ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-3/W4-2025
Latin American GRSS and ISPRS Remote Sensing Conference (LAGIRS) 2025, 10–13 November 2025, Foz do Iguaçu, Paraná, Brazil

# Urban Thermal Dynamics at Pixel Resolution: Neighborhood-Specific Analysis Using Machine Learning and Multi-source Geospatial Data in Guadalajara, Mexico

Moises Orencio Ahumada Murillo[1], Juan Pablo Rivera Caicedo[2], Himer Avila George[1], Hugo Armando Ordoñez Erazo[3]

[1] Universidad de Guadalajara, México — moises.ahumada1774@alumnos.udg.mx, himer.avila@academicos.udg.mx
[2] Universidad Autónoma de Nayarit, México — jprivera@uan.edu.mx
[3] Universidad del Cauca, Colombia — hugoordonez@unicauca.edu.co

**Keywords:** Thermal patterns, pixel-level analysis, machine learning, remote sensing, neighborhood-scale modeling, spatial granularity

## Abstract

This research examines the relationship between urban physical characteristics and land surface temperature within discrete thermal pixels obtained from Landsat 8 Collection 2 Level-2 products in Guadalajara, Mexico. The city's extensive collection of high-quality geospatial data enables urban thermal analysis with high granularity. We integrate multiple datasets: a LiDAR-derived urban tree inventory with individual tree metrics; building footprints with roof material classifications; green space polygons; transportation infrastructure including roads and sidewalks; and water body delineations. Our methodology focuses on the pixel level—for each 30-meter thermal pixel (900 square meters), we precisely quantify all urban features within its boundaries, creating a comprehensive dataset where each pixel contains measurements of all elements present. This spatial integration enables multivariate regression modeling using machine learning, where predictor variables represent the quantity of each urban element and the target variable is pixel temperature. Using interpretable machine learning techniques, we identify the relative importance of urban elements in predicting thermal patterns, achieving $R^2$ values ranging from 0.69 to 0.83 across different urban contexts. This pixel-level approach provides granular understanding of urban thermal dynamics, supporting evidence-based urban planning decisions for thermal management.

## 1. Introduction

Urban areas worldwide are undergoing significant transformations, resulting in altered thermal environments that affect energy consumption, public health, and overall quality of life. Identifying the drivers of urban thermal variations is essential for sustainable urban planning, particularly in rapidly expanding metropolitan regions such as the Guadalajara Metropolitan Area (AMG), Mexico. The expansion of the metropolitan area has led to substantial changes in land use, including vegetation loss, which contributes to documented climate challenges and increased vulnerability to heat events, as reported in local climate action plans (Gobierno de Guadalajara and C40 Cities and IMEPLAN and Gobierno de Jalisco, 2020).

The spatial distribution of land surface temperature (LST) is shaped by a complex interplay of the physical characteristics of the urban landscape, including building morphology, vegetation cover, surface materials, and urban geometry (Sun et al., 2019). In the Guadalajara Metropolitan Zone, reported studies show temperature differences of up to 4 °C between densely built areas and adjacent vegetated zones, evidencing the Urban Heat Island (UHI) phenomenon (Camacho Sandoval et al., 2024). Addressing these thermal disparities is a key aspect of the metropolitan climate action framework (Gobierno de Guadalajara and C40 Cities and IMEPLAN and Gobierno de Jalisco, 2020).

Traditional urban climate studies often operate on broader scales, employing frameworks such as Local Climate Zones (LCZs) to classify urban landscapes. While useful for identifying broad spatial patterns, macro-scale approaches may fail to capture fine-grained variations that influence localized thermal conditions relevant to site-specific interventions (Demuzere et al., 2022). There is a need to understand how specific urban features within smaller spatial units, such as satellite thermal pixels, contribute to observed temperatures.

Guadalajara provides a relevant case study for such granular analysis due to the availability of high-resolution geospatial datasets. A comprehensive public urban tree inventory, developed using LiDAR, provides detailed information on more than 1.1 million trees. This detailed urban forest characterization, when integrated with data on building footprints, construction materials, additional green spaces, and infrastructure, enables the precise quantification of urban components on the thermal pixel scale (Morales Manilla et al., 2018).

The present analysis examines the relationship between urban morphology and thermal patterns at 30-meter pixel resolution in Guadalajara. Selected neighborhoods exhibiting distinct urban forms are analyzed, as the variables influencing thermal conditions vary significantly across different areas of the city. We employ machine learning techniques, similar to the approaches used in other subtropical high-density cities (Sun et al., 2019), to model and interpret the influence of specific urban elements quantified within each pixel on its LST.

The practical applications of this research include:

- Identifying and ranking urban elements by their relative importance in predicting local temperatures
- Supporting data-driven decisions for urban interventions in existing neighborhoods

This research aims to apply a pixel-level analysis methodology to neighborhoods with high-quality geospatial data, developing

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-1-2026  

multivariate regression models that quantify the influence of urban elements on land surface temperature and determine their relative importance in different urban contexts in Guadalajara.

## 2. Materials and Methods

### 2.1 Study Area

This study examines specific neighborhoods within the Guadalajara Metropolitan Area (AMG) of Jalisco, Mexico. We focus our analysis on three neighborhoods: Jardines Alcalde, Panorámica de Huentitán, and Zona Centro. These neighborhoods were selected based on the availability of high-quality data and their distribution across different sectors of the city, allowing a geographically representative analysis. The spatial distribution of the selected neighborhoods enables the capture of diverse urban configurations throughout Guadalajara, while preserving the data resolution required for detailed pixel-level thermal analysis. By examining neighborhoods in different parts of the city, we can identify how urban elements influence the temperature of the land surface in diverse urban contexts within the metropolitan area, with potential applications to other neighborhoods with similar characteristics.

### 2.2 Data Sources

A diverse collection of high-resolution geospatial datasets covering the Guadalajara Metropolitan Area was compiled, enabling detailed characterization of the urban environment at the pixel level. The primary datasets employed are described below:

- **LiDAR Urban Tree Inventory:** Individual tree data were obtained from a public inventory generated using LiDAR technology, available through the "Visor Urbano" platform (Gobierno de Guadalajara, 2024). The dataset includes geographic coordinates, canopy radius, and height measurements for trees in public spaces.
- **Landsat 8 Surface Temperature Product:** Land Surface Temperature data were obtained from Landsat 8 Collection 2 Level-2 Science Products (USGS, 2024). According to the Landsat 8-9 Calibration and Validation Algorithm Description Document, the TIRS (Thermal Infrared Sensor) bands natively capture thermal data at 100-meter spatial resolution (USGS EROS, 2023, Section 2.2). The delivered products undergo resampling to 30-meter resolution to align with the multispectral OLI bands. This resampling process employs cubic convolution interpolation combined with the Akima method to achieve subpixel alignment (USGS EROS, 2023, Section 4.3.3), resulting in thermally corrected products that maintain geometric consistency across all Landsat 8 bands. It is important to note that although the LST products are provided at 30m pixel resolution, the effective thermal information remains at the original 100m scale. Clear-sky imagery with minimal cloud cover was selected from the available archive.
- **Cadastre, Sidewalks, and Road Network:** Vector datasets representing built environment components were obtained from (Instituto de Información Estadística y Geográfica de Jalisco, 2024). These include building footprints with roof material classifications, sidewalk polygons, and road networks. Roof material classifications are particularly relevant as different materials exhibit distinct thermal properties: concrete roofs have high thermal mass with significant heat storage capacity (Kubota et al., 2020), metal sheets show rapid heating and cooling cycles (Hernández-Pérez et al., 2014), while tile materials demonstrate intermediate thermal performance (Santamouris, 2017). These properties justify including roof materials as separate predictor variables in LST modeling.
- **Green Areas and Water Bodies:** Additional geospatial datasets were obtained from FIPRODEFO (Fideicomiso para la Administración del Programa de Desarrollo Forestal del Estado de Jalisco, 2024), including polygons representing surface green areas (distinct from LiDAR trees) and surface water bodies.

### 2.3 Data Preprocessing and Feature Extraction

All geospatial datasets were reprojected to a common coordinate reference system for spatial consistency. The feature extraction process quantified urban elements within each 30m LST pixel boundary. Using the LST pixels as a base grid (Figure 1(b)), we performed spatial overlay operations to calculate: (1) intersection areas for polygonal features (buildings by roof type, roads, sidewalks, green spaces, water bodies), and (2) count and average metrics for point features (tree inventory). This produced a structured dataset where each row represents one pixel containing the temperature value and all quantified urban features within its bounds (Table 1).

This processing pipeline culminated in a structured tabular dataset where each row represents a single 30 m pixel with a unique `pixelID`. Table 1 displays a representative sample of this resulting dataset, ready for input into our regression model. Each row contains the target variable (Land Surface Temperature) and all predictor variables that quantify urban morphology within that pixel. The sum of predictor areas (m²) per pixel may be less than the total pixel area (900 m²) as the available data layers do not account for all surface types, resulting in uncharacterized gaps. This format provides a clean, analysis-ready dataset where all relevant urban elements have been precisely quantified at the pixel level, making it suitable for direct input into machine learning regression algorithms. These variables derived form the basis for the subsequent exploratory analysis (Section 2.4) and regression modeling (Section 2.5).

### 2.4 Exploratory Data Analysis

Before constructing the predictive model (Section 2.5), an exploratory data analysis (EDA) was performed on the final dataset generated in Section 2.3. This step aimed to understand the distribution of the target variable and identify relationships between the urban form variables and land surface temperature.

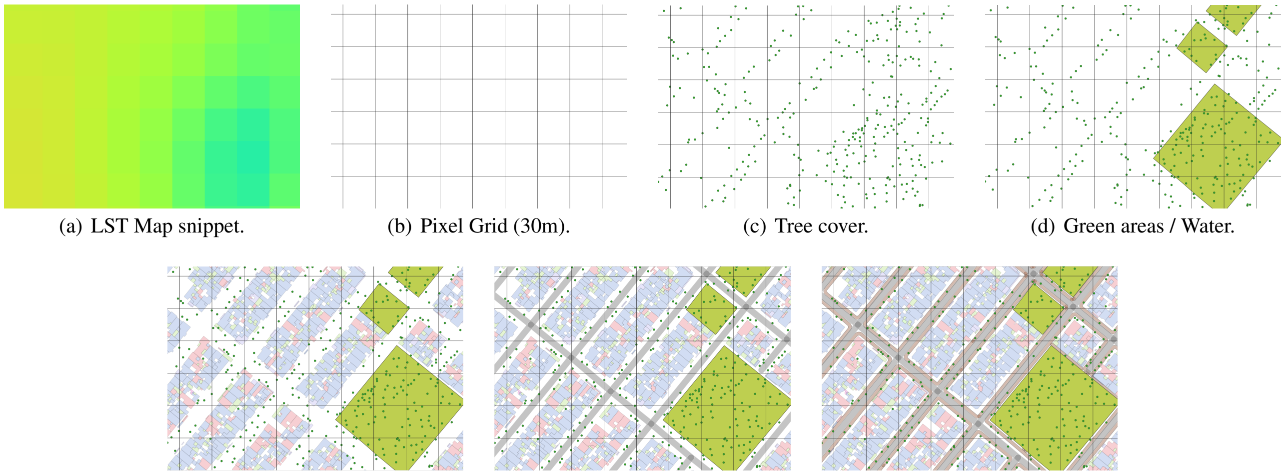
We first examined the statistical distribution of pixel temperatures across all analyzed neighborhoods, as shown in Figure 2. The histogram reveals a bimodal distribution with a primary peak around 42.12°C (the median) and a secondary peak near 36°C. This pattern suggests two distinct thermal regimes within the study area, potentially corresponding to different urban morphologies. The overall temperature range spans from 29.2°C to 48.4°C, with a mean of 41.41°C and standard deviation of 2.92°C, demonstrating considerable thermal variation within the urban neighborhoods.

Next, we examined the Pearson correlation coefficients between all derived predictor variables and the target variable (LST). Figure 3 visualizes these correlations, providing initial insights

| Pixel ID | Temp. (°C) Target | Concrete Roof (m²) | Metal Sheet (m²) | Tile Roof (m²) | Under Const. (m²) | Dome Roof (m²) | Sports Court (m²) | Other Roof (m²) | Avg. Tree Ht(m) | Total Trees | Green Area (m²) | Road Area (m²) | Sidewalk Area (m²) | Water Bodies (m²) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 42.68 | 482.8 | 171.1 | 2.3 | 0.0 | 1.2 | 0.0 | 3.5 | 3.2 | 1 | 5.8 | 87.3 | 45.2 | 0.0 |
| 2 | 42.51 | 391.1 | 209.8 | 0.0 | 4.5 | 0.0 | 2.7 | 0.0 | 8.7 | 1 | 5.3 | 47.7 | 67.7 | 0.0 |
| 3 | 42.95 | 174.1 | 11.3 | 59.1 | 0.0 | 0.0 | 15.4 | 1.8 | 4.8 | 5 | 32.7 | 185.1 | 188.7 | 0.0 |
| 4 | 43.14 | 347.2 | 6.2 | 12.6 | 7.3 | 2.1 | 0.0 | 0.0 | 5.1 | 2 | 23.1 | 217.8 | 128.6 | 0.0 |
| 5 | 40.93 | 563.8 | 0.0 | 85.7 | 1.9 | 0.0 | 8.6 | 3.2 | 2.8 | 2 | 5.0 | 112.6 | 75.3 | 0.0 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

*Note: Temperature is the target variable, while all other variables except PixelID are predictors. The sum of predictor areas (m²) per pixel may be less than the total pixel area (900 m²) as the available data layers do not account for all surface types, resulting in uncharacterized gaps.*

Table 1. Sample pixel data showing target variable (temperature) and predictor variables for urban form metrics.



(a) LST Map snippet. (b) Pixel Grid (30m). (c) Tree cover. (d) Green areas / Water.

(e) Buildings (by roof type). (f) Roads. (g) Sidewalks.

Figure 1. Conceptual illustration of the pixel characterization process. (a) Land Surface Temperature map for a sample area. (b) The 30m analysis pixel grid. (c-g) Different urban feature layers overlaid: (c) Tree cover, (d) Green areas and water bodies, (e) Buildings coloured by roof material, (f) Road surfaces, (g) Sidewalk surfaces. The methodology quantifies the area/presence of each feature within every pixel.

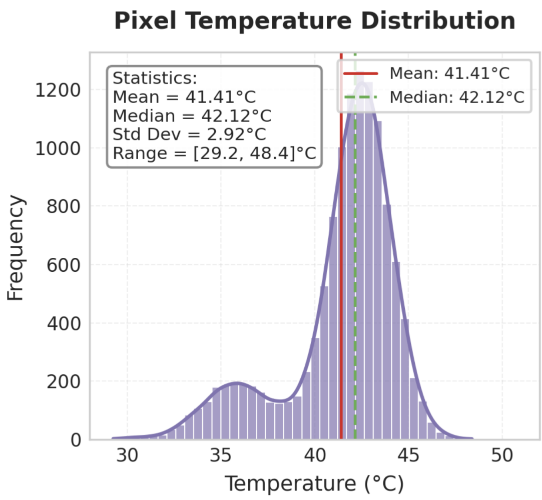



Figure 2. Distribution of land surface temperature showing bimodal pattern (mean=41.41°C, range=29.2-48.4°C).

into which urban elements most strongly influence surface temperature. Green infrastructure elements showed the strongest cooling effect, with green area having the strongest negative correlation (-0.60), followed by tree height (-0.49) and tree count (-0.31). Among the warming factors, concrete roofs exhibited the strongest positive correlation (0.34), followed by sidewalks (0.25) and road areas (0.24). Metal roofs also showed a moderate warming influence (0.19), while tile roofs, water bodies, and sports courts showed minimal cooling effects.

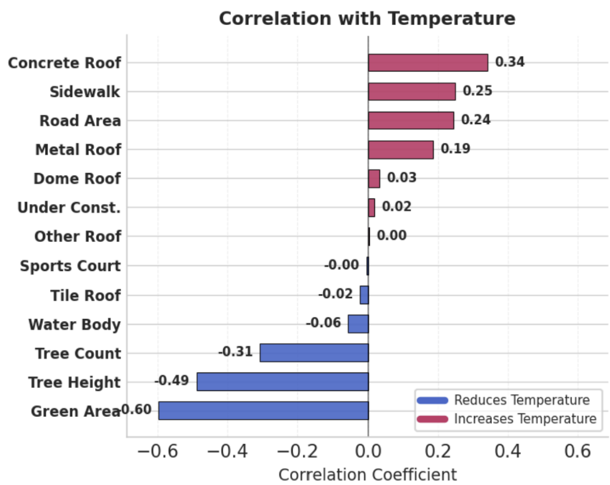



Figure 3. Pearson correlation coefficients between urban form variables and land surface temperature. Blue bars indicate cooling factors (negative correlation), while red bars show warming factors (positive correlation).

To explore these relationships in greater depth, we categorized temperature values into five distinct levels using quintiles, as shown in Table 2. This categorization provided a foundation for examining how different urban elements vary across temperature thresholds.

| | Temperature | |
|---|---|---|
| **Level** | **Min (°C)** | **Max (°C)** |
| Level 1 (Very low) | 29.24 | 40.56 |
| Level 2 (Low) | 40.56 | 41.74 |
| Level 3 (Medium) | 41.74 | 42.49 |
| Level 4 (High) | 42.49 | 43.24 |
| Level 5 (Very high) | 43.24 | 52.82 |

Table 2. Temperature threshold categories based on quintile analysis.

Using these temperature levels, we analyze how key urban form variables change across the thermal gradient. Figure 4(a) reveals a clear inverse relationship between the average height of the tree and the temperature levels. Pixels with the lowest temperatures (Level 1) have trees averaging 10.48 m in height, more than twice the average height (4.73 m) found in the hottest pixels (Level 5). Similarly, Figure 4(b) shows that cooler areas have nearly twice as many trees per pixel on average (4.10 trees in Level 1) compared to the hottest areas (2.14 trees in Level 5), reinforcing the cooling effect of urban trees. The most pronounced pattern is shown in Figure 4(c), where the coolest zones contain over five times more green space (384.19 m²) than the hottest zones (68.78 m²). This substantial contrast highlights the strong cooling influence of vegetated surfaces in urban settings.

Examining building materials, Figure 5 shows that both the total roof area and the proportion of concrete roofs increase with temperature levels. The coolest areas (Level 1) have approximately 175 m² of total roof coverage, while the hottest areas (Level 5) contain nearly 440 m². Concrete roofs, which have high thermal mass and absorption, constitute the largest proportion across all temperature levels, particularly in hotter areas.

These exploratory visualizations provide valuable intuition about the dataset before proceeding to the modeling phase. The clear patterns observed across temperature levels validate our approach of examining urban form elements at the pixel level and inform expectations for the machine learning model. In particular, the consistent inverse relationship between green infrastructure metrics and temperature levels, coupled with the positive association between impervious surfaces and temperature, suggests that these will be dominant factors in predicting LST.

## 2.5 Modeling

To investigate the quantitative relationships identified during EDA and build a predictive model, a machine learning regression approach was employed. Specifically, the XGBoost (Extreme Gradient Boosting) algorithm was selected, known for its high predictive performance, computational efficiency, and built-in regularization techniques that help prevent overfitting (Chen and Guestrin, 2016).

Before modeling, we identified which pixels belonged to each of our target neighborhoods (Colonia Jardines Alcalde, Colonia Panorámica de Huentitán, and Zona Centro). Using GeoPandas, we performed a spatial join between our pixel grid and neighborhood polygon boundaries, enabling us to filter the dataset from Table 1 to include only pixels within these specific urban zones. This critical step allowed us to apply our methodology to each

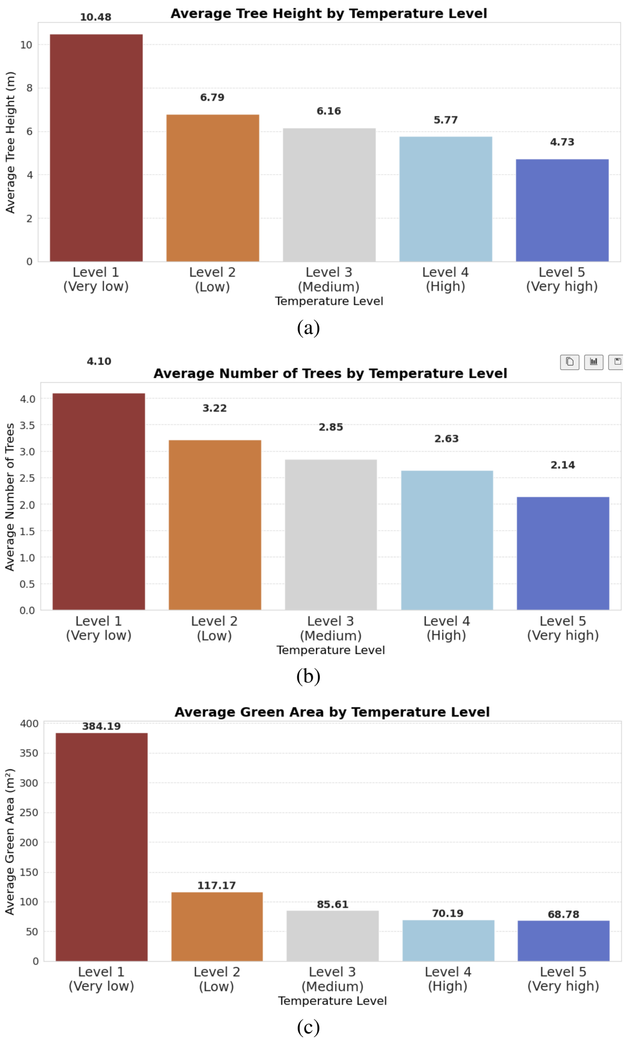



Figure 4. (a) Average tree height across temperature levels, showing taller trees in cooler pixels and progressively shorter trees as temperature increases. (b) Average number of trees per pixel across temperature levels, showing higher tree density in cooler areas. (c) Average green area across temperature levels, showing the dramatic decrease in vegetated surfaces as temperatures increase.

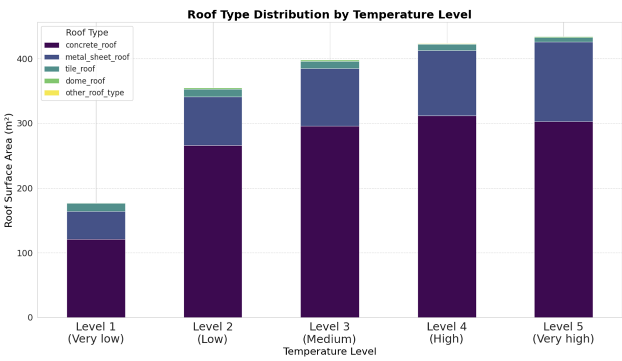



Figure 5. Distribution of roof types across temperature levels, showing increasing roof area and predominance of concrete roofs in hotter pixels.

selected neighborhood independently, enabling comparison of thermal dynamics across different urban morphologies within Guadalajara.

The filtered dataset for each neighborhood, comprising the LST (target variable) and predictor variables for each 30 m pixel as structured in Table 1, served as input for the modeling process. The data was split into training (80%) and test (20%) sets using random sampling.

Hyperparameter optimization was performed exclusively on the training set using Randomized Search Cross-Validation (RandomizedSearchCV). This process efficiently explored the hyperparameter space (learning rate, number of estimators, maximum depth, subsampling rates) using 5-fold cross-validation within the training data to identify optimal settings based on minimizing RMSE.

The final model performance was evaluated by training an XGBoost model with the optimal hyperparameters on the full training set and then evaluating it on the held-out test set. The $R^2$ and RMSE values reported in Section 3 represent this test set performance, ensuring unbiased estimates of model generalization ability. Modeling tasks were implemented in Python, using the XGBoost library and Scikit-learn for cross-validation and hyperparameter search utilities.

## 3. Results and Discussion

This section provides a detailed analysis of neighborhood-specific models constructed to examine urban thermal dynamics on the pixel scale. The analysis focuses on three neighborhoods within Guadalajara each characterized by distinct urban morphologies and thermal attributes. This diversity facilitates the evaluation of how relationships between urban elements and land surface temperature vary across different urban settings. Following the modeling approach described previously, thermal model results are analyzed individually for each neighborhood, followed by a comparative assessment across all studied areas, and conclude with practical implications for urban planning and thermal management strategies.

### 3.1 Thermal model results by neighborhood

The neighborhoods analyzed—Jardines Alcalde, Zona Centro, and Panorámica de Huentitán—exhibit distinctive thermal signatures and model performance metrics, reflecting the complex interaction between urban form elements and surface temperature at the fine-grained pixel level. The following subsections detail the results for each area, highlighting the key urban features that drive thermal dynamics in each specific context.

| Urban Form Variable | Importance |
|---|---|
| Green area | 0.3322 |
| Water body | 0.1486 |
| Road area | 0.1197 |
| Metal sheet roof | 0.0877 |
| Average tree height | 0.0817 |
| Sidewalk area | 0.0396 |
| Concrete roof | 0.0393 |
| Under construction | 0.0363 |
| Sports court | 0.0343 |
| Total trees | 0.0297 |

Table 3. Feature importance ranking for Jardines Alcalde neighbor model.

**3.1.1 Jardines Alcalde** The model for Jardines Alcalde exhibits strong predictive performance, explaining approximately 83% of the LST variability ($R^2 = 0.8290$) and achieving high precision (RMSE = 0.5115°C).

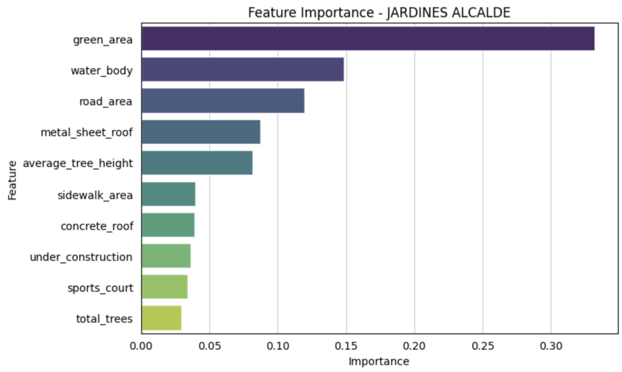



Figure 6. Feature importance ranking for Jardines Alcalde neighbor, showing the relative influence of each urban form variable on land surface temperature prediction.

As shown in Figure 6 and Table 3, green infrastructure features play a dominant role in shaping thermal dynamics in Jardines Alcalde. Green area emerges as the most influential predictor (0.3322), more than twice as important as the second-rank feature. Water bodies rank second (0.1486), likely reflecting local water features that provide cooling effects, followed by road area (0.1197), metal sheet roofing (0.0877), and average tree height (0.0817).

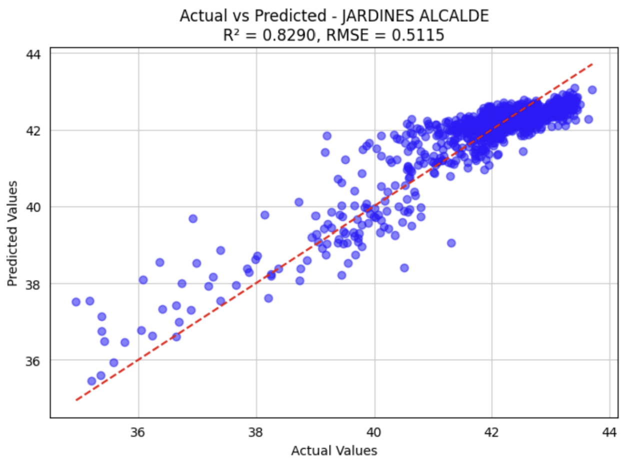



Figure 7. Actual vs. predicted temperature values for Jardines Alcalde ($R^2 = 0.8290$).

Figure 7 demonstrates the model's predictive accuracy, with points clustering tightly around the diagonal line. The temperature range spans approximately 35.5°C to 44°C, with most observations concentrated between 40°C and 43°C.

Figures 8 and 9 provide model diagnostics. The residuals show uniform distribution around zero with no systematic bias, and the histogram confirms normality (mean = -0.0004). Most errors fall within ±1°C, supporting the model's validity and indicating that selected features adequately capture LST influencing factors in this neighborhood.

**3.1.2 Zona Centro** The model for Zona Centro achieved $R^2 = 0.7043$ and RMSE = 0.6526°C. Although slightly lower than Jardines Alcalde, the model remains robust considering the greater complexity and urban density of the historical city center.

As shown in Figure 10 and Table 4, the feature importance profile differs from Jardines Alcalde. While green area remains the

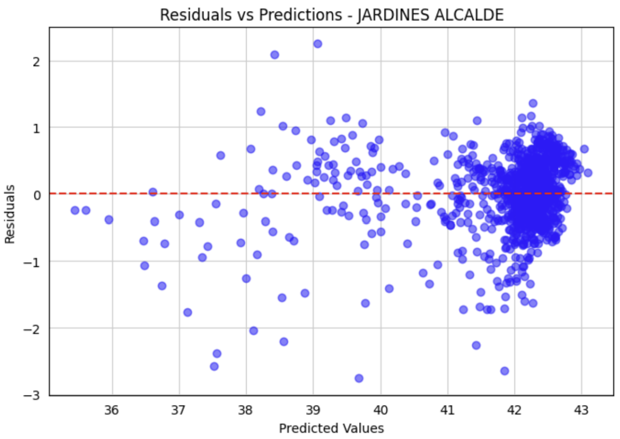



Figure 8. Residuals versus predicted values for Jardines Alcalde neighbor model, showing uniform scatter around zero with no apparent patterns.

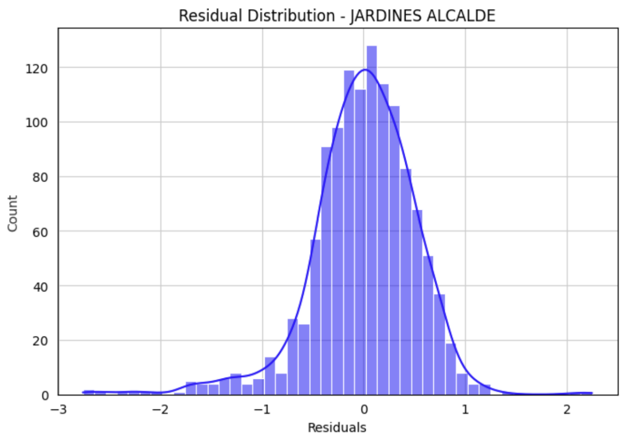



Figure 9. Distribution of model residuals for Jardines Alcalde neighbor, showing an approximately normal distribution centered near zero.

| Urban Form Variable | Importance |
|---|---|
| Green area | 0.2640 |
| Average tree height | 0.2582 |
| Concrete roof | 0.0799 |
| Road area | 0.0579 |
| Sidewalk area | 0.0563 |
| Water body | 0.0501 |
| Metal sheet roof | 0.0499 |
| Dome roof | 0.0410 |
| Under construction | 0.0378 |
| Total trees | 0.0377 |

Table 4. Feature importance ranking for Zona Centro neighbor model.

most important predictor (0.2640), average tree height closely follows (0.2582), suggesting that in the dense urban fabric, vertical vegetation structure becomes almost equally important as horizontal green coverage. Concrete roofs emerge as the third factor (0.0799), significantly higher than in Jardines Alcalde, reflecting their prevalence in the historical center. Water bodies show reduced importance, dropping to sixth place (0.0501).

The residual plots confirm normally distributed errors with mean near zero (-0.0008) and most residuals within ±1.5°C. The actual vs. predicted plot shows good agreement, with observations concentrated in the 40-44°C range, indicating the generally warmer conditions of this densely built urban center.

| Urban Form Variable | Importance |
|---|---|
| Average tree height | 0.2510 |
| Concrete roof | 0.1923 |
| Total trees | 0.1064 |
| Green area | 0.0948 |
| Sidewalk area | 0.0914 |
| Tile roof | 0.0631 |
| Metal sheet roof | 0.0588 |
| Road area | 0.0550 |
| Water body | 0.0341 |
| Under construction | 0.0306 |

Table 5. Feature importance ranking for Panorámica de Huentitán neighbor model.

#### 3.1.3 Panorámica de Huentitán

The model for Panorámica de Huentitán shows distinct performance with moderate explanatory capacity ($R^2 = 0.6937$), higher error (RMSE = 1.5921°C), and greater thermal variability (CV = 7.62%). These results reflect the neighborhood's heterogeneous topography and transitional urban development characteristic of peripheral metropolitan areas.

The feature importance pattern (Figure 11 and Table 5) reveals the most distinctive thermal dynamics among neighborhoods. Average tree height emerges as the dominant factor (0.2510), followed by concrete roof area (0.1923), which shows much greater importance than in other neighborhoods. Notably, green area ranks only fourth (0.0948), suggesting that in this peripheral neighborhood with steeper terrain, vertical vegetation structure exerts greater influence than horizontal green coverage.

Residuals show wider distribution (-6°C to +6°C) with normal distribution centered near zero (mean = 0.0024). The actual vs. predicted plot reveals broader temperature range (34-46°C) and greater scatter, consistent with higher RMSE. This wider thermal variability reflects the neighborhood's location at the urban periphery, where built environment transitions to natural landscape.

### 3.2 Comparative analysis between neighborhoods

The comparative analysis across the three neighborhoods highlights distinct patterns in model performance and urban form variable influence on land surface temperature. Jardines Alcalde yielded the highest predictive performance ($R^2 = 0.8290$, RMSE = 0.5115°C), followed by Zona Centro and Panorámica de Huentitán. These variations reflect differences in urban morphological complexity, where more homogeneous areas exhibit more predictable thermal patterns.

Feature importance rankings reveal how LST determinants differ across urban contexts. Green infrastructure consistently acts as a significant cooling factor, though its relative influence varies by local characteristics. In Jardines Alcalde, horizontal green coverage dominates thermal regulation; in Zona Centro, both green area and tree height are equally influential; in Panorámica de Huentitán, tree height supersedes green area importance, while building materials exert stronger warming effects.

The elevated RMSE in Panorámica de Huentitán indicates that peripheral areas with greater heterogeneity in topography and development may require more sophisticated modeling approaches. Despite variations in form (horizontal coverage versus vertical structure), green infrastructure's consistent identification as a dominant cooling factor confirms its universal importance in urban thermal regulation, with cooling efficiency depending heavily on surrounding urban context—a finding with significant implications for targeted greening strategies.

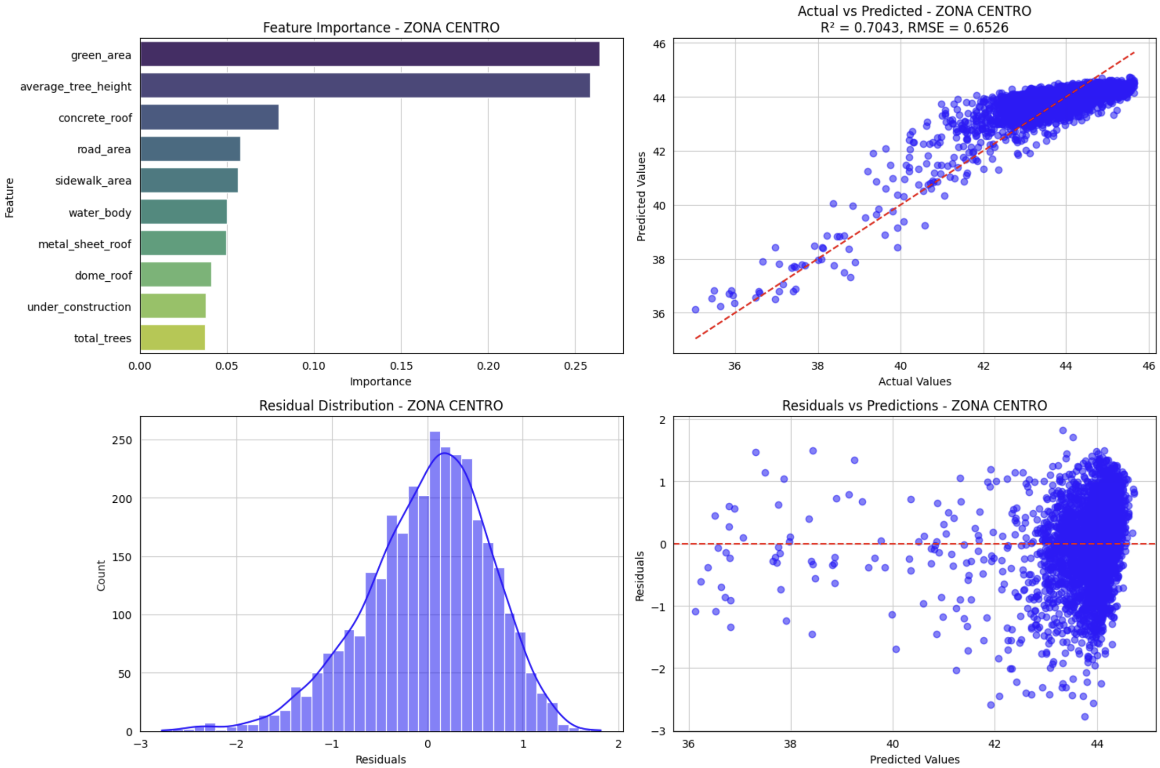



Figure 10. Diagnostic plots for Zona Centro neighbor model showing: (top left) feature importance ranking, (top right) actual vs. predicted values, (bottom left) residual distribution, and (bottom right) residuals vs. predictions scatter plot.

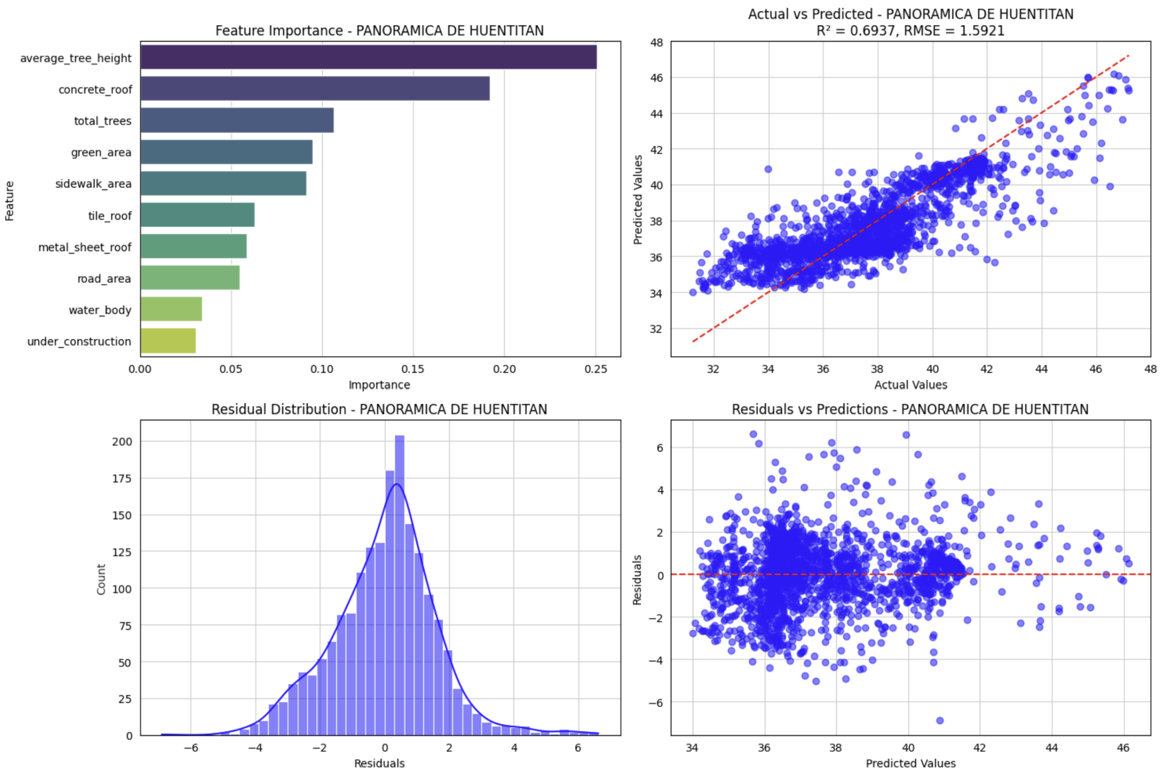



Figure 11. Diagnostic plots for Panorámica de Huentitán neighbor model showing: (top left) feature importance ranking, (top right) actual vs. predicted values, (bottom left) residual distribution, and (bottom right) residuals vs. predictions scatter plot.

### 3.3 Practical implications for urban planning

The neighborhood-specific findings provide practical guidelines that urban planners can directly apply when designing new developments or retrofitting existing urban spaces to optimize thermal conditions.

In residential neighborhoods (similar to Jardines Alcalde), priority should focus on expanded horizontal green coverage and strategic water feature placement. For dense urban centers (similar to Zona Centro), effective thermal management requires balanced approaches combining horizontal green coverage with vertical vegetation structure, plus reduction of concrete roof areas. In peripheral areas (similar to Panorámica de Huentitán), interventions should emphasize conservation and planting of taller tree species while carefully regulating concrete roof implementation.

These strategies are quantitatively supported by model results. For example, increasing green area by 300 m² within a pixel could reduce surface temperatures by up to 5°C in residential areas, while in dense centers, combining 150 m² of green area with 2-3 meter increases in tree height achieves similar cooling effects. The quantitative insights support evidence-based decision-making in urban planning, aligning with Guadalajara's climate adaptation objectives outlined in the Metropolitan Climate Action Plan (Gobierno de Guadalajara and C40 Cities and IMEPLAN and Gobierno de Jalisco, 2020).

## 4. Conclusions

The study highlights the advantages of pixel-level thermal analysis in urban environments, offering high spatial granularity for understanding the relationships between urban form and land surface temperature. The quantification of diverse urban features at the 30-meter thermal pixel scale enables the transformation of intuitive notions of urban heat dynamics into a data-driven analytical framework with direct applications for evidence-based planning.

The high-resolution methodology revealed distinct thermal patterns across neighborhoods in Guadalajara, with significant variations in the importance of urban elements. Green infrastructure consistently functioned as a dominant cooling factor, though its influence varied by context: horizontal green coverage showed greatest impact in residential areas, while tree height became equally important in dense urban cores, and building materials exerted stronger influence in peripheral zones. These findings underscore the importance of context-sensitive approaches for effective urban thermal regulation.

From a practical perspective, this research provides planners with a quantitative framework to predict thermal outcomes of design interventions. For example, increasing green coverage by approximately 300 m² within a thermal pixel may reduce surface temperatures by up to 5°C, depending on the surrounding urban context.

The methodology exhibits strong predictive performance ($R^2$ = 0.69-0.83), though higher prediction errors in peripheral areas indicate that incorporating additional environmental variables may improve model accuracy in complex urban environments. The pixel-level approach provides the foundation for operationalizing Guadalajara's Metropolitan Climate Action Plan objectives, enabling targeted interventions that enhance cooling efficiency and optimize resource allocation across the metropolitan region.

## References

Camacho Sandoval, T., Magaña Rueda, V. O., Ramos de Robles, S. L., Gran Castro, J. A., 2024. Vegetation as a Regulator of Urban Climate: the Case of the Guadalajara Metropolitan Zone, Jalisco, Mexico. *Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM*, 114, e60849.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 785–794.

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., Bechtel, B., 2022. A global map of Local Climate Zones to support earth system modelling and urban scale environmental science. *Earth System Science Data*, 14(8), 3835–3873.

Fideicomiso para la Administración del Programa de Desarrollo Forestal del Estado de Jalisco, 2024. Sistema de información territorial de jalisco (site jalisco). `https://fiprodefo.jalisco.gob.mx/`. Datasets including Green Areas and Water Bodies polygons.

Gobierno de Guadalajara, 2024. Inventario arbolado lidar. `https://visorurbano.guadalajara.gob.mx/lidar`.

Gobierno de Guadalajara and C40 Cities and IMEPLAN and Gobierno de Jalisco, 2020. Plan de acción climática del Área metropolitana de guadalajara (pacmetro). Technical report, Instituto Metropolitano de Planeación del Área Metropolitana de Guadalajara (IMEPLAN), Guadalajara, Jalisco, México.

Hernández-Pérez, I., Álvarez, G., Xamán, J., Zavala-Guillén, I., Arce, J., Simá, E., 2014. Thermal performance of reflective materials applied to exterior building components—A review. *Energy and Buildings*, 80, 81–105.

Instituto de Información Estadística y Geográfica de Jalisco, 2024. Información cartográfica y estadística. `https://iieg.gob.mx/ns/?page_id=161`. Datasets including Cadastre (Buildings, Roof Materials), Sidewalks, and Road Network.

Kubota, T., Miura, M., Tominaga, Y., Mochida, A., 2020. Characterising thermal behaviour of buildings and its effect on urban heat island in tropical areas. *International Journal of Energy and Environmental Engineering*, 11(1), 85–100.

Morales Manilla, L. M., Hernández Magaña, A. I., Arizaga Pérez, S., Coba Pérez, P. C., García Martínez, C. B., Vargas Ramírez, N., Ramírez de Arellano, F., 2018. Inventario del arbolado público urbano de la zona metropolitana de guadalajara: Informe final. Technical report, Centro de Investigaciones en Geografía Ambiental, Universidad Nacional Autónoma de México (CIGA-UNAM), Morelia, Michoacán, México.

Santamouris, M., 2017. The urban heat island effect, its causes, and mitigation, with reference to the thermal properties of asphalt concrete. *Journal of Cleaner Production*, 156, 244–257.

Sun, Y., Gao, C., Li, J., Wang, R., Liu, J., 2019. Quantifying the Effects of Urban Form on Land Surface Temperature in Subtropical High-Density Urban Areas Using Machine Learning. *Remote Sensing*, 11(8), 959.

USGS, 2024. Landsat 8-9 collection 2 level 2 science products. `https://doi.org/10.5066/P9OGBGM6`.

USGS EROS, 2023. Landsat 8-9 Calibration and Validation Algorithm Description Document. Technical Report LSDS-1747, Version 5.0, Earth Resources Observation and Science Center, Sioux Falls, South Dakota.