Where Heat Meets Density

Exploring the Urban Heat Island Effect in New York City

Xiaoyue Yan
yanxy277@mit.edu
Department of Architecture
SMArchs Building Technology

Introduction

Urban Heat Islands (UHI) occur when cities become warmer than surrounding rural or less developed areas due to dense buildings, paved surfaces, limited vegetation, and human activities. Under global warming, UHI effects have become increasingly important in large metropolitan areas, where elevated temperatures can worsen outdoor thermal comfort, increase cooling demand, and intensify heat-related health risks during extreme heat events. In New York City, uneven distributions of building density, green space, and waterfront exposure may further create strong spatial differences in urban heat exposure.

Land Surface Temperature (LST), commonly derived from satellite remote sensing, is widely used to map UHI patterns. Compared with ground-based weather stations, satellite-based LST provides broader spatial coverage and captures fine-scale thermal variations across different land covers and urban forms. In this project, LST is used as an indicator of urban heat intensity to visualize heat distribution across New York City and examine its relationship with physical urban environment parameters.

Methodology

This project used a grid-based spatial workflow to examine how urban heat varies with building density and other urban-environmental conditions across New York City. Geospatial data processing and visualization were conducted using QGIS and GeoPandas in Python.

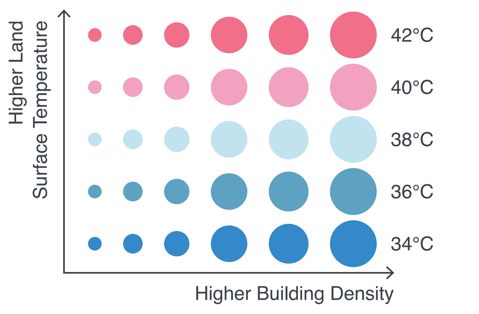
Step 1: Spatial unit and data integration.
All datasets were organized at a 200 m grid-cell level. The LST data, built environment variables, visual environment indicators, natural environment indicators, and socio-demographic variables were joined by a common cell ID. All spatial layers were projected to a consistent coordinate system to ensure reliable spatial overlay and mapping.

Step 2: Urban heat mapping.
From the July 2024 LST dataset, all records corresponding to 14:00 local time were selected and averaged for each 200 m grid cell. The main map uses color to represent land surface temperature, while circle size represents building density.

Step 3: Machine-learning analysis.
To further examine the relationship between urban heat and urban form, an XGBoost regression model was trained using the grid-level mean 2 PM LST as the outcome variable. Predictors included transportation accessibility, road density, building density, park density, POI density, population and income indicators, visual-environment variables, and Normalized Difference Vegetation Index (NDVI). Higher NDVI values indicate denser and healthier vegetation, which is generally associated with stronger local cooling effects.

Step 4: Model interpretation with SHAP.
The dataset was split into training and testing subsets, and model performance was evaluated using RMSE, MAE, and R^2. SHAP values were then calculated to identify the most influential predictors and to visualize nonlinear relationships between each urban-environmental variable and afternoon urban heat.

Map Description

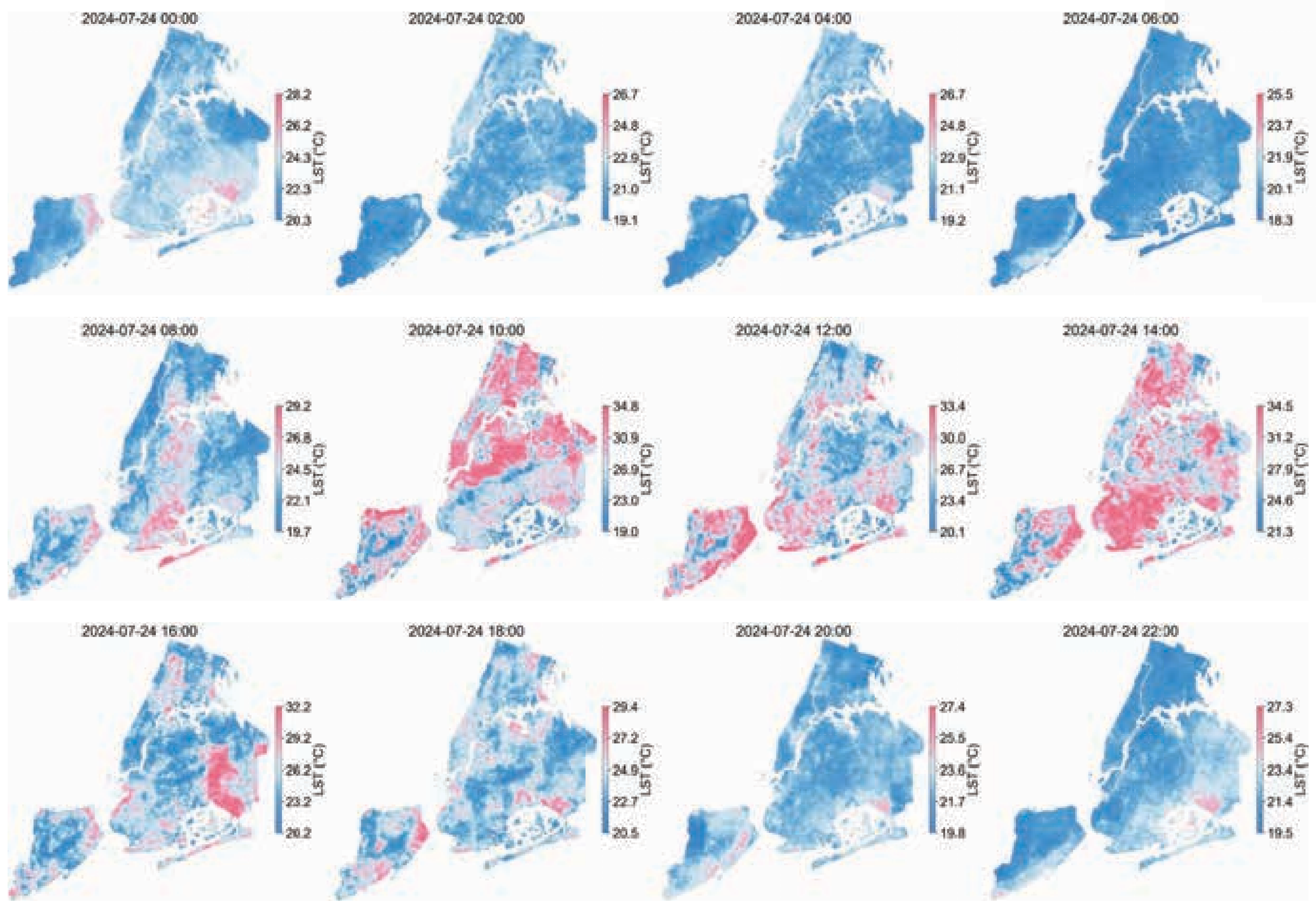


Circle color: Mean Land Surface Temperature of July 2024, 14:00.
Circle size: building density.

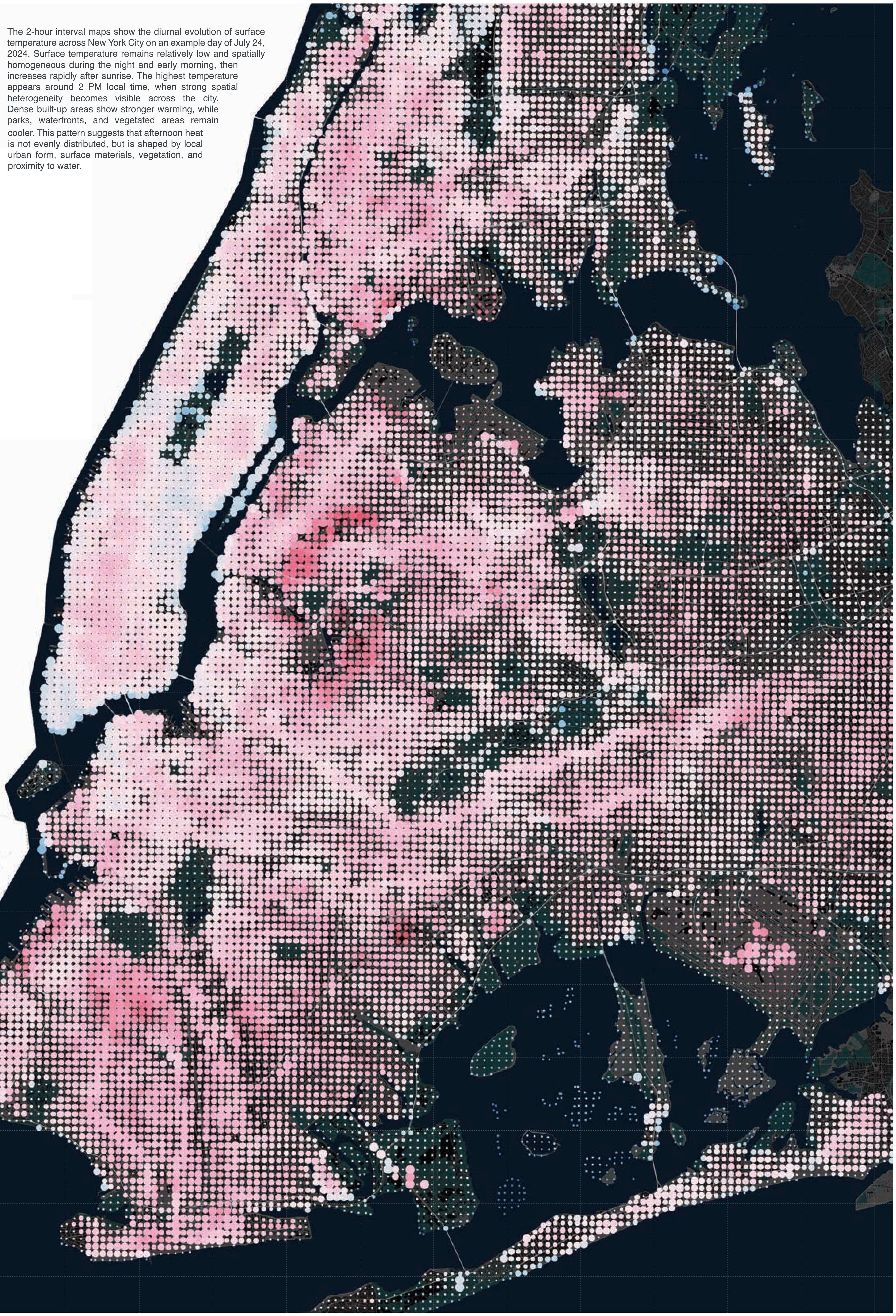
Warmer colors indicate higer land surface temperature, and larger circles indicate denser built environments.

Hotter areas generally overlap with higher building density. Long Island City appears as one of the most prominent heat hotspots, while Manhattan is not uniformly the hottest despite its high density. In contrast, areas with concentrated urban greenery, such as large parks and vegetated waterfront zones, show visibly lower temperatures, highlighting the cooling role of vegetation and open space.

Daily evolution of Land Surface Temperature



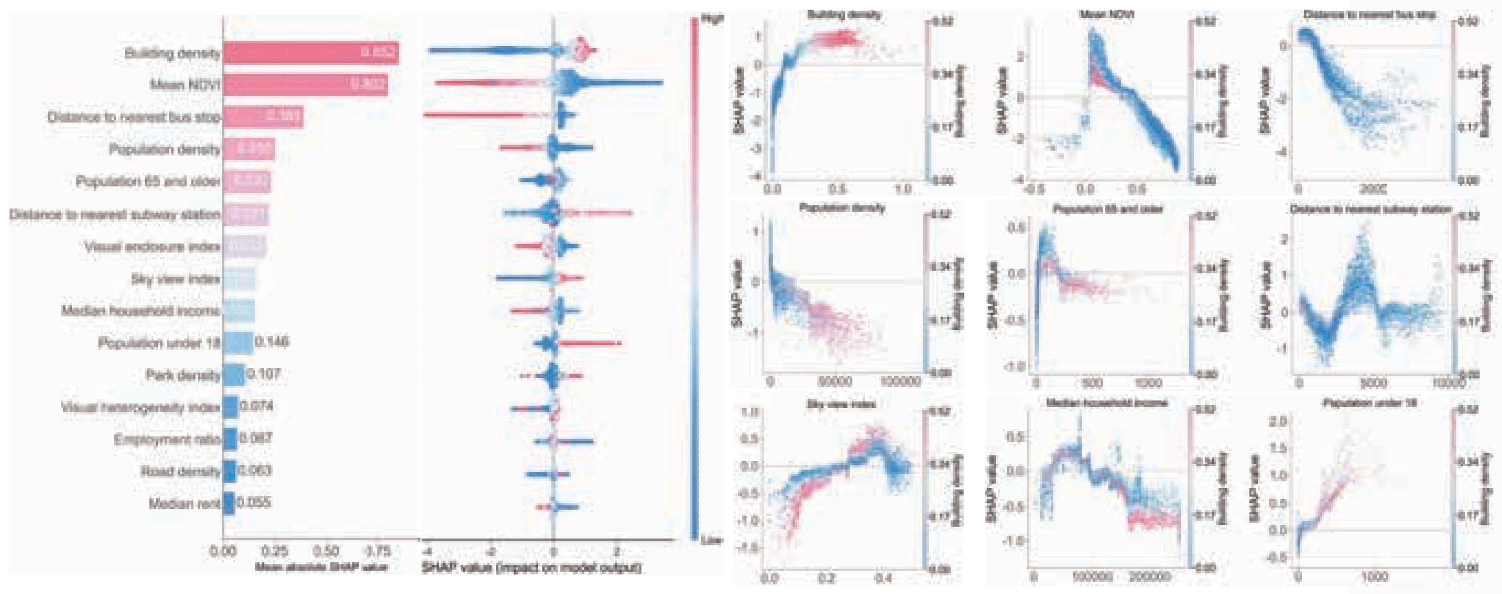
The 2-hour interval maps show the diurnal evolution of surface temperature across New York City on an example day of July 24, 2024. Surface temperature remains relatively low and spatially homogeneous during the night and early morning, then increases rapidly after sunrise. The highest temperature appears around 2 PM local time, when strong spatial heterogeneity becomes visible across the city. Dense built-up areas show stronger warming, while parks, waterfronts, and vegetated areas remain cooler. This pattern suggests that afternoon heat is not evenly distributed, but is shaped by local urban form, surface materials, vegetation, and proximity to water.



Factors associated with the urban heat island effect

Building density is the most important predictor, with higher density generally increasing the predicted heat level, especially at low-to-moderate density ranges before the effect gradually levels off. Mean NDVI is the second most influential factor and shows a strong cooling association: higher vegetation greenness corresponds to lower predicted heat.

Transit-accessibility variables, especially distance to the nearest bus stop, also rank highly, likely reflecting broader urban intensity and centrality rather than a direct causal effect. Overall, the results suggest that NYC's urban heat pattern is primarily structured by the combined effects of built density, vegetation, and local urban form.



Data Source

Land Surface Temperature
USGS Landsat 8–9 Collection (https://earthexplorer.usgs.gov/)
ECOSTRESS Level-2 Land Surface Temperature and Emissivity (https://appeears.earthdatacloud.nasa.gov/)

Built environment
New York Major Roads; New York City Parks; New York City Shoreline; New York City Building Footprints; New York City Street Centerline; New York City Open Space;New York City Census Blocks, New York City, State Assembly Districts (https://geodata.libraries.mit.edu/)

Visual environment
Google Street View (https://www.google.com/maps)

Natural environment
USGS Landsat 8–9 Collection (https://earthexplorer.usgs.gov/)

Socio-demographic factors
American Community Survey (https://www.census.gov/en.html)