the proposed forecasting framework is robust and transferable across vastly different urban structural contexts.

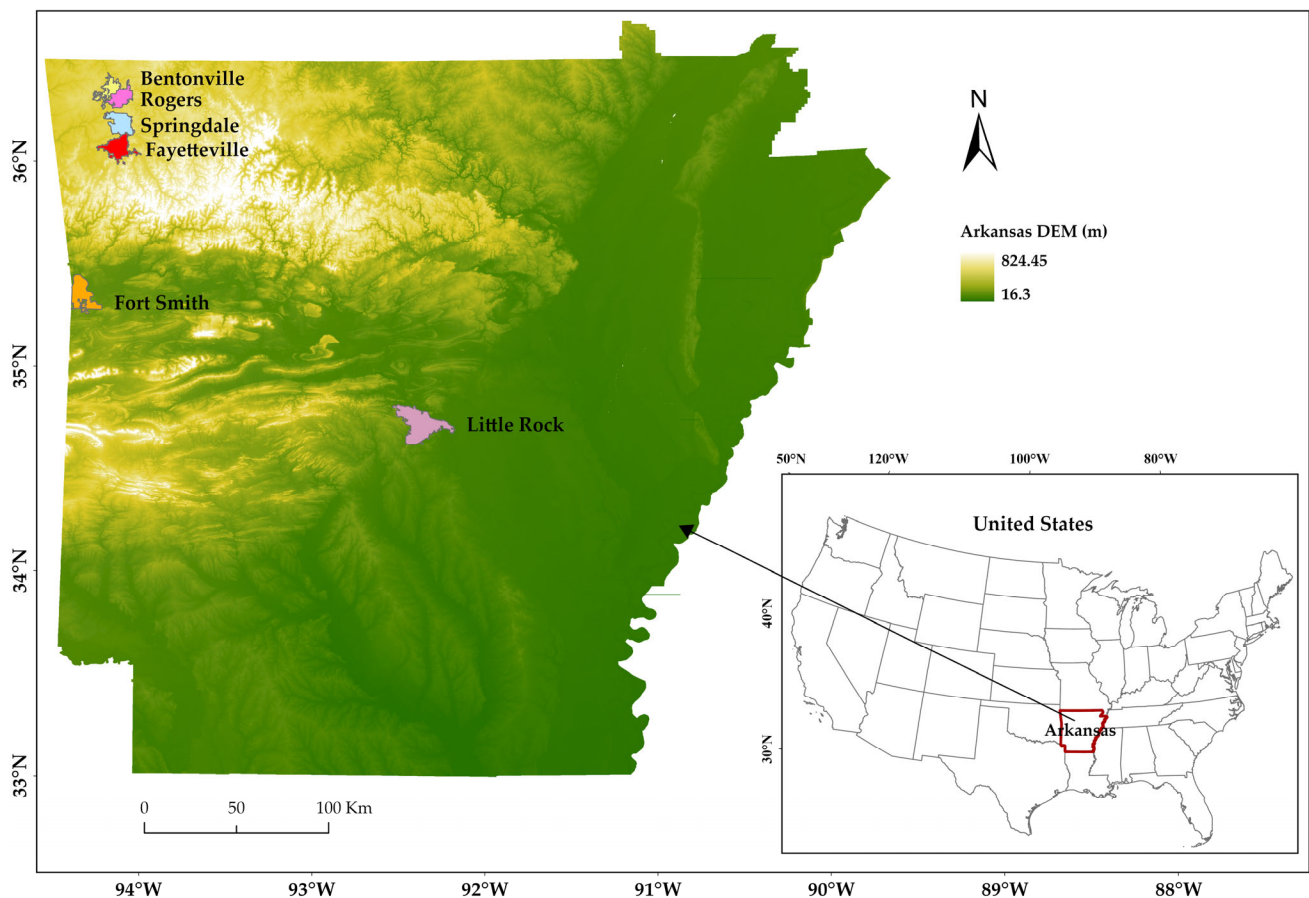


Figure 1. Arkansas map showing the locations of cities considered in the study, along with the elevation in the background.

3. Materials and Methods

3.1. Datasets

This study used multi-temporal Landsat imagery to analyze seasonal LST dynamics across three urban regions in Arkansas: Fort Smith, Little Rock, and Northwest Arkansas. Satellite data were selected for five reference years—2001, 2006, 2011, 2016, and 2021—across both summer and winter seasons to capture temporal variability in surface thermal conditions under contrasting seasonal conditions.

Landsat 5 Thematic Mapper (TM) imagery was used for 2001, 2006, and 2011, while Landsat 8 Operational Land Imager/Thermal Infrared Sensor (OLI/TIRS) imagery was used for 2016 and 2021. All scenes were obtained from the United States Geological Survey (USGS) archive through Google Earth Engine (GEE), using Collection 2 products to ensure radiometric and geometric consistency across years. Table 1 summarizes the selected Landsat images, including sensor type, the Worldwide Reference System (WRS) path/row, and acquisition dates for each study area and season [28]. The selection of USGS Tier 1 products guarantees an absolute clear-sky radiometric calibration accuracy of less than 1 °C within the thermal infrared bands, rendering these datasets highly robust for multi-decadal climate monitoring without the structural degradation common in uncalibrated raw datasets.

To reduce the effects of cloud contamination and seasonal inconsistency, scenes were selected within representative summer and winter periods and filtered to retain images

with low cloud cover [29]. In cases where cloud-free winter imagery was limited, the clearest available scenes nearest to the target season were used to preserve spatial coverage and image quality. Cloud and shadow pixels were masked using the QA_PIXEL band. The remaining small gaps were reduced using a 3×3 focal mean filter [30]. All raster layers were resampled to a uniform spatial resolution of 30 m, clipped to the study area boundaries, and projected to a common coordinate reference system to support consistent multi-temporal analysis. This 30 m spatial resolution represents the optimal open-source baseline for resolving microclimatic contrasts across mid-sized municipal infrastructures.

The Landsat imagery served as the basis for deriving thermal and spectral predictors used in subsequent modeling, including brightness temperature, vegetation indices, built-up indices, and other surface-condition variables [31]. These datasets formed the multi-temporal input framework used to model LST variability across the three study regions [32].

Table 1. Landsat images with the Worldwide Reference System (WRS) Path/Row used for seasonal LST analysis.

City	Year	Sensor	WRS Path/Row	Summer Acquisition Date(s)	Winter Acquisition Date(s)
Fort Smith	2001	Landsat 5 TM	026/035	22 July 2001	29 December 2001
	2006	Landsat 5 TM	026/035	2 June 2006	24 October 2006
	2011	Landsat 5 TM	026/035	3 August 2011	6 December 2010
	2016	Landsat 8 OLI-TIRS	026/035	13 July 2015, 17 September 2016	20 November 2016
	2021	Landsat 8 OLI-TIRS	026/035	13 July 2021	18 November 2021
Little Rock	2001	Landsat 5 TM	024/036	25 August 2001	28 October 2001
	2006	Landsat 5 TM	024/036	6 July 2006	13 December 2006
	2011	Landsat 5 TM	024/036	6 September 2011, 17 May 2011	9 November 2011
	2016	Landsat 8 OLI-TIRS	024/036	2 August 2016	6 November 2016
	2021	Landsat 8 OLI-TIRS	024/036	31 July 2021	4 November 2021
NWA	2001	Landsat 5 TM	026/035	22 July 2001	29 December 2001
	2006	Landsat 5 TM	026/035	2 June 2006	24 October 2006
	2011	Landsat 5 TM	026/035	3 August 2011	6 December 2010
	2016	Landsat 8 OLI-TIRS	026/035	13 July 2015, 17 September 2016	20 November 2016
	2021	Landsat 8 OLI-TIRS	026/035	13 July 2021	18 November 2021

3.2. Retrieval of LST

LST was derived from Landsat thermal imagery using a single-channel retrieval approach implemented in GEE. This method was applied to all selected Landsat scenes after image preprocessing and cloud masking. The workflow was designed to provide internally consistent LST estimates across multiple years and seasons for subsequent ML analysis.

For Landsat 5, thermal band 6 from the USGS Collection 2 Level 2 product was used, while for Landsat 8, thermal band 10 was used. Brightness temperature (BT) was obtained from the thermal bands after applying the appropriate scaling factors provided in the Landsat Collection 2 data products. Cloud and shadow pixels were masked using the QA_PIXEL band based on USGS bitmask criteria. To reduce the impact of small data gaps due to residual cloud contamination, a 3×3 focal mean filter was applied to the masked raster layers.

Surface emissivity was estimated dynamically from vegetation conditions derived from NDVI. Fractional vegetation cover (FV) was first calculated from NDVI values, and emissivity was then estimated as a function of FV following established single-channel LST retrieval procedures. The derived emissivity values were combined with brightness

temperature to compute LST using the Planck-based single-channel equation, which has been widely applied in Landsat-based thermal studies [33,34].

The general form of the retrieval is expressed as:

$$F_V = \left(\frac{NDVI - NDVI_{Min}}{NDVI_{Max} - NDVI_{Min}} \right)^2 \quad (1)$$

$$\varepsilon = 0.97 \times (1 - F_V) + 0.99 \times F_V \quad (2)$$

Brightness temperature (in Kelvin) and emissivity were then used to compute LST with Equation (3), the Planck-based single-channel equation [35,36]:

$$LST = \frac{T_B}{(1 + ((\lambda \times T_B)/\rho) \times \ln(\varepsilon))} - 273.15 \quad (3)$$

where:

TB = Brightness temperature (Kelvin), λ = Wavelength of thermal radiation (10.895 μm for Landsat 8), $\rho = h \times c / \sigma \approx 1.438 \times 10^{-2} \text{ m} \cdot \text{K}$, ε = Surface emissivity.

The resulting LST rasters were used as the dependent variable in subsequent modeling. Because these LST estimates were derived rather than directly observed, their interpretation should be considered in the context of methodological uncertainty associated with emissivity estimation, atmospheric effects, and thermal band calibration. To improve transparency, the study treats the retrieved LST maps as internally consistent thermal surfaces suitable for comparative multi-temporal modeling across the three study regions.

3.3. Feature Extraction and Selection

To characterize the land surface conditions associated with spatial variability in LST, a set of spectral, thermal, and spatial predictor variables was derived from the Landsat imagery for each study area. These variables were selected to represent key surface properties related to vegetation cover, built-up intensity, surface moisture, and geographic location.

The predictor set included reflectance-based spectral bands, BT, and several commonly used remote sensing indices derived from Landsat imagery. These included the NDVI, NDBI, normalized difference moisture index (NDMI), green soil-adjusted vegetation index (GSAVI), enhanced vegetation index (EVI2), bareness index (BI), and built-up soil index (BSI). Spatial coordinates (latitude and longitude) were also included to account for locational variation and spatial heterogeneity in LST patterns. Tables 2 and 3 summarize the predictor variables used in the modeling framework.

To explore the relationships between these variables and LST, Pearson correlation coefficients were calculated separately for the Landsat 5 and Landsat 8 datasets [37]. The resulting correlation matrices were used as an exploratory tool to evaluate the direction and relative strength of associations between predictor variables and LST, rather than as the sole basis for variable inclusion. This step helped assess the relevance of vegetation, built-up, and moisture-related indices in explaining thermal variability across the study areas [38].

The selected predictor variables were then assembled into the ML input dataset used for model development. Together, these variables provided a multi-dimensional representation of land surface characteristics and their relationship with observed LST patterns [39,40].

Table 2. Spectral bands and indices are used as predictors for LST modeling.

Feature	Description	Source
Latitude/Longitude	Spatial coordinate of each pixel	Metadata
Blue	Blue band reflectance	Landsat 8 OLI Band 2
Green	Green band reflectance	Landsat 8 OLI Band 3
Red	Red band reflectance	Landsat 8 OLI Band 4
NIR	Near-infrared reflectance	Landsat 8 OLI Band 5
SWIR1	Shortwave infrared 1 reflectance	Landsat 8 OLI Band 6
SWIR2	Shortwave infrared 2 reflectance	Landsat 8 OLI Band 7
BT	Brightness Temperature	Landsat 8 TIRS Band 10
Blue	Blue band reflectance	Landsat 8 OLI Band 2
Green	Green band reflectance	Landsat 8 OLI Band 3
Red	Red band reflectance	Landsat 8 OLI Band 4
NIR	Near-infrared reflectance	Landsat 8 OLI Band 5

Table 3. Spectral indices derived from Landsat 8 bands.

Index	Description	Formula
NDVI	Normalized Difference Vegetation Index	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$
NDBI	Normalized Difference Built-up Index	$(\text{SWIR} - \text{NIR}) / (\text{SWIR} + \text{NIR})$
NDMI	Normalized Difference Moisture Index	$(\text{NIR} - \text{SWIR1}) / (\text{NIR} + \text{SWIR1})$
GSAVI	Green Soil Adjusted Vegetation Index	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green} + 0.5) \times (1 + 0.5)$
EVI2	Enhanced Vegetation Index (2-band)	$2.5 \times (\text{NIR} - \text{Red}) / (\text{NIR} + 2.4 \times \text{Red} + 1)$
BI	Bare Soil Index	$(\text{SWIR1} + \text{Red} - \text{NIR} - \text{Blue}) / (\text{SWIR1} + \text{Red} + \text{NIR} + \text{Blue})$

3.4. Training of Selection Points for Predicting LST

LST modeling models were trained using a stratified random sampling strategy. Sample points for about 400 were distributed across urban areas in Fort Smith, Little Rock, and NWA to ensure balanced spatial and seasonal coverage. Each sampled point included all spectral and index-based predictors, as well as the corresponding LST value retrieved using the single-channel algorithm (Figure 2).

The sampled dataset was divided into training (70%) and validation (30%) subsets to support model calibration and independent performance assessment. Sample locations were distributed across the study area to provide broad spatial coverage and to capture variability in surface characteristics represented in the multi-temporal predictor stack. The resulting training and validation subsets were then used in the XGBoost (version 2.1.0) modeling framework described below [24,41,42].

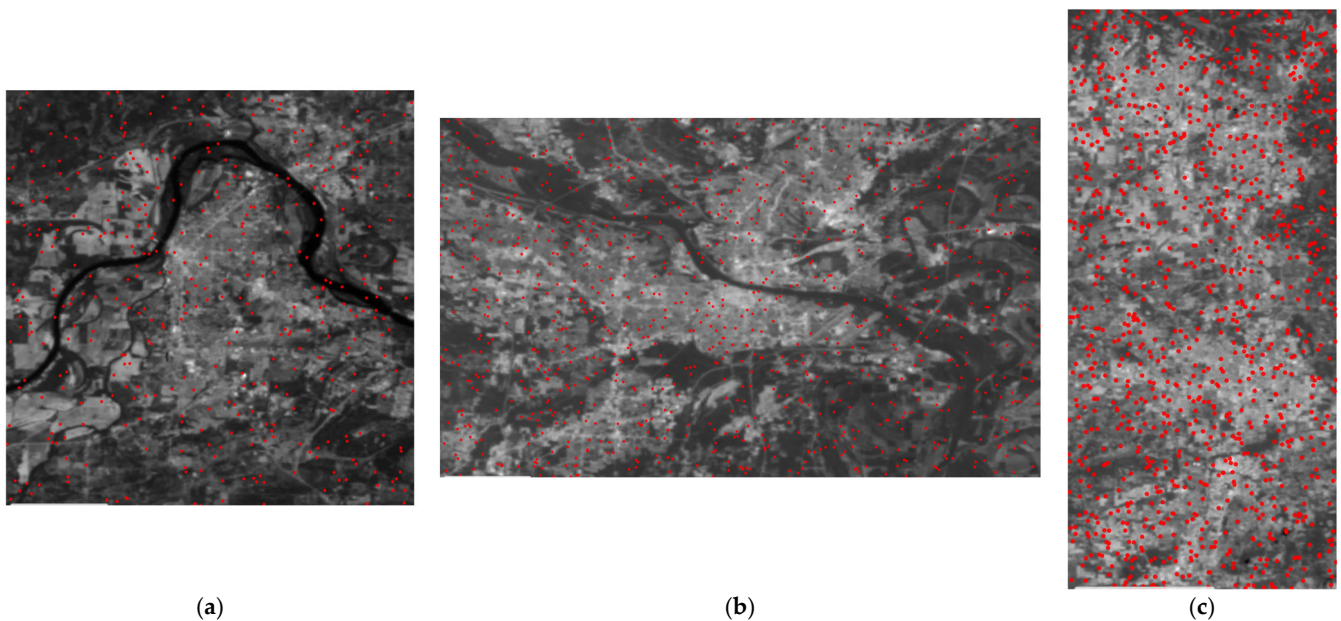


Figure 2. Spatial distribution of sampling points used for training and validation of the XGBoost model across the three urban study areas: (a) Fort Smith, (b) Little Rock, and (c) Northwest Arkansas. Red dots indicate selected pixels overlaid on LST images derived from the single-channel method. The sampling design ensures balanced spatial coverage for robust model performance.

3.5. XGBOOST Regression

The extracted spectral, thermal, and spatial predictors were used to model LST using the XGBoost algorithm. XGBoost is a supervised ML method based on gradient-boosted decision trees. It is widely recognized for its computational efficiency, robustness, and ability to capture nonlinear relationships in high-dimensional datasets [43].

The model was implemented in Python 3.11 using the XGBoost library in Google Colab. The predictor set included multi-temporal spectral and index-based variables derived from Landsat imagery, together with spatial coordinates and previously derived thermal information. These variables were used to represent both environmental conditions and spatial variability associated with surface thermal patterns [44].

The sampled dataset was divided into training and validation subsets using a 70:30 split. Model performance was evaluated using three standard regression metrics: coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) [18]. These metrics were used to assess how well the model reproduced the spatial variability of the target LST surface [17].

Model tuning focused on identifying an appropriate learning rate while maintaining the same overall model structure during repeated testing. The final calibrated model was then applied to the full predictor stack to generate spatially continuous LST estimates for each study area. This workflow allowed assessment of the model's ability to represent urban thermal variability from multi-temporal remote sensing inputs [45]. Figure 3 illustrates the comprehensive methodological framework of this study, which is divided into two integrated phases: Google Earth Engine (GEE)-based data extraction (right) and ML-based modeling (left).

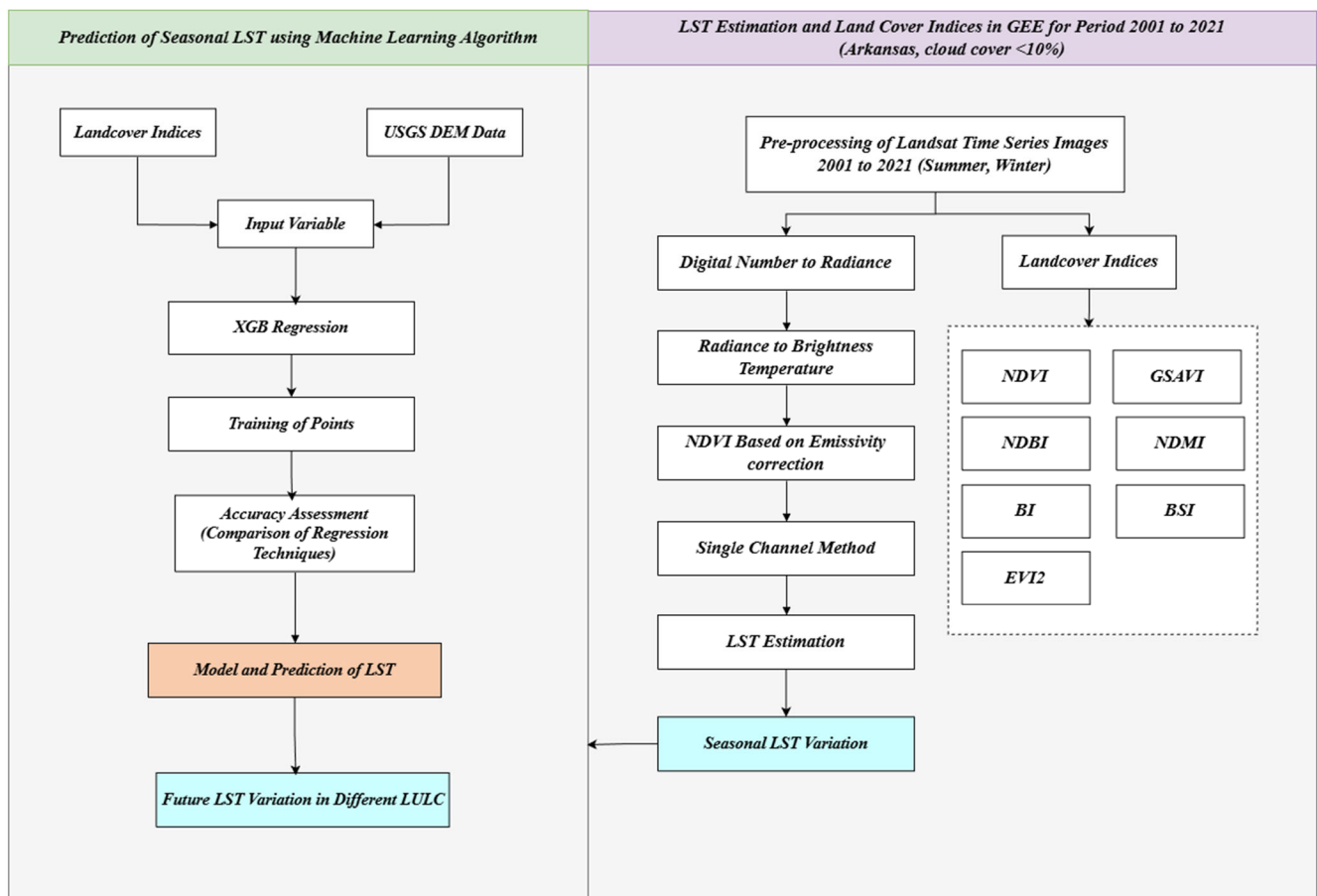


Figure 3. Workflow of the methodology used in this research.

3.6. Generation of Future Predictor Variables

For the 2026 and 2031 LST predictions, spatial predictor variables were developed using future LULC scenarios and a combined Cellular Automata–Markov (CA–Markov) model. The modeling of LULC, along with its calibration and validation, was detailed in our previous studies [46]. Briefly, future urban expansion and vegetation decline across the three case study areas were simulated using the CA–Markov model based on NLCD (National Land Cover Database) datasets from 2001 to 2021. Transition probability matrices and spatial transition rules were established using a 3×3 Moore neighborhood configuration and terrain-based constraints.

Since spectral indices (e.g., NDVI, NDBI, NDMI, BI) need satellite reflectance data, spectral proxy layers were built for 2026 and 2031 using CA–Markov results. Specifically, the mean index value associated with LULC 2021 (based on the Landsat 8 2021 base image) for each classification (Water, Vegetation, Cropland, Built-Up, and Bare Land) was extracted. These class-specific spectral signatures were then assigned to the newly transitioned pixels in the 2026 and 2031 simulated LULC maps. This substitution assumes that the fundamental physical properties of the land-cover classes remain relatively stable over the short-to-medium term. The resulting simulated predictor stacks served as input variables for the calibrated XGBoost model to project the future spatial distribution of LST.

4. Results

4.1. Model Accuracy

Model performance was evaluated across the three study areas using R^2 , RMSE, and MAE (Table 4). The results indicate that the predictor variables captured a substantial

proportion of the spatial variability in the target surface, with R^2 values ranging from 0.74 to 0.78.

Northwest Arkansas showed the highest model performance ($R^2 = 0.78$), followed by Fort Smith ($R^2 = 0.77$) and Little Rock ($R^2 = 0.74$). RMSE values ranged from 1.35 °C to 1.46 °C, while MAE values remained at or below 1.00 °C across all three study areas. These results indicate that the model reproduced spatial thermal patterns with reasonable consistency despite differences in topography, vegetation cover, and urban structure among the three regions.

The calibrated models also showed differences in selected learning rates among the study areas, reflecting variation in model behavior under different urban and environmental conditions. Overall, the validation statistics support the suitability of the XGBoost framework for modeling spatial variability in the target surface from multi-temporal remote sensing predictors. Figure 4 compares the observed values derived from the single-channel method with the corresponding model estimates for the three study areas.

Table 4. Accuracy assessment of the XGBoost regression method from spectral variables.

Parameter	XGBoost Model		
	Fort Smith	Little Rock	Northwest Arkansas
R^2	0.77	0.74	0.78
RMSE	1.35	1.46	1.45
MAE	0.91	0.94	1.00
Learning Rate	0.15	0.20	0.18
Number of samples	426	944	1067

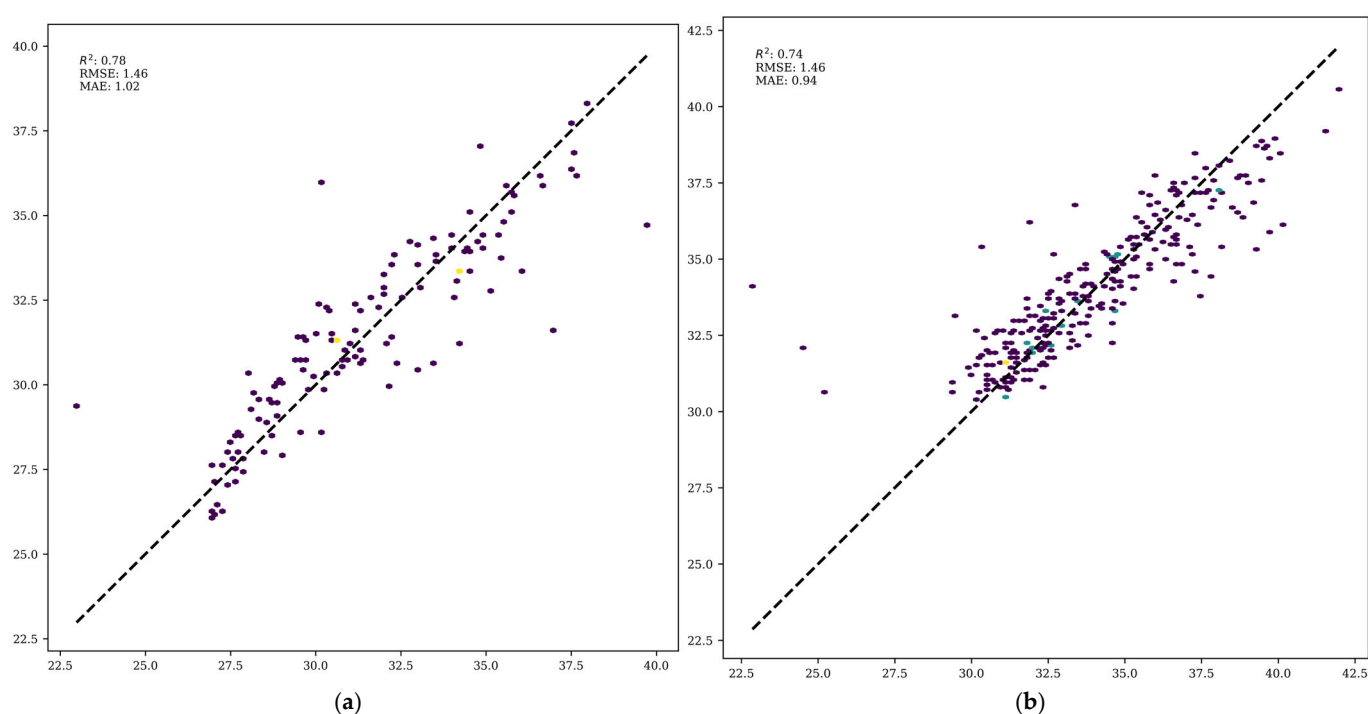


Figure 4. Cont.

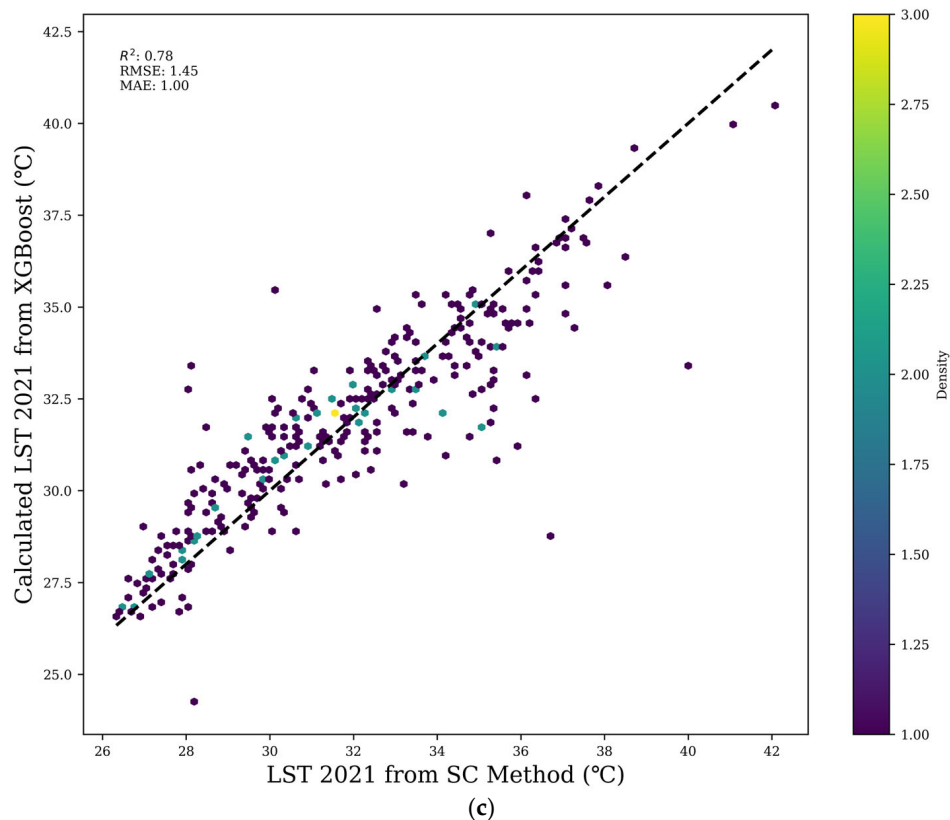


Figure 4. Performance evaluation of XGBoost Regression method for estimated and measured LST in 2021 using data from 2011 and 2016: (a) Fort Smith, (b) Little Rock, and (c) NW Arkansas.

4.2. Spatial Distribution of LST

Figure 5a–c illustrates the spatial distribution of LST across Fort Smith, Little Rock, and Northwest Arkansas from 2001 to 2021 at 5-year intervals. The maps show a clear intensification of surface warming over time, particularly in urban cores and expanding developed areas. This pattern is consistent with increasing impervious surface cover and reduced vegetation across the study regions.

Fort Smith and Little Rock show pronounced warming in downtown, industrial, and newly developed zones. At the same time, Northwest Arkansas exhibits notable increases in thermal intensity across rapidly urbanizing areas such as Fayetteville, Springdale, Rogers, and Bentonville. In contrast, surrounding vegetated and less-developed areas generally remain cooler, highlighting the moderating role of vegetation and lower development intensity on surface thermal conditions.

Overall, the spatial patterns indicate a strong urban heat effect in all three regions, with warming becoming more extensive and spatially concentrated over time. These observed distributions provide a spatial basis for understanding recent urban thermal change and for evaluating the relationship between land development and surface heat intensity in Arkansas.

Temperature statistics across the three study regions further support these mapped patterns (Table 5). During summer, mean temperatures generally increased over time, with Fort Smith and Northwest Arkansas showing especially strong warming by 2021. Little Rock also exhibited a warming trend, although with slightly lower mean summer values than the other two regions. Maximum summer values consistently remained high across all study areas, indicating persistent exposure to intense surface heating during warm-season conditions. Winter temperatures showed greater variability than summer temperatures, but the overall pattern suggests milder thermal conditions in later years for

all three study areas. Mean and maximum winter values generally increased over time, while extremely low values became less dominant in later periods. The tight alignment between physical urban expansion and LST inflation confirms a potent regional UHI effect driven by rapid LULC transformations [8,10,47]. As native vegetation and soil are replaced by high-mass, low-albedo engineering materials (e.g., asphalt, concrete), the landscape's natural latent heat flux via evapotranspiration is heavily suppressed. This drives a drastic spike in sensible heat flux, converting expanding metropolitan zones into highly efficient thermal traps.

The multi-centric warming pattern unique to NWA is a direct byproduct of its specific infrastructure boom. The rapid merger of four distinct municipalities into a single metropolitan strip has created a continuous corridor of high-heat-retaining surfaces. In contrast, the centralized warming shapes seen in Fort Smith and Little Rock reflect classic monocentric urban density profiles.

These findings indicate that if the current pace of impervious surface expansion continues without strategic urban forestry interventions, regional climate vulnerabilities will escalate. The resulting high baseline surface temperatures will likely increase cooling energy demand, accelerate localized smog formation, and pose severe cardiovascular and thermal health risks to vulnerable populations residing in urban cores.

Together, these seasonal patterns (Figure 5) indicate stronger warming during summer and more moderate warming during winter, consistent with recent urban thermal intensification across the region.

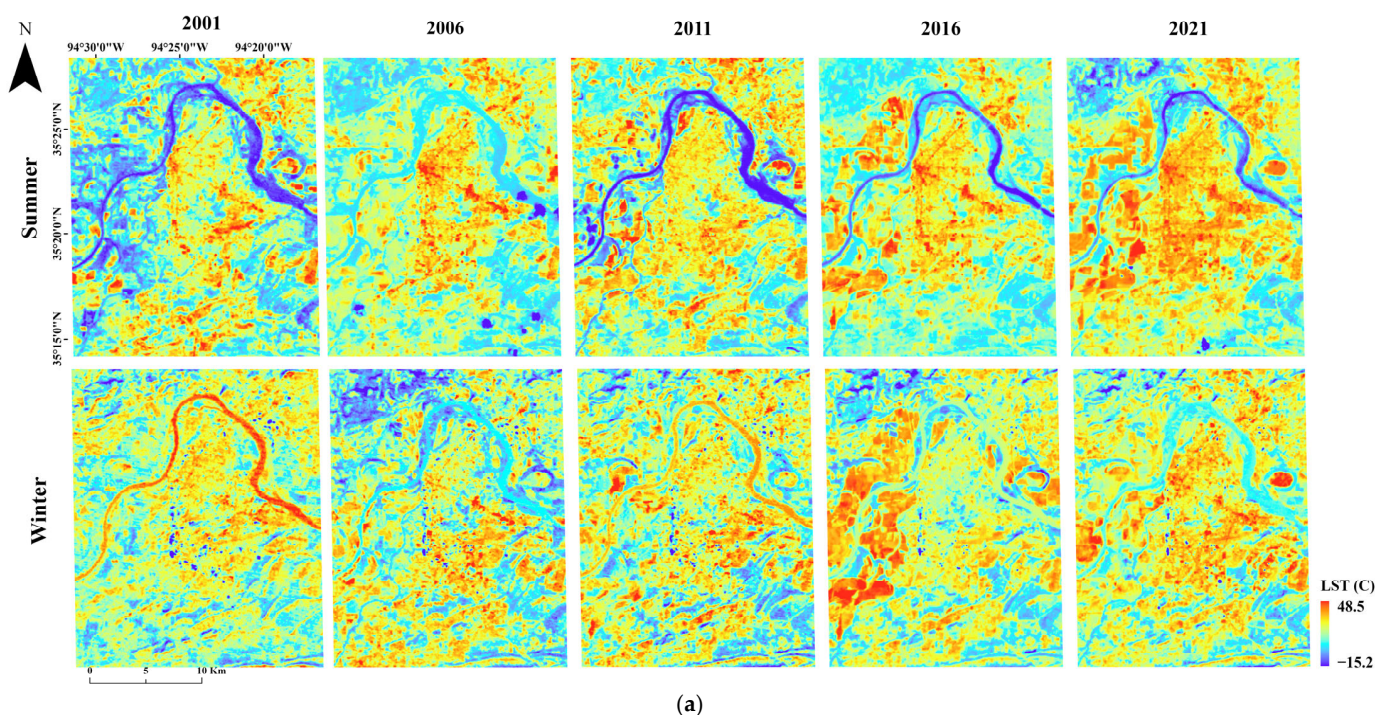


Figure 5. Cont.

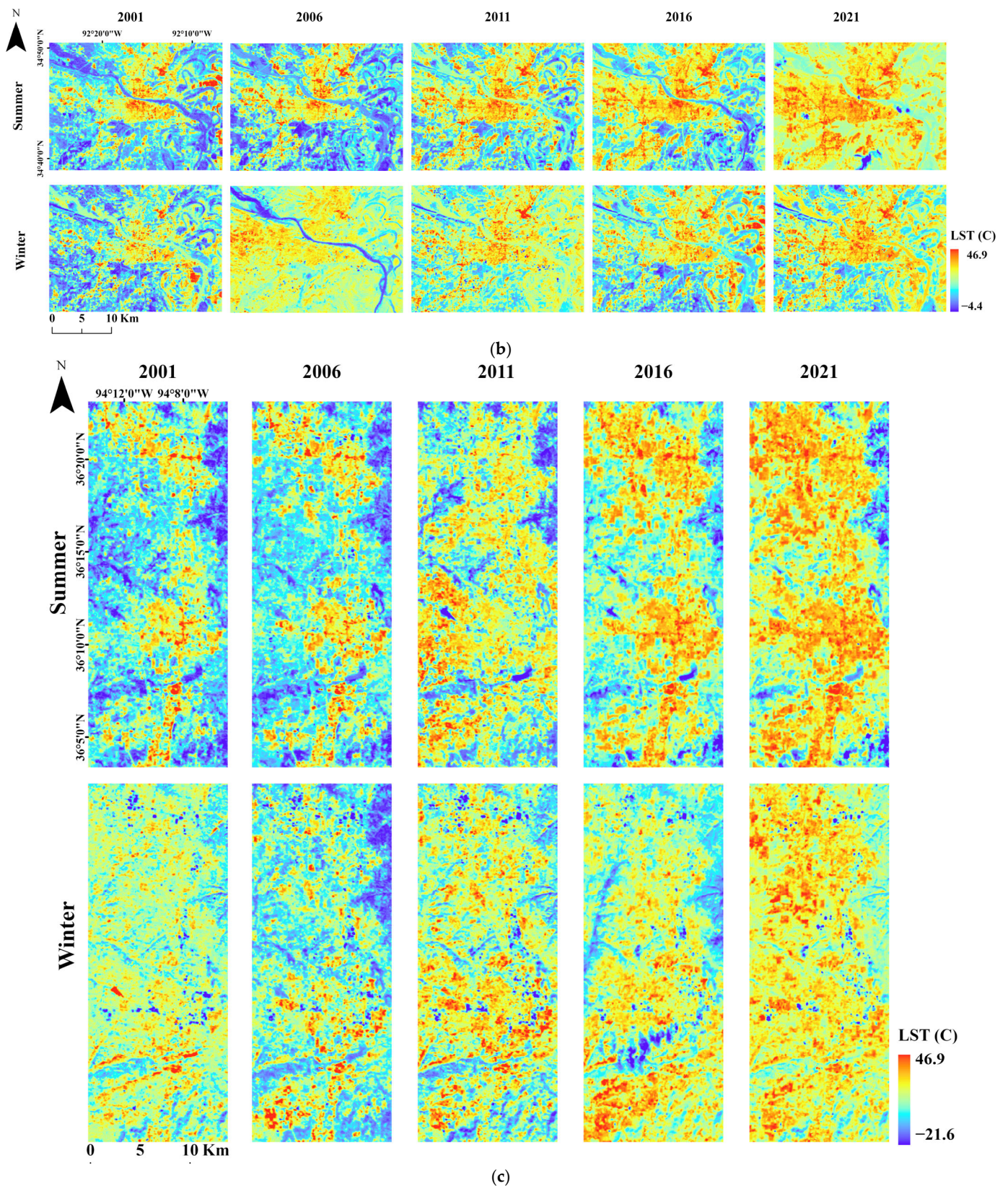


Figure 5. Spatial distribution of LST across the study regions ((a): Fort Smith, (b): Little Rock, (c): NWA).

Table 5. LST variation in (C) over the study area from 2001 to 2021 during the summer and winter seasons.

City	Season	Year	Min	Max	Mean	STD
Fort Smith	Summer	2001	12.76	48.41	28.91	2.67
		2006	26.92	45.10	32.01	2.32
		2011	18.47	45.43	31.58	2.99
		2016	27.57	46.15	34.09	2.66
		2021	28.49	46.72	37.47	2.76
	Winter	2001	−15.23	9.16	3.23	1.42
		2006	−14.91	16.26	6.97	1.53
		2011	−10.23	23.67	15.78	1.43
		2016	−1.53	23.41	16.09	2.27
		2021	3.73	27.57	17.54	1.62
Little Rock	Summer	2001	19.91	39.18	29.83	2.43
		2006	17.67	45.11	30.35	3.15
		2011	8.91	41.49	27.32	3.42
		2016	25.45	45.21	32.46	2.56
		2021	4.52	46.88	34.04	2.88
	Winter	2001	−4.45	24.23	14.04	1.75
		2006	−3.42	22.74	15.45	1.41
		2011	−3.41	25.39	15.31	1.92
		2016	−1.15	28.42	16.75	2.18
		2021	6.31	31.54	21.43	1.97
NWA	Summer	2001	24.11	43.09	31.18	2.02
		2006	11.91	41.92	27.12	2.31
		2011	27.43	43.89	32.51	2.13
		2016	21.59	45.73	31.99	3.03
		2021	24.14	46.91	39.10	2.24
	Winter	2001	−21.63	5.83	0.07	1.29
		2006	−20.94	15.78	5.18	1.58
		2011	−5.94	19.24	12.95	1.31
		2016	−4.43	19.09	13.56	1.66
		2021	−8.68	29.29	17.40	1.79

Temperature research at the three study sites—Fort Smith, Little Rock, and NWA—reveals different seasonal and temporal patterns from 2001 to 2021. During the summer season, average temperatures typically range from 27 °C to 39 °C across the research period, with significant variation between cities and years. For example, Fort Smith and Northwest Arkansas have recently had higher mean summer temperatures (2021: 37.5 °C and 39.1 °C, respectively), indicating a progressive warming trend.

Little Rock has slightly lower summer mean temperatures, ranging from 27 °C to 34 °C, but also shows a warming trend. The minimum summer temperatures range from 12.8 °C to 28.5 °C, while maximum values consistently fall within the 40–47 °C range, indicating intense heat periods during the summer months.

Winter temperatures show greater variability and lower averages than summer temperatures. Across all locations, mean winter temperatures range from near freezing or below (e.g., the minimum mean in Northwest Arkansas of 0.07 °C in 2001) to approximately

21.4 °C (Little Rock, 2021), reflecting the region's typical cold conditions. Maximum winter temperatures have been increasing, with values reaching as high as 31.5 °C in recent years. Minimum winter temperatures have also shifted, with some exceptionally low minimums seen in Northwest Arkansas (down to −21.6 °C in 2001), although warming has occurred in subsequent years, with fewer extreme lows.

This seasonal dynamic suggests both extended cold spells and gradual warming trends. Overall, the statistics show a shift toward higher-temperature zones during the summer over the last two decades, with substantial increases in mean and maximum temperatures, most notably in 2021 across all three cities. In contrast, the winter months are mildly warming, as evidenced by higher mean and maximum values and somewhat moderated minimum temperatures [48]. These temperature patterns are consistent with regional climate warming and point to increased heat exposure throughout the summer, as well as a milder cold season. Understanding these trends is critical for forecasting urban heat island effects, energy demand, and public health during extreme temperature occurrences in the research area [49].

The combined LST distribution Table 6 reveals a consistent warming trend across all three cities, particularly evident during the summer months. Fort Smith demonstrates a dramatic transition from predominantly mid-range summer temperatures (25–<30 °C, approximately 63% in 2001) toward the hottest temperature bins by 2021. By then, over 83% of the city area falls within the 35–<40 °C and 40–<45 °C ranges, underscoring a substantial urban heat intensification. Winter LSTs in Fort Smith are mostly in the cooler ranges (<5 °C and 5–<10 °C), with little transition to higher ranges, showing greater seasonal variance.

Little Rock follows a similar course. While summer distributions in the early years have a little higher fraction of mid-range values, there is a clear trend toward higher LST bins over time. By 2021, more than one-third of the area falls within the 35–<40 °C band, and an increasing share falls into the 40–<45 °C category. Winter patterns in Little Rock show a gradual shift upward, with the 15–<20 °C bin becoming increasingly dominant from 2006 onward. NWA has lower summer temperatures due to its higher elevation and increased vegetation cover. In 2001, more than 63% of the area remained in the 30–35 °C range. However, by 2021, NWA also experiences a striking rise, with nearly 94% of its urban area distributed within the 35–<40 °C and 40–<45 °C bands. Winter LSTs in NWA are initially confined to the coldest ranges (<5 °C), but by 2021, there is modest warming into the 10–<15 °C bin, suggesting early signs of seasonal temperature elevation.

These findings, drawn from high-resolution temporal analysis, highlight the growing burden of urban heat across Arkansas. The findings support the presence of both temporal warming (as evidenced by the upward trend between LSTs) and seasonal divergence, particularly the escalation of summer heat compared to relatively stable winter patterns. Such patterns highlight the need for heat-mitigation methods in urban development and land cover management [50].

Table 6. Distribution of different ranges of LST of summer and winter seasons over the study area from 2001 to 2021.

City	Year	LST < 20		20 ≤ LST < 25		25 ≤ LST < 30		30 ≤ LST < 35		35 ≤ LST < 40		40 ≤ LST < 45		LST ≥ 45		
		Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	
FS	2001	2.96	0.39	20.59	2.68	487.31	63.44	247.38	32.21	9.73	1.27	0.07	0.01	0.013	0.002	Summer
	2006	0	0	0	0	158.42	20.62	514.76	67.02	94.46	12.29	0.43	0.06	0.002	0.0002	
	2011	0.11	0.02	0.38	0.05	276.25	35.96	372.91	48.55	115.86	15.08	2.54	0.33	0.005	0.0005	
	2016	0	0	0	0	53.59	6.97	410.58	53.45	297.7	38.76	6.18	0.8	0.005	0.0007	
	2021	0	0	0	0	13.66	1.78	115.34	15.02	519.58	67.64	119.44	15.55	0.05	0.007	
LR	2001	0.004	0.0005	0.34	0.04	482.21	56.73	342.88	40.33	24.58	2.89	0.002	0.0002	0	0	
	2006	0.015	0.002	2.66	0.31	422.79	49.74	348.7	41.02	74.97	8.82	0.87	0.1	0.005	0.0006	
	2011	0.122	0.014	252.03	29.65	388.98	45.76	196.97	23.17	11.85	1.39	0.07	0.008	0	0	
	2016	0	0	0	0	169.05	19.88	526.96	61.99	150.66	17.72	3.34	0.39	0.13	0.0015	
	2021	1	0.12	4.65	0.54	19.99	2.35	511.89	60.22	298.34	35.09	14.08	1.66	0.05	0.006	
NWA	2001	0	0	0.06	0.007	255.39	31.93	507.85	63.49	36.37	4.55	0.19	0.02	0	0	
	2006	0.34	0.04	131.47	16.43	563.78	70.48	103.24	12.9	1.03	0.13	0.007	0.001	0	0	
	2011	0	0	0	0	78.98	9.87	596.32	74.55	123.8	15.48	0.76	0.095	0	0	
	2016	0	0	0.31	0.04	231.88	28.99	416.12	52.02	149.99	18.75	1.55	0.19	0.006	0.0008	
	2021	0	0	0.007	0.001	0.31	0.039	46.65	5.83	441.67	55.22	311.17	38.9	0.064	0.008	
FS	2001	12.47	1.62	683.72	89.02	71.88	9.36	0	0	0	0	0	0	0	0	Winter
	2006	0.46	0.06	80.75	10.51	679.18	88.42	7.66	0.99	0.02	0.002	0	0	0	0	
	2011	0.03	0.003	0.06	0.008	0.23	0.03	239.04	31.12	527.26	68.64	1.45	0.19	0	0	
	2016	0.001	0.001	0.06	0.008	0.25	0.04	251.16	32.70	478.62	62.31	37.96	4.94	0	0	
	2021	0	0	0.014	0.002	0.20	0.026	43.70	5.69	695.53	90.55	28.54	3.72	0.08	0.01	
LR	2001	0.017	0.002	0.044	0.005	0.437	0.052	585.56	68.88	263.25	30.97	0.71	0.08	0.014	0.0015	
	2006	0.001	0.001	0.074	0.008	0.90	0.11	317	37.29	530.66	62.42	1.38	0.16	0	0	
	2011	0.027	0.003	0.08	0.009	25.36	2.98	289.82	34.09	532.94	62.69	1.78	0.21	0.002	0.0002	
	2016	0.009	0.001	0.05	0.006	0.28	0.031	214.48	25.23	581.99	68.46	52.97	6.23	0.23	0.03	
	2021	0	0	0	0	0.019	0.002	0.054	0.006	288.91	26.93	582.21	68.49	38.82	4.57	
NWA	2001	388.75	48.60	410.78	51.35	0.33	0.042	0	0	0	0	0	0	0	0	
	2006	4.44	0.55	318.34	39.79	476.96	59.62	0.12	0.015	0.002	0.0002	0	0	0	0	
	2011	0.07	0.009	0.42	0.053	10.12	1.26	738.28	92.30	50.96	6.37	0	0	0	0	
	2016	0.07	0.009	0.46	0.058	19.43	2.43	631.64	78.97	148.26	18.53	0	0	0	0	
	2021	0.06	0.007	0.17	0.02	0.92	0.11	70.59	8.82	684.63	85.59	43.42	5.42	0.08	0.01	

4.3. LST Forecasting and Scenario Projections

The calibrated XGBoost framework was applied to the full predictor stack to generate spatially continuous estimates of the target LST surface for each study area. These modeled surfaces complement the observed distributions by showing how the multi-temporal spectral, thermal, and spatial predictors collectively represent urban thermal variability across Fort Smith, Little Rock, and Northwest Arkansas.

The estimated LST patterns were generally consistent with the observed spatial structure of warming identified in the earlier analysis. Areas with stronger built-up signatures and reduced vegetation tended to correspond to higher modeled thermal values, while less developed and vegetated areas were associated with lower thermal estimates. This agreement supports the utility of the selected predictor set for representing spatial heterogeneity in surface temperature.

Overall, the modeled spatial patterns reinforce the observed urban heat contrast among dense urban zones, transitional development areas, and surrounding vegetated land. These results demonstrate that the ML framework is effective for representing spatial variation in urban thermal conditions from multi-temporal remote sensing inputs.

Correlation Analysis

The correlation matrices (Figure 6) reveal significant associations between spectral bands, indices, and LST, indicating the physical dynamics of the land surface. High positive correlations ($r > 0.80$) among the visible bands (Blue, Green, Red) and between the two SWIR bands specify extensive data redundancy, as these bands capture similar reflectance and moisture characteristics. A near-perfect alignment ($r \approx 0.99$) among vegetation indices (NDVI, GSAVI, and EVI2) confirms that they provide nearly identical estimates of plant density. These greenness indicators are strongly inversely related to built-up and bare soil indices (NDBI, BI, and BSI), highlighting the distinct spectral split between natural and developed land. Likewise, moisture-sensitive indices like NDMI show strong negative correlations with dry, impervious surfaces. LST demonstrates a clear thermal footprint: it is positively correlated with built-up indices (0.59–0.70), linked to heat retention in urban areas, and negatively correlated with vegetation, which provides cooling through evapotranspiration. As expected, LST and the BT are almost perfectly coupled ($r > 0.98$) [20].

Remarkably, geographic coordinates (latitude/longitude) show minimal influence compared to land cover type. Because many predictors—particularly within the vegetation and urban clusters—are highly multicollinear, future modeling should utilize dimensionality reduction (such as PCA) or selective variable pruning to ensure statistical robustness and eliminate redundancy. BT showed the strongest positive association with LST, as expected given its close thermal relationship within the retrieval and modeling framework. Overall, the correlation patterns support the relevance of the selected predictor variables for representing variation in surface thermal conditions across the study areas.

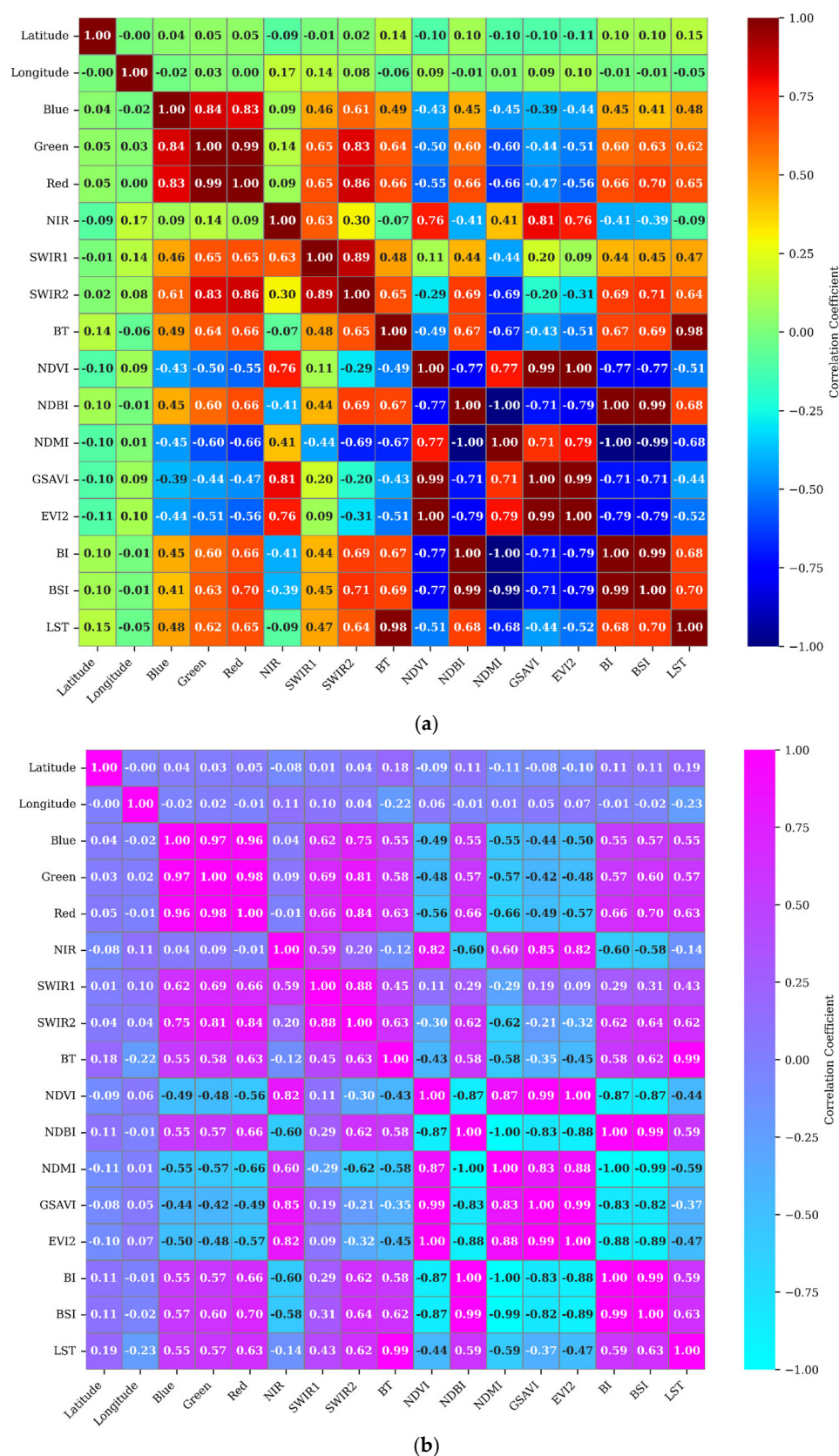


Figure 6. Correlation analysis between Landsat spectral features and LST: (a) Landsat-5 TM (b) and 8 OLI.

Spatial modeling of LST for the winter and summer seasons of 2026 and 2031 (Figure 7) shows that urban heat will continue to intensify, particularly in densely populated urban

cores and rapidly urbanizing peripheries. Compared to the reported LST in 2021, modeling shows increased and enhanced thermal hotspots in all three cities:

Fort Smith: Summer LST zones will migrate primarily to higher temperature categories (35–45 °C) by 2026 and 2031, with cooler and intermediate ranges (25–30 °C and 30–35 °C) decreasing. The 40–45 °C zone expands dramatically, illustrating severe urban heating and emphasizing essential areas for heat reduction measures.

Little Rock: Predicted LST maps show significant development of high-temperature zones (35–45 °C) by 2026, with much greater growth by 2031. This intensification is commensurate with the growth of urbanization and impervious surfaces.

NWA: This region, which has typically been colder because of its higher elevation and dense vegetation, is expected to warm significantly by 2031, reaching 40–45 °C. This emphasizes the rapid impacts of urbanization and the urgent need for UHI reduction through increasing plant cover and reflecting surfaces.

Projected changes in LST distribution across temperature ranges (Table 7) reinforce observed warming trends. For instance, by 2031, over 94% of Northwest Arkansas's urban areas are projected to fall within the 35–45 °C range, an increase from less than 5% in 2001. Similar upward shifts are observed in Fort Smith and Little Rock, underscoring the growing urban heat challenge.

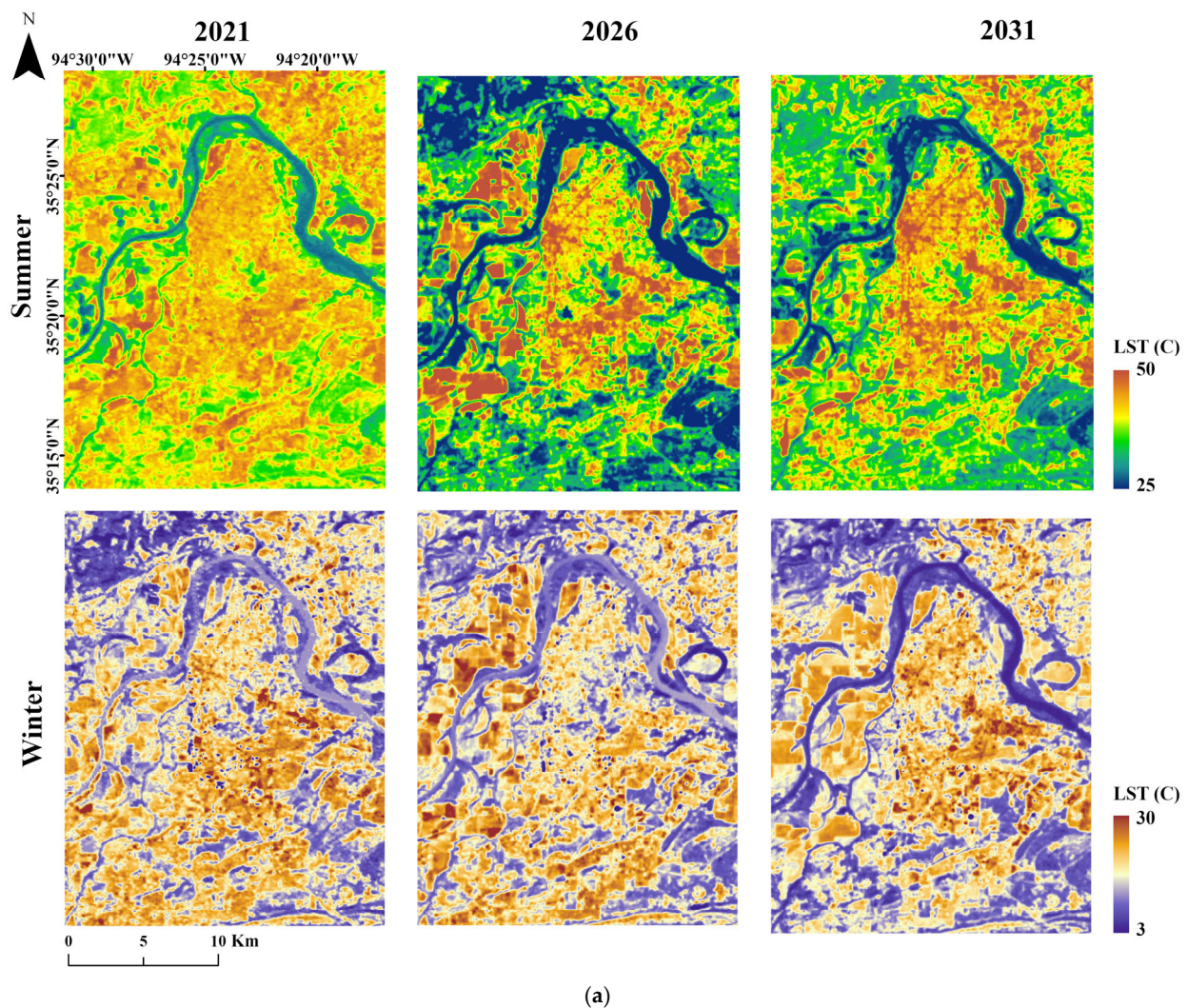


Figure 7. Cont.

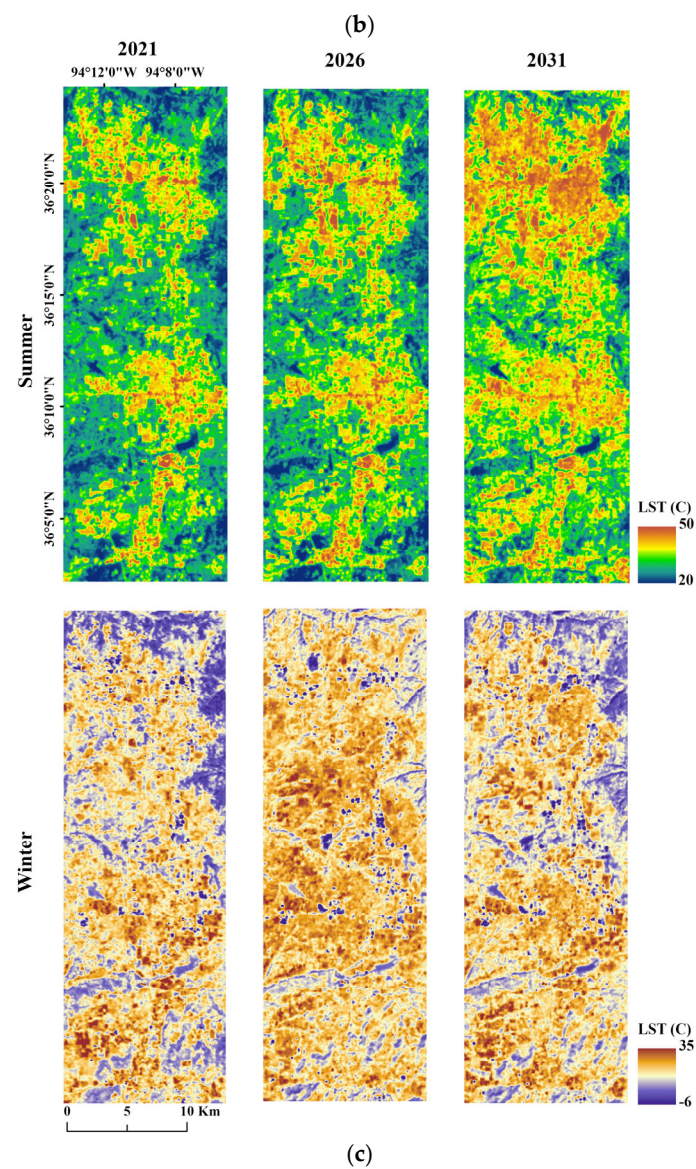
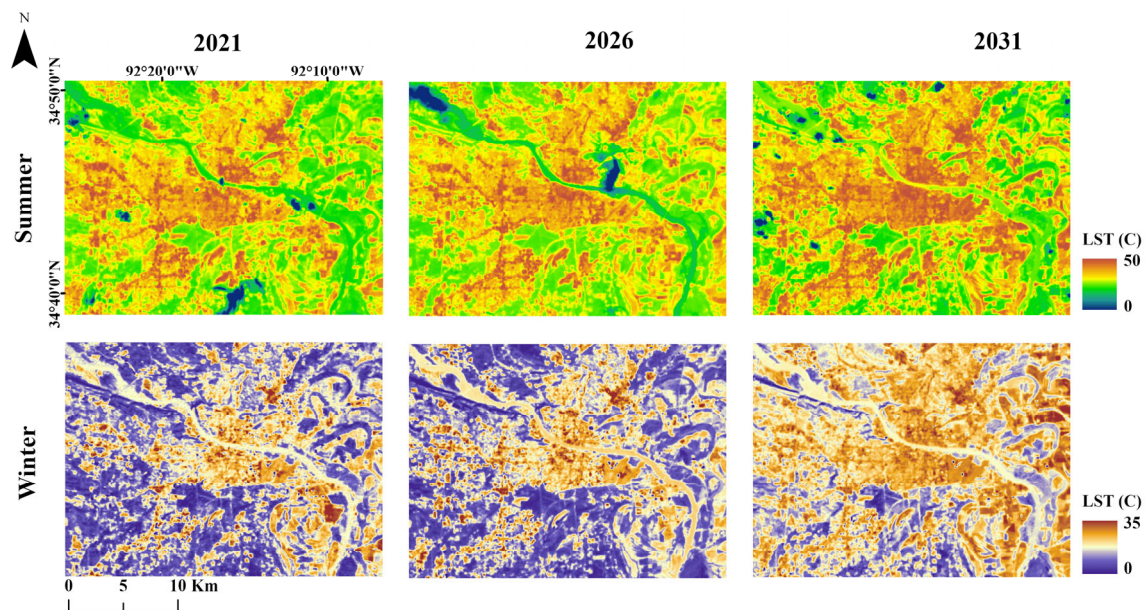


Figure 7. Forecasted winter LST map of the study area in 2026 and 2031 compared to 2021: (a) Fort Smith, (b) Little Rock, and (c) NW Arkansas.

Table 7. Distribution of different ranges of forecasted LST of summer and winter seasons over the study area.

City	Season	Year	LST < 20		20 ≤ LST < 25		25 ≤ LST < 30		30 ≤ LST < 35		35 ≤ LST < 40		40 ≤ LST < 45		LST ≥ 45	
			Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)
FS	Summer	2021	0	0	0	0	13.66	1.78	115.34	15.02	519.58	67.64	119.44	15.55	0.050	0.007
		2026	0	0	0	0	6.46	0.84	56.45	7.35	562.30	73.21	136.71	17.80	6.15	0.8
		2031	0	0	0	0	0	0	39.78	5.18	524.14	68.24	189.10	24.62	15.06	1.96
LR	Summer	2021	1	0.12	4.65	0.54	19.99	2.35	511.89	60.22	298.34	35.09	14.08	1.66	0.05	0.006
		2026	0.51	0.06	1.02	0.12	12.58	1.48	367.12	43.19	418.54	49.24	43.52	5.12	6.71	0.79
		2031	0.09	0.01	0.76	0.09	9.77	1.15	294.01	34.59	400.69	47.14	127.33	14.98	17.34	2.04
NWA	Summer	2021	0	0	0.007	0.001	0.31	0.039	46.65	5.83	441.67	55.22	311.17	38.90	0.064	0.008
		2026	0	0	0	0	0.16	0.02	27.75	3.47	342.88	42.87	410.14	51.28	18.87	2.36
		2031	0	0	0	0	0.08	0.01	15.68	1.96	314.24	39.29	444.93	55.63	24.87	3.11
City	Season	Year	LST < 0		0 ≤ LST < 5		5 ≤ LST < 10		10 ≤ LST < 15		15 ≤ LST < 20		20 ≤ LST < 25		LST ≥ 25	
			Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)	Area (Km ²)	Area (%)
FS	Winter	2021	0	0	0.014	0.002	0.20	0.026	43.70	5.69	695.53	90.55	28.54	3.72	0.08	0.01
		2026	0	0	0.008	0.001	0.07	0.01	18.51	2.41	677.98	88.27	68.43	8.91	3.07	0.4
		2031	0	0	0	0	0.04	0.005	14.90	1.94	614.53	80.01	125.73	16.37	12.90	1.68
LR	Winter	2021	0	0	0	0	0.019	0.002	0.054	0.006	288.91	26.93	582.21	68.49	38.82	4.57
		2026	0	0	0	0	0.009	0.001	0.076	0.009	168.47	19.82	631.64	74.31	49.81	5.86
		2031	0	0	0	0	0.019	0.002	0.043	0.005	113.31	13.33	667.42	78.52	69.19	8.14
NWA	Winter	2021	0.06	0.007	0.17	0.02	0.92	0.11	70.59	8.82	684.63	85.59	43.42	5.42	0.08	0.01
		2026	0.03	0.004	0.08	0.01	0.64	0.08	45.11	5.64	548.11	68.53	175.80	21.98	30.07	3.76
		2031	0.01	0.001	0.05	0.006	0.24	0.03	23.75	2.97	440.53	55.08	294.09	36.77	41.12	5.14

5. Discussion

This study provides a foundational benchmark as the first to use ML methods to model interconnected LULC changes and predict LST in Arkansas. While ML models such as gradient boosting have been widely applied to major global cities, applying them to fast-growing, mid-sized Southern U.S. metropolitan areas addresses a significant gap in urban climate research. The demonstrated adaptability of the XGBoost approach across different types of regional urban areas—both single-center cities and sprawling multi-center corridors—offers an important guide for climate-resilient land management at the community scale.

5.1. Algorithm Performance and Model Validation

The XGBoost model performed well in forecasting, with R^2 values between 0.74 and 0.78 for all three regions. This shows it can effectively model the complex environmental factors that affect urban LST. The model also had low RMSE values (≤ 1.46 °C) across all simulations, which are consistent with the typical uncertainties observed in satellite-based LST measurements and machine-learning temperature predictions in remote sensing studies.

To find the best modeling approach, this study compared several common ML methods for LST and environmental modeling. Earlier research has shown models like the Support Vector Machine (SVM) [51], Random Forest (RF) [12], boosted regression approaches [14], and conventional regression techniques [16] for surface temperature estimation. These methods have shown reliable performance under specific environmental conditions and predictor structures. [52].

XGBoost was chosen for this research because it can model nonlinear relationships, manage many correlated predictor variables, and capture complex interactions between spectral, thermal, and land-cover features. Unlike traditional regression methods that assume simple relationships, or ensemble models that can be affected by redundant inputs, XGBoost improves predictions step by step and uses regularization to prevent overfitting [14].

This was especially important in this study because correlation analysis showed a lot of overlap among variables. For example, vegetation indices like NDVI, GSAVI, and EVI2 were almost identical ($r \approx 0.99$), and visible spectral bands were strongly correlated ($r > 0.80$). Even with these overlaps, the model's predictions stayed consistent across all regions, showing that the framework worked well under different environmental and spatial conditions [16,53,54].

Using XGBoost also provides practical benefits for future LST forecasting. It can include different types of predictors from multiple sources, such as remotely sensed spectral data and projected LULC scenarios from the CA-Markov model. This enables a forecasting framework that connects future landscape changes to thermal responses.

However, the overlap among predictors shows there is room for improvement. Future studies could try feature selection and dimensionality reduction methods, or compare results directly with Random Forests, Support Vector Machines, deep learning models, and hybrid ensembles to make the models more efficient and easier to interpret.

5.2. Biophysical Drivers and Surface Energy Balance Mechanisms

Changes in LST across the study regions are closely tied to basic biophysical processes that control energy exchange at the surface. The strong positive link between LST and Brightness Temperature ($r > 0.98$) shows the workflow is consistent and highlights the main role of emitted thermal radiation in setting surface temperatures [16]. Also, the moderate-to-strong positive correlations between LST and built-up indicators like the Normalized

Difference Built-up Index (NDBI; $r = 0.59\text{--}0.70$) suggest that changes in land cover are a main cause of higher temperatures in Arkansas cities [32].

These findings align with established urban heat island theory, whereby conversion of vegetated and permeable landscapes into impervious urban surfaces modifies the local surface energy balance [41,44]. Materials like concrete, asphalt, and roofing hold more heat and cool less through evaporation than natural surfaces. When vegetation is removed and hard surfaces increase, less heat is lost through evapotranspiration and more is released as sensible heat, which raises surface temperatures [2].

The negative correlations between LST and vegetation indices further show that vegetation and green infrastructure help cool the environment. Vegetated areas reduce heat by providing shade, holding moisture, and cooling through evaporation, which helps control local climate conditions [45,46]. The strong correlation among vegetation indicators (NDVI, GSAVI, and EVI2) means that, even though these variables overlap, they all point to the vital role of vegetation in managing urban temperatures.

The thermal patterns seen in the study show that surface temperature depends on many factors, like land cover, surface reflectance, moisture, urban density, and location. This complexity highlights why ML models like XGBoost are useful, since they can capture these complicated, nonlinear relationships. This multi-factor supports the use of ML approaches such as XGBoost, which can capture complex nonlinear interactions among environmental predictors.

Future LST simulations in this study used land use and cover scenarios from the CA-Markov model. The links found between vegetation loss, more impervious surfaces, and higher LST help explain the warming expected in the 2026 and 2031 forecasts. These results suggest that if urban growth continues as it has, it could keep changing surface energy patterns and make regional temperatures rise.

5.3. Methodological Limitations and Multi-Source Uncertainties

While this integrated framework is accurate and reproducible for regional heat mapping, it has some limitations. First, combining CA-Markov land changes with XGBoost assumes that past land-use patterns and the link between land cover and temperature will stay the same over time. In reality, factors like economic changes, new zoning laws, or unexpected conservation efforts could change these patterns.

Second, this analysis only uses satellite-based surface conditions and not air temperatures (canopy urban heat island, UHI). While high LST shows physical environmental stress, it measures the surface temperature at the moment the satellite passes, not the ongoing conditions people feel at street level. Local field studies with weather stations can record frequent air temperature data, but they do not cover enough area to map large regions. So, using past satellite images is the only practical way to model temperature over many decades and across large areas where field data is missing.

Finally, the predictive framework only used physical remote sensing variables and did not include changing atmospheric factors such as wind speed and humidity, or socioeconomic data like population density and income. Future research could improve on this by adding climate change projections, demographic information, and atmospheric models to create a more complete forecasting tool.

5.4. Future Scenarios, Planning Implications, and Climate Resilience

Projections for 2026 and 2031 show that ongoing urban growth in Arkansas could worsen thermal conditions. CA-Markov simulations predicted more built-up land and fewer agricultural and vegetated areas. When these land-use scenarios were used in the

XGBoost model, the LST forecasts showed that high-temperature zones would spread across all regions studied.

These projected patterns suggest that increased impervious surfaces could raise environmental stress and change seasonal temperatures. While LST does not directly measure how people experience heat, higher surface temperatures are often linked to high energy use in buildings, changes in local weather, and wider urban climate issues [5,9].

From a planning point of view, these results suggest using climate-friendly land management, such as preserving vegetated areas, increasing urban tree cover, adding permeable infrastructure, and choosing development strategies that consider heat [11,44]. Forecasting tools like the one in this study can help identify areas that may face more heat in the future [55].

6. Conclusions and Future Directions

This study applied an XGBoost-based ML framework and multi-temporal Landsat data to analyze urban thermal patterns across three rapidly urbanizing regions in Arkansas: Fort Smith, Little Rock, and Northwest Arkansas. The results indicate a clear intensification of urban heat between 2001 and 2021, particularly in summer, with higher-temperature zones becoming more widespread in developed areas. Furthermore, the integrated forecasting framework confirms that these thermal patterns will significantly expand into the 2026 and 2031 horizons if recent landscape urbanization trajectories persist, escalating baseline temperatures across all three metropolitan networks. These patterns highlight the growing influence of urban expansion and land surface change on thermal conditions across the study region.

The modeling results showed that XGBoost effectively captured spatial variability in the target surface, with validation statistics ranging from $R^2 = 0.74$ to 0.78 and RMSE values of at most 1.46 °C. Predictor variables related to vegetation, built-up intensity, moisture conditions, and thermal properties contributed to the model's ability to capture urban thermal contrasts across different geographic settings.

Overall, the integration of remote sensing and ML provided a useful framework for assessing urban heat dynamics and identifying spatial patterns of thermal intensification. The findings are relevant to urban planning and land management, particularly in rapidly developing regions where vegetation loss and the growth of impervious surfaces can increase heat exposure. Future work could strengthen this framework by incorporating independent LST validation, additional explanatory variables, and scenario-based land-cover simulations to enable more robust long-term thermal assessment.

Author Contributions: All authors contributed to the design of the methodology and writing of the article. R.V. conducted the paper as part of his PhD program, collected the data, and performed the analysis. M.H.A. provided supervision and resources and contributed to the review, editing, and guidance throughout the research process. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement: The datasets used in this study are publicly available from open-access sources. Satellite data were obtained from USGS Earth Explorer. All data sources are properly cited within the manuscript. The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, A.; Zhang, M.; Kafy, A.-A.; Tong, B.; Hao, D.; Feng, Y. Predicting the Impacts of Urban Land Change on LST and Carbon Storage Using InVEST, CA-ANN and WOA-LSTM Models in Guangzhou, China. *Earth Sci. Inform.* **2022**, *16*, 437–454. [\[CrossRef\]](#)
2. Hao, L.; Sun, G.; Huang, X.; Tang, R.; Jin, K.; Lai, Y.; Chen, D.; Zhang, Y.; Zhou, D.; Yang, Z.-L.; et al. Urbanization Alters Atmospheric Dryness through Land Evapotranspiration. *npj Clim. Atmos. Sci.* **2023**, *6*, 149. [\[CrossRef\]](#)
3. Ismaila, A.-R.B.; Muhammed, I.; Adamu, B. A Spatial Regression Approach to Modeling Urban Land Surface Temperature. *MethodsX* **2023**, *10*, 102022. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Tanoori, G.; Soltani, A.; Modiri, A. Machine Learning for Urban Heat Island (UHI) Analysis: Predicting Land Surface Temperature (LST) in Urban Environments. *Urban Clim.* **2024**, *55*, 101962. [\[CrossRef\]](#)
5. Saha, M.; Kafy, A.-A.; Bakshi, A.; Faisal, A.-A.; Almulhim, A.I.; Rahaman, Z.A.; Al Rakib, A.; Fattah, M.A.; Akter, K.S.; Rahman, M.T.; et al. Modeling Microscale Impacts Assessment of Urban Expansion on Seasonal Surface Urban Heat Island Intensity Using Neural Network Algorithms. *Energy Build.* **2022**, *275*, 112452. [\[CrossRef\]](#)
6. Fameli, K.; Assimakopoulos, V.; Kotroni, V.; Retalis, A. Effect of the Land Use Change Characteristics on the Air Pollution Patterns Above the Greater Athens Area (GAA) After 2004. *Glob. NEST J.* **2013**, *15*, 169–177. [\[CrossRef\]](#)
7. Amir Siddique, M.; Wang, Y.; Xu, N.; Ullah, N.; Zeng, P. The Spatiotemporal Implications of Urbanization for Urban Heat Islands in Beijing: A Predictive Approach Based on CA-Markov Modeling (2004–2050). *Remote Sens.* **2021**, *13*, 4697. [\[CrossRef\]](#)
8. Assaf, G.; Hu, X.; Assaad, R. Predicting Urban Heat Island Severity on the Census-Tract Level Using Bayesian Networks. *Sustain. Cities Soc.* **2023**, *97*, 104756. [\[CrossRef\]](#)
9. Sahoo, S.; Majumder, A.; Swain, S.; Gareema; Pateriya, B.; Al-Ansari, N. Analysis of Decadal Land Use Changes and Its Impacts on Urban Heat Island (UHI) Using Remote Sensing-Based Approach: A Smart City Perspective. *Sustainability* **2022**, *14*, 11892. [\[CrossRef\]](#)
10. Carrillo-Niquete, G.A.; Andrade, J.L.; Valdez-Lazalde, J.R.; Reyes-García, C.; Hernández-Stefanoni, J.L. Characterizing Spatial and Temporal Deforestation and Its Effects on Surface Urban Heat Islands in a Tropical City Using Landsat Time Series. *Landsc. Urban Plan.* **2022**, *217*, 104280. [\[CrossRef\]](#)
11. Konda, V.; Chejarla, V.; Mandla, V.; Voleti, V.; Chokkavarapu, N. Vegetation Damage Assessment Due to Hudhud Cyclone Based on NDVI Using Landsat-8 Satellite Imagery. *Arab. J. Geosci.* **2018**, *11*, 35. [\[CrossRef\]](#)
12. Lee, P.; Park, J. An Effect of Urban Forest on Urban Thermal Environment in Seoul, South Korea, Based on Landsat Imagery Analysis. *Forests* **2020**, *11*, 630. [\[CrossRef\]](#)
13. Liu, C.; Zhang, M. Investigating LST Evolution and Heatwave Patterns Using Machine Learning in the Beijing-Tianjin-Hebei Urban Agglomeration. *iScience* **2026**, 116545. [\[CrossRef\]](#)
14. Samb, R.; Jena, A.; Manavvi, S.; Roy, U.K.; Yadav, B. Machine Learning Driven Land Surface Temperature Prediction and Urban Heat Risk Assessment in the Gambia. *Discov. Cities* **2026**, *3*, 49. [\[CrossRef\]](#)
15. Xie, T.; Jiang, M.; Zeng, C.; Li, H.; Zhang, G.; Li, C.; Shen, H. A mechanism-learning deeply coupled model for single-channel land surface temperature retrieval. *Remote Sens. Environ.* **2026**, *338*, 115350.
16. Li, X.; Jia, H.; Wang, L. Remote Sensing Monitoring of Drought in Southwest China Using Random Forest and eXtreme Gradient Boosting Methods. *Remote Sens.* **2023**, *15*, 4840. [\[CrossRef\]](#)
17. Mohammad, P.; Goswami, A.; Chauhan, S.; Nayak, S. Machine Learning Algorithm Based Prediction of Land Use Land Cover and Land Surface Temperature Changes to Characterize the Surface Urban Heat Island Phenomena over Ahmedabad City, India. *Urban Clim.* **2022**, *42*, 101116. [\[CrossRef\]](#)
18. Duan, S.-B.; Lian, Y.; Zhao, E.; Chen, H.; Han, W.; Wu, Z. A Novel Approach to All-Weather LST Estimation Using XGBoost Model and Multisource Data. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5004614. [\[CrossRef\]](#)
19. Pandey, S.; Kumari, N. Prediction and Monitoring of LULC Shift Using Cellular Automata-Artificial Neural Network in Jumar Watershed of Ranchi District, Jharkhand. *Environ. Monit. Assess.* **2023**, *195*, 130. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Badugu, A.; Arunab, K.S.; Mathew, A. Predicting Land Surface Temperature Using Data-Driven Approaches for Urban Heat Island Studies: A Comparative Analysis of Correlation with Environmental Parameters. *Model. Earth Syst. Environ.* **2023**, *10*, 1043–1076. [\[CrossRef\]](#)
21. Pande, C.B.; Egbueri, J.C.; Costache, R.; Sidek, L.M.; Wang, Q.; Alshehri, F.; Din, N.M.; Gautam, V.K.; Chandra Pal, S. Predictive Modeling of Land Surface Temperature (LST) Based on Landsat-8 Satellite Data and Machine Learning Models for Sustainable Development. *J. Clean. Prod.* **2024**, *444*, 141035. [\[CrossRef\]](#)
22. Lukas, P.; Melesse, A.M.; Kenea, T.T. Prediction of Future Land Use/Land Cover Changes Using a Coupled CA-ANN Model in the Upper Omo–Gibe River Basin, Ethiopia. *Remote Sens.* **2023**, *15*, 1148. [\[CrossRef\]](#)
23. Raza, D.; Khushi, M.; Shu, H.; Aslam, H.; Saleem, M.S.; Ahmad, A.; Mirza, S.; Saeed, U.; Khan, S.U. CA-ANN Based LULC Prediction and Influence Assessment on LST-NDVI Using Multi-Temporal Satellite Images. *Environ. Earth Sci.* **2024**, *83*, 144. [\[CrossRef\]](#)
24. Li, J.; Li, K.; Qian, Y.; Dou, X.; Han, Q.; Zeng, J.; Zhao, H.; Lan, Q.; Xu, Z.; Bai, J.; et al. An Urban Land Surface Temperature and Emissivity Separation Algorithm from ASTER TIR Data and Its Application. *Int. J. Digit. Earth* **2025**, *18*, 2459346. [\[CrossRef\]](#)

25. *Regional Growth*; Northwest Arkansas Council: Springdale, AR, USA, 2026.
26. United States Census Bureau. Census.Gov | U.S. Census Bureau Homepage. Available online: <https://www.census.gov/en.html> (accessed on 8 March 2026).
27. US Department of Commerce, National Weather Service. Little Rock, AR—Climate Data. Available online: <https://www.weather.gov/lzk/wxcntl3.htm> (accessed on 8 March 2026).
28. Earth Resources Observation and Science Center. *Landsat 8 Operational Land Imager/Thermal Infrared Sensor Level-1, Collection 1 Legacy Sample*; Earth Resources Observation and Science Center: Sioux Falls, SD, USA, 2024.
29. Bulgin, C.E.; Maidment, R.I.; Ghent, D.; Merchant, C.J. Stability of Cloud Detection Methods for Land Surface Temperature (LST) Climate Data Records (CDRs). *Remote Sens. Environ.* **2024**, *315*, 114440. [\[CrossRef\]](#)
30. Feng, S.; Cook, J.M.; Anesio, A.M.; Benning, L.G.; Tranter, M. Long Time Series (1984–2020) of Albedo Variations on the Greenland Ice Sheet from Harmonized Landsat and Sentinel 2 Imagery. *J. Glaciol.* **2023**, *69*, 1225–1240. [\[CrossRef\]](#)
31. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and Product Vision for Terrestrial Global Change Research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [\[CrossRef\]](#)
32. Komeh, Z.; Hamzeh, S.; Memarian, H.; Attarchi, S.; Alavipanah, S.K. A Remote Sensing Approach to Spatiotemporal Analysis of Land Surface Temperature in Response to Land Use/Land Cover Change via Cloud Base and Machine Learning Methods, Case Study: Sari Metropolis, Iran. *Int. J. Environ. Res.* **2025**, *19*, 98. [\[CrossRef\]](#)
33. Kumar, B.P.; Babu, K.R.; Anusha, B.N.; Rajasekhar, M. Geo-Environmental Monitoring and Assessment of Land Degradation and Desertification in the Semi-Arid Regions Using Landsat 8 OLI/TIRS, LST, and NDVI Approach. *Environ. Chall.* **2022**, *8*, 100578. [\[CrossRef\]](#)
34. Wang, L.; Yue, P.; Yang, Y.; Sha, S.; Hu, D.; Ren, X.; Wang, X.; Han, H.; Jiang, X. Land Surface Condition-Driven Emissivity Variation and Its Impact on Diurnal Land Surface Temperature Retrieval Uncertainty. *Remote Sens.* **2025**, *17*, 2353. [\[CrossRef\]](#)
35. Parastatidis, D.; Mitrika, Z.; Chrysoulakis, N.; Abrams, M. Online Global Land Surface Temperature Estimation from Landsat. *Remote Sens.* **2017**, *9*, 1208. [\[CrossRef\]](#)
36. Yu, X.; Guo, X.; Wu, Z. Land Surface Temperature Retrieval from Landsat 8 TIRS—Comparison between Radiative Transfer Equation-Based Method, Split Window Algorithm and Single Channel Method. *Remote Sens.* **2014**, *6*, 9829–9852. [\[CrossRef\]](#)
37. Landsat 8 | U.S. Geological Survey. Available online: <https://www.usgs.gov/landsat-missions/landsat-8> (accessed on 16 June 2025).
38. Mumtaz, F.; Tao, Y.; de Leeuw, G.; Zhao, L.; Fan, C.; Elnashar, A.; Bashir, B.; Wang, G.; Li, L.; Naeem, S.; et al. Modeling Spatio-Temporal Land Transformation and Its Associated Impacts on Land Surface Temperature (LST). *Remote Sens.* **2020**, *12*, 2987. [\[CrossRef\]](#)
39. Alexander, C. Normalised Difference Spectral Indices and Urban Land Cover as Indicators of Land Surface Temperature (LST). *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *86*, 102013. [\[CrossRef\]](#)
40. Santhosh, L.G.; Shilpa, D.N. Assessment of LULC Change Dynamics and Its Relationship with LST and Spectral Indices in a Rural Area of Bengaluru District, Karnataka India. *Remote Sens. Appl. Soc. Environ.* **2023**, *29*, 100886. [\[CrossRef\]](#)
41. Elmes, A.; Alemohammad, H.; Avery, R.; Caylor, K.; Eastman, J.R.; Fishgold, L.; Friedl, M.A.; Jain, M.; Kohli, D.; Laso Bayas, J.C.; et al. Accounting for Training Data Error in Machine Learning Applied to Earth Observations. *Remote Sens.* **2020**, *12*, 1034. [\[CrossRef\]](#)
42. He, P.; Chen, Z.; Zhang, L.; Ma, C.; Luo, C. Machine Learning Prediction of Future Land Surface Temperature from SAR Optical Fusion under Urban Expansion in Changsha, China. *Sci. Rep.* **2025**, *16*, 1258. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Xu, F.; Fan, J.; Yang, C.; Liu, J.; Zhang, X. Reconstructing All-Weather Daytime Land Surface Temperature Based on Energy Balance Considering the Cloud Radiative Effect. *Atmos. Res.* **2022**, *279*, 106397. [\[CrossRef\]](#)
44. Wang, J.; Tang, B.-H.; Zhu, X.; Fan, D.; Li, M.; Chen, J. A Comparative Analysis of Five Land Surface Temperature Downscaling Methods in Plateau Mountainous Areas. *Front. Earth Sci.* **2025**, *12*, 1488711. [\[CrossRef\]](#)
45. Ullah, S.; Khan, M.; Qiao, X. Evaluating the Impact of Urbanization Patterns on LST and UHI Effect in Afghanistan's Cities: A Machine Learning Approach for Sustainable Urban Planning. *Environ. Dev. Sustain.* **2025**. [\[CrossRef\]](#)
46. Vahid, R.; Aly, M.H. An Integrated CA–Markov Modeling Framework for Forecasting Land Use and Land Cover Dynamics in Arkansas, USA. *Geomatics* **2025**, *5*, 62. [\[CrossRef\]](#)
47. Kawamoto, Y. Effect of Urbanization on the Urban Heat Island in Fukuoka-Kitakyushu Metropolitan Area, Japan. *Procedia Eng.* **2016**, *169*, 224–231. [\[CrossRef\]](#)
48. Lee, H.S.; Trihamdani, A.R.; Kubota, T.; Iizuka, S.; Phuong, T.T.T. Impacts of Land Use Changes from the Hanoi Master Plan 2030 on Urban Heat Islands: Part 2. Influence of Global Warming. *Sustain. Cities Soc.* **2017**, *31*, 95–108. [\[CrossRef\]](#)
49. Alamdari, N.; Claggett, P.; Sample, D.; Easton, Z.; Yazdi, M. Evaluating the Joint Effects of Climate and Land Use Change on Runoff and Pollutant Loading in a Rapidly Developing Watershed. *J. Clean. Prod.* **2022**, *330*, 129953. [\[CrossRef\]](#)
50. Chen, Y.; Zhang, N. Urban Heat Island Mitigation Effectiveness under Extreme Heat Conditions in the Suzhou-Wuxi-Changzhou Metropolitan Area, China. *J. Appl. Meteorol. Climatol.* **2018**, *57*, 235–253. [\[CrossRef\]](#)

51. Feyzolahpour, M. Measuring the Spring and Autumn Fluctuations of Land Use Indices (LULC) Using Support Vector Machine (SVM) Method and Analyzing the Correlation Relationships of LST with NDBI, MNDWI and NDVI Indices (in the Heyran Pass Area). *Hydrogeomorphology* **2024**, *11*, 19–39. [[CrossRef](#)]
52. Vahid, R.; Aly, M.H. A Comprehensive Systematic Review of Machine Learning Applications in Assessing Land Use/Cover Dynamics and Their Impact on Land Surface Temperatures. *Urban Sci.* **2025**, *9*, 234. [[CrossRef](#)]
53. Chen, H.; Huang, J.J.; Dash, S.S.; McBean, E.; Wei, Y.; Li, H. Assessing the Impact of Urbanization on Urban Evapotranspiration and Its Components Using a Novel Four-Source Energy Balance Model. *Agric. For. Meteorol.* **2022**, *316*, 108853. [[CrossRef](#)]
54. Nasiri, V.; Heidarlou, H.B.; Alchin, A.A.; Moradi, F.; Rahmanian, S.; Afshari, S.; Maftai, C.; Griess, V.C. How Do Conservation Policies, Climate and Socioeconomic Changes Impact Hyrcanian Forests of Northern Iran? *Ecol. Inform.* **2023**, *78*, 102351. [[CrossRef](#)]
55. Mansourmoghaddam, M.; Rousta, I.; Ghafarian Malamiri, H.; Sadeghnejad, M.; Krzyszcak, J.; Ferreira, C.S.S. Modeling and Estimating the Land Surface Temperature (LST) Using Remote Sensing and Machine Learning (Case Study: Yazd, Iran). *Remote Sens.* **2024**, *16*, 454. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.