

Volume 2, Issue 3

Research Article

Date of Submission: 20 August, 2026

Date of Acceptance: 15 September, 2026

Date of Publication: 23 September, 2026

Temperature Prediction with Satellite Imagery Features of New York City Using Machine Learning

Junrong Wang^{1*}, **Gladys Elizabeth Kenyon²**, **Lingbo Liu³** and **Shuming Bao⁴**

¹The Bartlett Centre for Advanced Spatial Analysis, University College London, UK

²University of Liverpool, UK

³Center for Geographic Analysis, Harvard University, USA

⁴China Data Institute, Ann Arbor, MI, USA

***Corresponding Author:**

Junrong Wang, The Bartlett Centre for Advanced Spatial Analysis, University College London, UK.

Citation: Wang, J., Kenyon, G. E., Liu, L., Bao, S. (2026). Temperature Prediction with Satellite Imagery Features of New York City Using Machine Learning. *J AI VR Hum Comput*, 2(3), 01-14.

Abstract

This project applies satellite imagery to study the temperature changes in cities with an advanced machine learning approach. For extracting features, we employ the MOSAIKS (Multi-Task Observation Using Satellites and Kitchen Sinks) which comprises the RCF (Random Convolution Features) version which is an unsupervised learning algorithm, deciding which features to extract [1]. Then Xtreme Gradient Boosting or Random Forest are used to predict temperature from features. Specifically, two RCF architectures and multiple regression methods are compared systematically to evaluate overall performance and interpret-ability, contributing to Artificial Intelligence. Multiple methods performs differently in various image sizes, revealing a new relationship between methods and image sizes, and guiding method choices for specific temperature prediction. Existing satellite-based temperature studies rely on physics-based LST (Land Surface Temperature) retrieval algorithms (mono-/split-window) combined with empirical regressions and sensor data, unsuitable of prediction in design [2,3]. Some research on machine-learning (Random Forest, gradient boosting, Decision Tree, Artificial Neural Network, Convolutional Neural Networks, Fast Library for Approximate Nearest Neighbors) depend on complex sensor data inputs, unsuitable for prediction from just imagery [4-6]. Despite of improved accuracy, they are computationally demanding, offering limited interpretability, whereas more transparent regression models underperform and hardly scale globally. In contrast, our approach requires only imagery and achieves relatively high accuracy [7,8]. More importantly, using our methods, solely satellite imagery can yield reasonable results, useful for designers to generate temperature predictions from fake satellite imagery they design, benefiting engineering.

Keywords: Feature Extraction, Heat Island Effect, Land, Surface Temperature, Machine Learning, Remote Sensing

Introduction

Over the last three decades, metropolitan areas have seen significant urban development which has had a considerable impact on the climates of local and regional contexts. According to the latest data, in 2016, more than half of the world's population were living in urban areas. This percentage is expected to rise significantly by the year 2030 [9]. As a result of the Urban Heat Island (UHI) effect which occurs in densely populated urban areas, man-made surfaces retain excessive heat from the natural landscapes and release it intermittently at night. Human health, environment, and energy consumption are impacted by the UHI [10-12].

Evaluation of the impacts of the UHI phenomenon now increasingly involves the use of remote sensing and GIS technologies, due to a lack of granular temperature data which is time consuming and costly to acquire. These techniques employ data on land surface temperature (LST) acquired with thermal infrared sensors including NOAA/AVHRR, TERRA/MODIS, and Landsat [13]. Landsat 8's thermal infrared sensor (TIRS) improves the LST estimation by using SW techniques [14]. Despite certain improvements in the measurement of temperature, one can predict results for artificially created spaces in landscape and urban design projects from using landscape patterns as features which can provide guidance for landscape architecture design [15].

New York City (NYC) is an ideal example of a densely urbanized area where understanding the impact of green and blue spaces on the urban thermal environment is well understood. The problems related to the UHI effect are particularly significant in New York City because it is one of the most densely populated cities in the world [11,15]. The fact that the characteristics of an urban landscape in New York City have a strong relationship with changes in temperature suggests that NYC is appropriate for analyzing the interdependence of satellite image characteristics and temperature.

New York City is characterized by a high UHI effect, where the built structures act as heat sinks and the inability to store moisture results in higher temperatures in the lower layer of the atmosphere than in the surrounding countryside [16]. When the UHI effect is not present, the mass air movement towards the front part of New York City is significantly impeded in the central metropolitan part. This can be attributed to the fact that the city area surface is relatively rougher than the surrounding region which results in more friction and therefore limits air circulation [16]. When full urban heat islands are developed, the speed of the front decreases over the upwind part of the city and sharply increases over the downwind half. Predictions also show that New York City is expected to experience increased mortality due to heat than other cities including Boston and Philadelphia in the 2020s, 2050s, and 2080s [17]. To evaluate the effectiveness of the measures, including the urban forestry, green roofs, and high albedo surfaces, a simulation was carried out to determine the possibility of reducing the scale of the urban heat island and lowering the air temperature in New York City. The best strategy was to optimize the vegetation [16].

This research employs the environmental and satellite imagery features of New York City (NYC) with specific machine learning techniques to measure local temperature variations and the varied UHI effect. The proposed method is an efficient approach to temperature prediction that is useful when city planning and designing or during any activity implying temperature forecast based on satellite data.

Literature Review

Impact of Urban Morphology on Thermal Environment

The thermal environment is significantly influenced by urban morphology. The structure of the metropolitan area contributes to the thermal conditions during the night [18]. It is also crucial that land surface temperatures rely on aspects such as the density, height, and proximity of vegetation and structures [19]. Studies on the effects of city morphology on temperature conditions in cities across the globe strengthens the significance of targeted measures to relieve UHI in cities [20].

Urban green and blue spaces play a crucial role for regulating temperatures. Present research shows that the positioning of green space in urban parks affects the temperatures of the neighbourhoods [21]. Wetlands have cooling effect on urban environment [22]. It also affects the land surface temperature as evidenced by the three-dimensional building structures and green spaces interaction [23].

Machine Learning in Environmental Feature Extraction

Techniques of machine learning are effective in the extraction of topographical features from satellite images and the study of thermal processes in cities [23]. The application of Sentinel-2A imagery along with several machine learning models such as random forest, SVM, Naïve Bayes, Decision Tree (CART) in numerous studies prove the effectiveness of such approaches in urban environment analysis [24]. Weather data interpolated to resolutions of thousands of meters together with machine learning algorithms help to understand the effects of climate changes when studied at the city level [24].

Machine learning algorithms are particularly suitable in identifying geographical patterns from satellite images and thus improving our understanding of the thermal conditions of urban landscapes. Other studies carried out on satellite data and several machine learning algorithms have demonstrated the potential of these techniques to identify urban features. Yuan, Zhou, Dang, Sun, Hu, and Mu used Boosted Regression Tree to investigate impacts of 3D building features and urban green space on land surface temperature [23]. Qing used Convolutional Neural Network to extract visual features of Urban Stress Vitality [25]. The incorporation of highly detailed weather information in machine learning algorithms provides useful information on the impact of climate change in cities [26].

Integration of 2D Features in Urban Studies

The integration of 2D features in the consideration of thermal conditions in cities introduces new approaches to understanding urban structures. Previous work has explored building and vegetation density affecting LST and has given insights into better urban thermal environments [27]. The application of boosted regression tree methods for

modeling of intensities of UHI evidences the role of the complex system of urban indicators in the field of environmental studies [28]. Guiyang town density is directly proportional with the Land Surface Temperature (LST). Furthermore, some variables such as height of the building, average wind velocity from the south direction, shape index of the surrounding land use and land cover also act as a mitigating factor for the LST [29,30].

Study Area and Data

Data

Satellite Imagery

Satellite imagery with a resolution of 10 meters was acquired from Sentinel-2 (<https://planetarycomputer.microsoft.com/api/stac/v1>). The band with a resolution of 10 meters is filtered and for the grids' URLs are divided in New York City Region. Taking a raw capture of the image patches from each URL will lead to the outer fringes of the raster that have been extracted from the URLs having incomplete parts. As shown in Figure 1, to avoid the creation of gaps between neighboring grids, all the raster tiff files are merged before the patching process is done.

The Spatiotemporal Asset Catalog (STAC) is used to retrieve the URL of a Sentinel-2 satellite image of New York City. This was done by using a shapefile of New York City neighborhood boundaries to specify the desired region for the satellite image retrieval. The New York City region is divided into grids of four distinct scales, with side widths ranging from 100m to 400m, incremented by 100m each time. Next, the grids are overlaid onto the satellite image to segment the image into smaller portions.

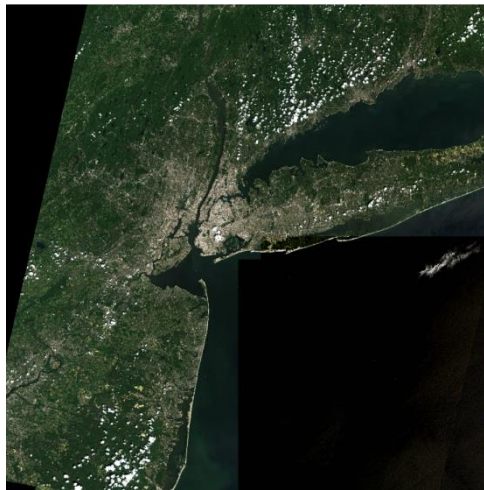


Figure 1: Merged Sentinel-2 Image of 10m Resolution

Ground Truth Temperature Data

The temperature data for New York City is sourced from the NYC Heat Map (Surface Temperature) provided by local council(newyorkcitycouncil.github.io) which shows the temperature of the site relative to the average temperature of the city. The data is predicted from the Landsat 8 satellite OLI/TIRS sensors provided by the US Geological Survey and is managed and analyzed by the Data Team of the New York City Council. The data is derived from the Landsat 8 satellite which is used by the US Geological Survey (USGS). It buys photos globally with a specified frequency of every 16 days, with a granularity of 30 meters per pixel.

The satellite data is distributed by Google Earth Engine (GEE), a cloud computing platform for the analysis of large-scale satellite data. All the Landsat 8 photos that include the New York City region in the period of June-September 2020-2022 are pulled from GEE and each image is first filtered to remove clouds and cloud shadows ensuring that only clear observations are used. The photos are cropped to the bounds of New York City, which means only data within the city is considered. The average surface temperature is then calculated for each pixel in New York City, including all the filtered images. Firstly, the average temperature over all days within the period of June-September 2020-2022 are extracted in every pixel. Then the average temperature of all locations is calculated. After that, the 'deviation from the mean temperature' is calculated. The temperature data is exported as a raster format, which is a gridded data format. The raster data is first standardized to Kelvin using scale and offset parameters obtained from GEE before being converted to Fahrenheit because the raw data on GEE are not in Kelvin but thermal bands of satellite imagery. The digital numbers from thermal bands are scaled into radiance and then converted to temperature in Kelvin using the GEE-specific metadata parameters.

The mean temperature for NYC is calculated, and then the result is subtracted from the temperature of each pixel to obtain the deviation from the mean. This underscores the extent to which each locality is either above or below average temperature of the whole city.

To strengthen neighborhood-level effects and make the data generated more evenly distributed, a moving window average is computed. This process involves averaging the temperature considering data within a square area which is approximately 0.5 miles width (27 pixels width and height). Overall, this approach provides a broader and more encompassing look at the variance of temperature in different groups of people. The temperature variances are restricted to plus/minus 8 degrees Fahrenheit to ensure that parks, aquatic bodies, and marshes do not dominate the entire map while also providing a more realistic temperature change within residential and commercial areas. The final processed data is represented on a map that shows the temperatures in the areas of New York City; it is especially important to highlight the areas that are hotter or colder than the average temperature in the city (hotter labeled as red, colder labeled as blue).

In this case, for the purpose of this study, the temperature raster is integrated into the satellite image patches through the calculation of the mean temperature within each grid based on the raster file as shown in figure 2.

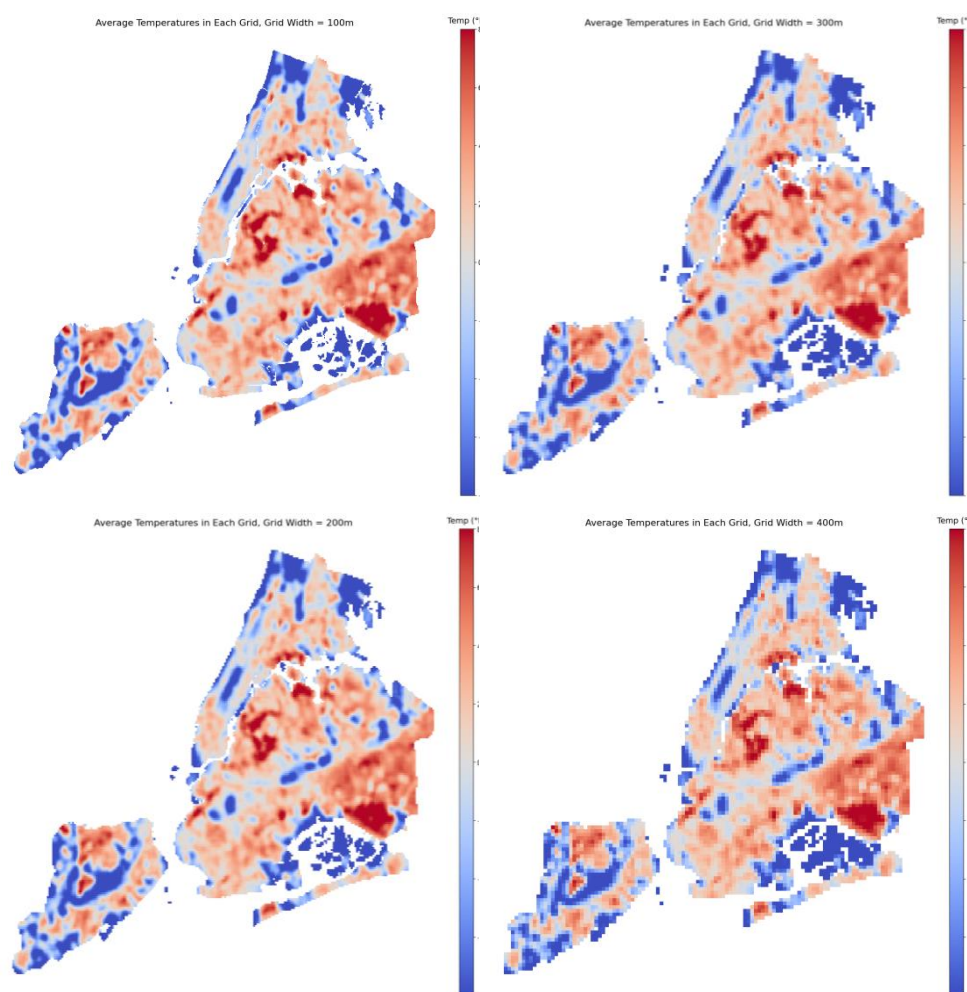


Figure 2: Relative Temperature in Each Grid in the New York City Region

Landcover of New York City

In order to assess the interpretability of the model, we utilize preprocessed landcover raster data obtained from NYC Open Data (Landcover Raster Data (2010) – 3ft Resolution | NYC Open Data (cityofnewyork.us)). This data has a resolution of 3ft which enables the determination of the proportion of green space, blue space and building space for each grid. These ratios are then used to analyze the regression of these features to the image features extracted by the model. The ratios are calculated by dividing the blue space area and building footprint area in each grid by the total area.

Methods

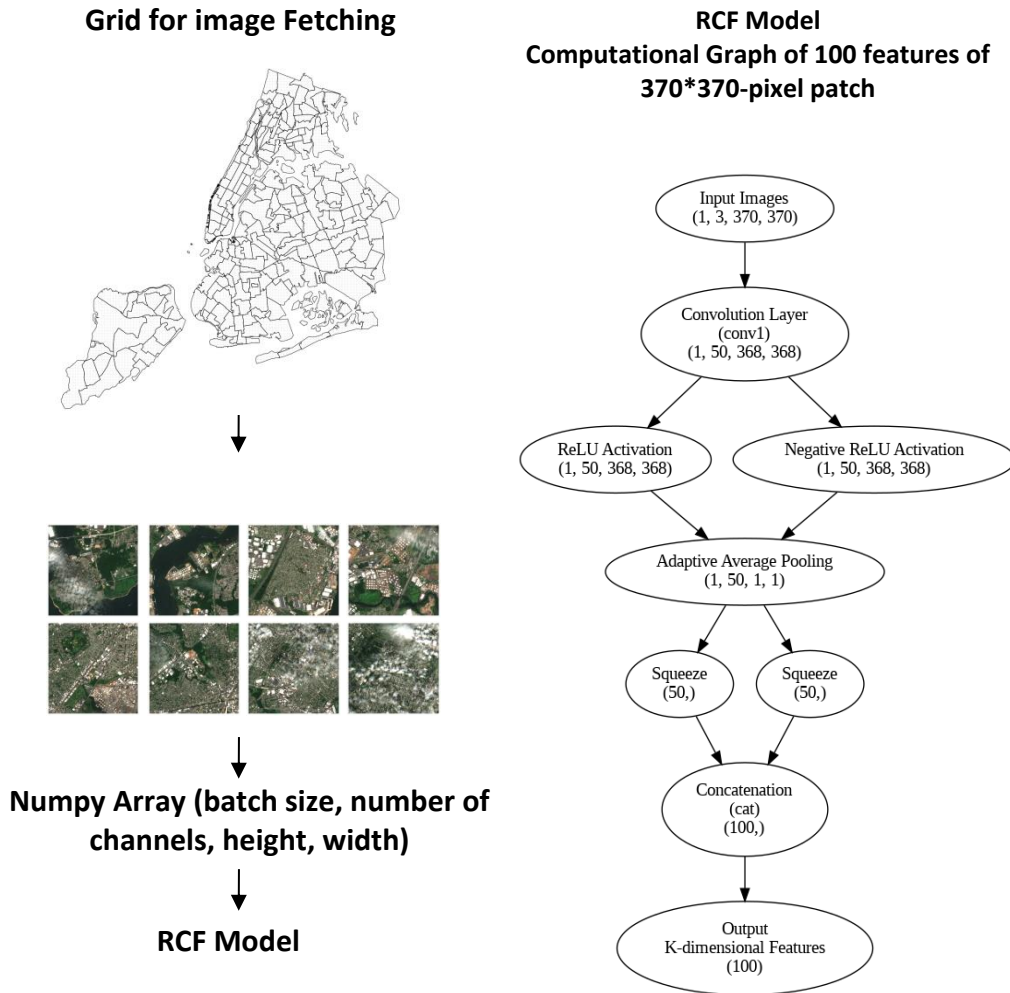


Figure 3: Flow Chart of Image Feature Extraction of New York City Orthophotos

RCF Feature Extraction Original RCF Model

As shown in Figure 3, the MOSAIKS model utilizes an RCF model to extract features from 2D satellite image patches [31,32]. The initial RCF model incorporates a 2D convolutional layer. The structure involves a ReLU (Rectified Linear Unit) activation and a negative ReLU activation. The activation function utilizes one nonlinear adjustment to the generated features thus providing the multi-layer neural network with a meaningful depth. The model employs both positive and negative ReLU activation function for once each. This methodology is to extract various characteristics originated from supplied data. The ReLU activation function is defined as follows:

$$ReLU(x) = \max(0, x)$$

The function maintains positive values and designates a value of zero to all negative values. The ReLU is usually applied in neural networks for its ability on introducing non-linearity, solve the problem of vanishing gradients, and provide computational efficiency.

The negative ReLU activation function means basically the application of ReLU on the negated input:

$$ReLU(-x) = -\min(0, x)$$

This function negates values from the negative original input into positive values, while zeroing the positive values. The traditional ReLU function gets the positive replies corresponding to cases of positive convolutional result. Conversely, the negative ReLU function captures the negative responses corresponding to cases of negative convolution result. This enables the model to retain and effectively utilize information from both types of responses, which might otherwise be lost when employing traditional ReLU activation functions. By concatenating the output of conventional and negative ReLU into $[ReLU(x), ReLU(-x)]$, the model efficiently increases the number of features, thereby obtaining a more comprehensive representation. For instance, a certain area in the image may show a significant positive activation feature, while in another area it may present as negative activation. Both forms of activation carry information value. Utilizing both positive and negative activations enables the model to fully retain the feature. The aggregated result

provides a balanced perspective in the feature space, giving equal importance to strong and weak responses. This is particularly advantageous in tasks where both positive and negative evidence jointly influence the final decision. The model x1a applies the conventional ReLU function to the convolution output, while x1b applies the negative ReLU function to the same output. By pooling and concatenating x1a and x1b separately, a result that simultaneously contains positive and negative feature responses is obtained.

To reduce the size of the feature maps generated by the convolutional layers, the model uses Adaptive Average Pooling, which is essential for compressing the information in the feature maps and ensuring that it remains the same regardless of changes of input size. This is especially useful in a model that is required to process inputs of different dimensions. Adaptive average pooling is a modified version of the conventional average pooling technique to maintain the output size. Unlike standard average pooling that uses manually preset window size and stride, the adaptive average pooling only need the user to set the desired output size. It would modify its window size and stride to generate an output of the preset dimensions, no matter what the input size is.

The formula for average pooling is:

$$Output(i, j) = \frac{1}{KH \times KW} \sum_{m=1}^{KH} \sum_{n=1}^{KW} Input(i \times s + m - P, j \times S + n - P)$$

The KH and KW are the height and width of the pooling window. P is the padding added to the input to control the spatial dimensions of the output.

In adaptive average pooling, the size of the pooling window and stride is dynamically determined to ensure that the output size is as specified. For adaptive pooling from input size (H_{in} , W_{in}) to output size (H_{out} , W_{out}), each output position (i , j) where $0 \leq i < H_{out}$, $0 \leq j < W_{out}$, the start and end indices in the input are calculated as:

$$kernel_{height} = \left\lceil \frac{H_{in}}{H_{out}} \right\rceil$$

$$kernel_{width} = \left\lceil \frac{W_{in}}{W_{out}} \right\rceil$$

$$stride_{height} = \left\lfloor \frac{H_{in}}{H_{out}} \right\rfloor$$

$$stride_{width} = \left\lfloor \frac{W_{in}}{W_{out}} \right\rfloor$$

A Modified Experimental RCF Model

To obtain more detail and checking for a better predicting power, a new RCF model is proposed by incorporating one convolutional layer and one dropout layer as a modified experimental model on the previous extraordinary model.

As shown in Figure 4, the RCF2 model enhances the initial set of the convolutions by adding a second convolutional layer known as conv2. Another useful component is the convolutional layers which are used to extend the model's ability to capture more complex and abstract features through series of convolutions.

The second convolutional layer increases the number of feature channels, thus giving the model the ability to detect more features. The output of conv1 which has $num_features/4$ channels is used as input to conv2 and the output of conv2 is of $num_features$ channels. This provides a new form of the RCF model if any change in predicting power is needed.

Regularization with Dropout

The second layer of the network is the ReLU, which is basically the positive ReLU activation followed by the negative ReLU activation to extract both the positive and negative feature map responses, in a manner analogous to the RCF model. But now the result is firstly passed through a dropout layer and then through the second convolutional layer. Dropout is a form of regularization which is an attempt to reduce overfitting by purposefully 'dropping out' a certain percentage of the weights or activation values during training.

This is because dropout helps the model not to be dependent on certain inputs and in the process of learning different features. This is most helpful especially when working with a small data set or when the model has many parameters.

By combining positive and negative feature responses from two layers, the model can produce a comprehensive and nuanced feature representation, which can be beneficial for tasks such as classification, detection, or segmentation.

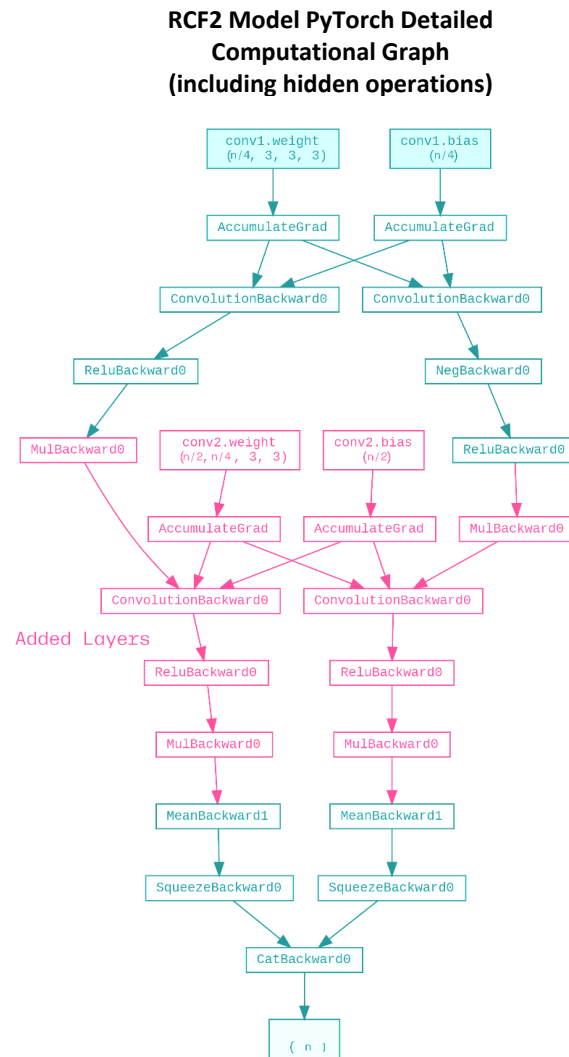


Figure 4: Computational Graph of RCF2 Model

Results

Model Interpretability Analysis Results

The image features used in the model are effective in predicting the Green Space Ratio, Blue Space Ratio, and Building Area Ratio, with high accuracy as indicated by the high R^2 values and low MSE values, as shown in Table 1. The slightly lower R^2 for Blue Space Ratio may suggest that these features are more variable or less consistently represented in the image data compared to Green Space and Building Area Ratios. The results suggest that the image features with a size of 100m capture key aspects of these spatial ratios effectively, with the strongest predictive power for Building Area Ratio, followed by Green Space Ratio, and then Blue Space Ratio, as shown in Figure 5.

Outcome variable	MSE	R2
Green Space Ratio	0.0052	0.91
Blue Space Ratio	0.0005	0.89
Building Area Ratio	0.0005	0.93
Temperature	1.5531	0.92

Table 1: Model Interpretability Analysis for Image Patch Size 100m, 100 Features

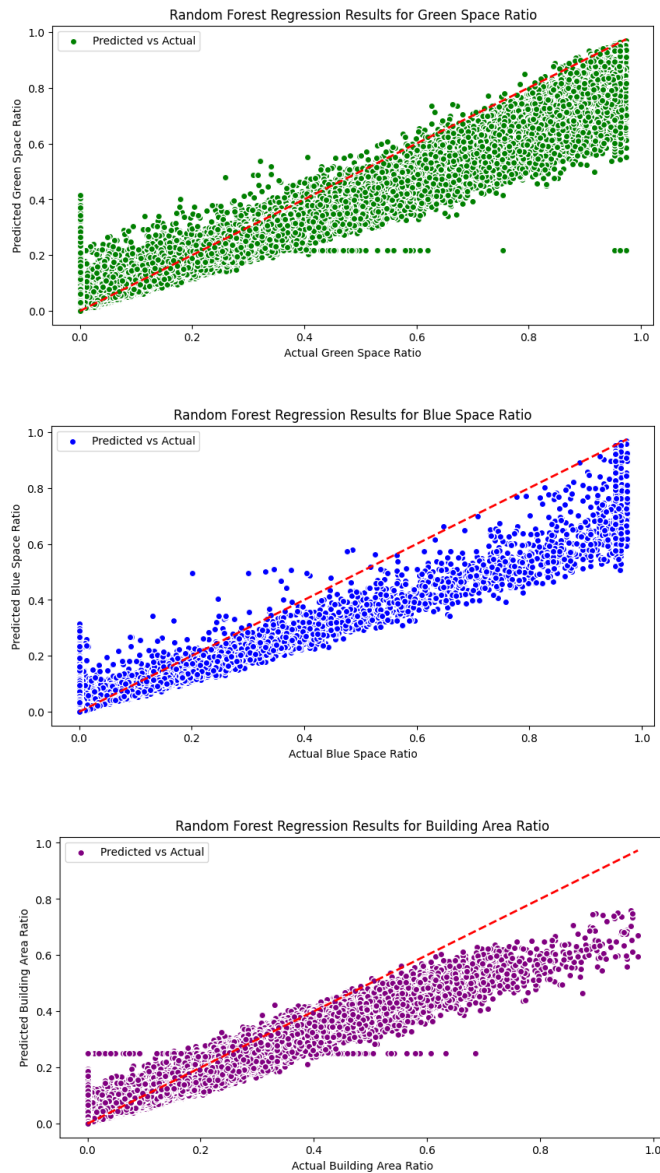


Figure 5: Model Interpretability Analysis for Image Patch Size 100m, 100 Features

Temperature Prediction Results

Temperature Prediction from Image Features Extracted from RCF with Random Forest

The Random Forest regression model performs well in predicting average temperature from image features, as indicated by the low MSE and high R^2 score. The model demonstrates strong predictive power, with most of the variance in temperature being explained by the image features. This indicates predicting temperature from the features extracted is an eligible method. As shown in Figure 6, the scatter plot further corroborates the model's effectiveness, showing that the predictions are closely aligned with the actual values. However, there is room for improvement, particularly in extreme cases, where prediction errors are slightly higher.

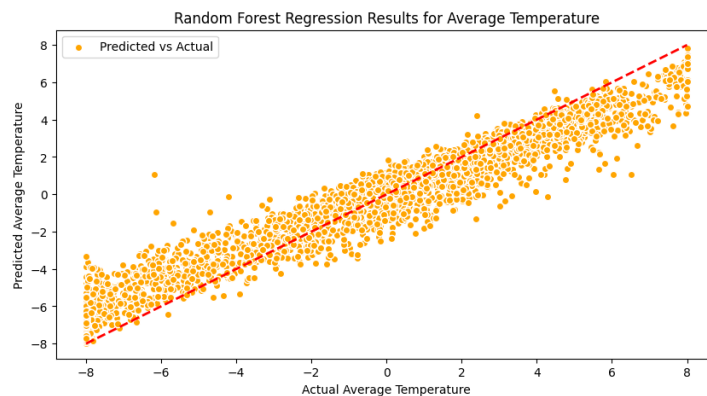


Figure 6: Temperature Prediction for Image Patch Size 100m, 100 Features

Scale Comparison

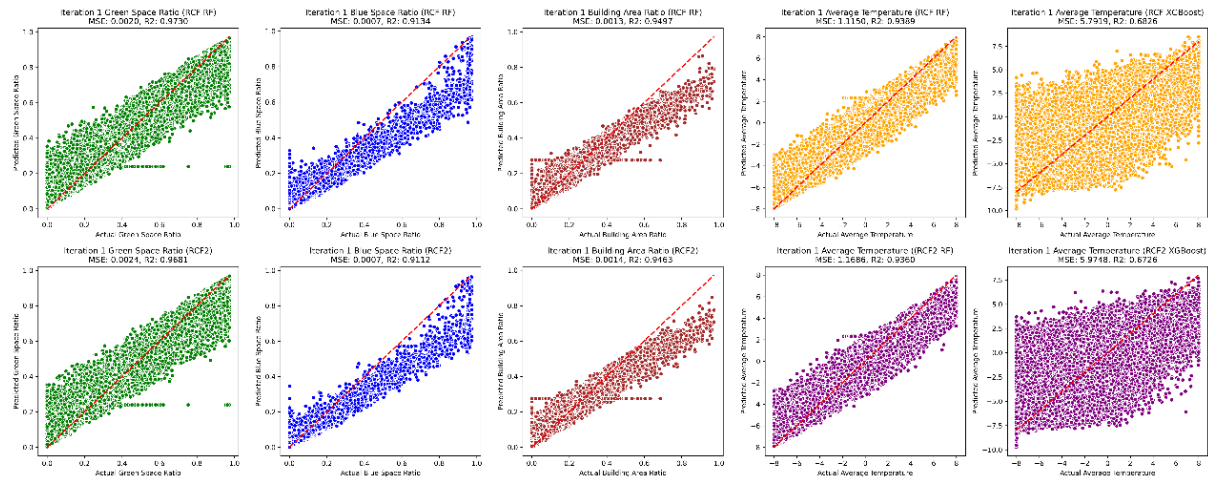


Figure 7: Analysis Comparison for Image Patch Size 100m, 100 Features

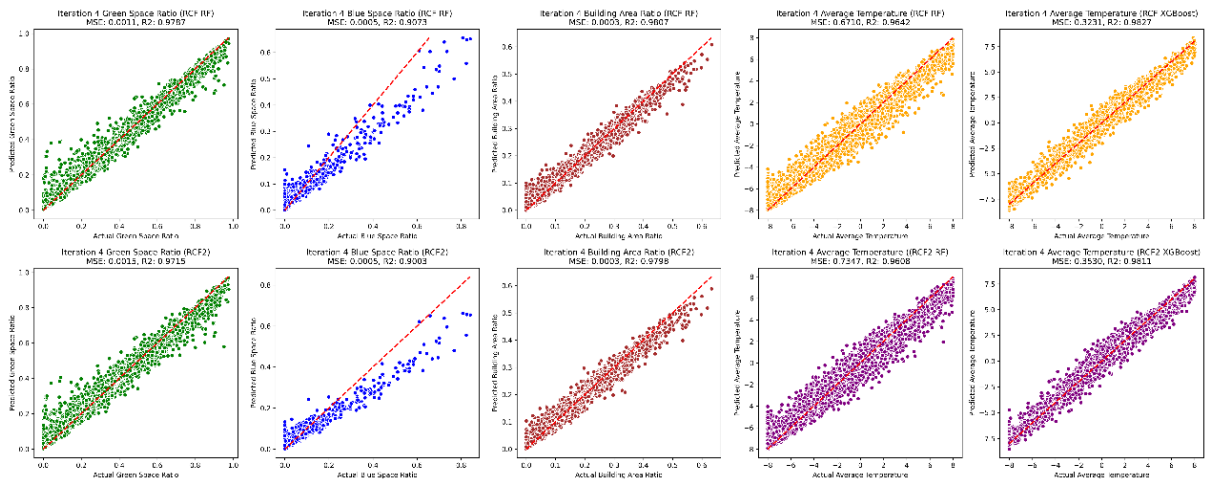


Figure 8: Analysis Comparison for Image Patch Size 400m, 100 Features

Green and Blue Space Ratio Predictions

For both green and blue space ratios, Random Forest performs consistently well across all image patch sizes as shown in the first 2 columns of plots in Figure 7 and 8. The predictions align closely with the actual values, as indicated by high R^2 scores and low MSE values.

For green space ratio, in the scenario when grid size is 100m, the evaluation metrics for Random Forest with RCF are MSE 0.0020, R^2 0.9730. The ones for Random Forest with RCF2 are MSE 0.0024, R^2 0.9681.

In the scenario when grid size is 400m, the evaluation metrics for Random Forest with RCF are MSE 0.0011, R^2 0.9787. The ones for Random Forest with RCF2 are MSE 0.0015, R^2 0.9715.

For blue space ratio, in the scenario when grid size is 100m, the evaluation metrics for Random Forest with RCF are MSE 0.0007, R^2 0.9134. The ones for Random Forest with RCF2 are MSE 0.0007, R^2 0.9112.

In the scenario when grid size is 400m, the evaluation metrics for Random Forest with RCF are MSE 0.0005, R^2 0.9073. The ones for Random Forest with RCF2 are MSE 0.0005, R^2 0.9003.

RCF2 introduces additional complexity with a deeper architecture, but this does not consistently outperform RCF for predicting green and blue space ratios using Random Forest. RCF slightly outperforms RCF2 in some cases, showing that the simpler model may be more effective for this task.

Building Area Ratio Predictions

Both RCF and RCF2 models perform well in predicting the building area ratio as shown in the third column of plots in Figure 7 and 8. R^2 scores are consistently high across all patch sizes, indicating strong predictive power. The difference between RCF and RCF2 is minimal, with RCF sometimes showing a slight edge.

For building area ratio, in the scenario when grid size is 100m, the evaluation metrics for Random Forest with RCF are MSE 0.0013, R2 0.9497. The ones for Random Forest with RCF2 are MSE 0.0013, R2 0.9464.

In the scenario when grid size is 400m, the evaluation metrics for Random Forest with RCF are MSE 0.0003, R2 0.9807. The ones for Random Forest with RCF2 are MSE 0.0003, R2 0.9798.

The experimental architecture of RCF2 does not significantly outperform RCF in this context, similar to the green and blue space ratio predictions.

Average Temperature Predictions

As shown in the last 2 columns of plots in Figure 7 and 8, comparing the rows, both RCF and RCF2 perform reasonably well, but there is a marginal improvement with RCF, which captures temperature variations slightly better.

For temperature, in the scenario when grid size is 100m, the evaluation metrics for Random Forest with RCF are MSE 1.1160, R2 0.9388. The ones for Random Forest with RCF2 are MSE 5.7919, R2 0.6826.

In the scenario when grid size is 400m, the evaluation metrics for Random Forest with RCF are MSE 0.6710, R2 0.9642. The ones for Random Forest with RCF2 are 0.7347, R2 0.9608.

In the scenario when grid size is 400m, the evaluation metrics for XGBoost with RCF are MSE 0.3231, R2 0.9827. The ones for XGBoost with RCF2 are MSE 0.3530, R2 0.9811.

As shown in the last 2 columns of plots in Figure 7 and 8, comparing the columns, XGBoost significantly outperforms Random Forest in predicting average temperature, particularly when using features extracted by RCF2. The performance gain is evident across all patch sizes, with XGBoost achieving higher R² scores and lower MSE values.

RCF2 with XGBoost is shown to be particularly effective, indicating that the altered version of RCF model better captures the detailed features which can facilitate accurate temperature prediction maintaining the function of the original model on feature extraction.

Effect of Image Patch Size

As shown in Figure 9 to 12, in smaller patches (<300m), Random Forest demonstrates a better performance than XGBoost.

For instance, when patches are 100m, the evaluation metrics for Random Forest RCF reach MSE 1.1150, R2 0.9389, while the evaluation metrics for XGBoost RCF stay 5.7919, R2 0.6826.

For the same patch size, the evaluation metrics for Random Forest RCF2 achieve 1.1686, R2 0.9360, while XGBoost RCF2 falls short of it with MSE 5.9748, R2 0.6726.

These results suggests that the original RCF often achieves slightly better results when the image size patch size is 100m, indicating that detailed features extracted by RCF2 may not be necessary for smaller patches.

In larger patches (300m), XGBoost prone to be more effective, outperforming Random Forest as the patch size increases. The combination of RCF2 with XGBoost benefits from the detailed feature extraction of RCF2, being more relevant for larger, more complex images.

For example, when patch size is 400m, the evaluation metrics for XGBoost RCF reaches MSE 0.3231 R2 0.9827, while the evaluation metrics for Random Forest RCF are MSE 0.6710, R2 0.9642.

When patches are 400m, the evaluation metrics for XGBoost RCF2 reaches MSE 0.3530, R2 0.9811, while the evaluation metrics for Random Forest RCF2 stays MSE 0.7347, R2 0.9608.

The most effective temperature prediction method is to use the new form RCF2 for feature extraction combined with XGBoost for regression. This approach takes advantage of the ability of RCF2 to capture detailed features and the superior predictive ability of XGBoost, especially for larger image patches.

When image patches are smaller than 300m, random forest generally performs better than XGBoost, probably because simple models are sufficient for less complex data. For image patches larger than 300m, XGBoost performs better than random forest thanks to the richer feature set extracted by RCF2 and the enhanced complexity handling by XGBoost model.

Image Feature Temperature Prediction Combining Green and Blue Spatial Ratio and Building Ratio

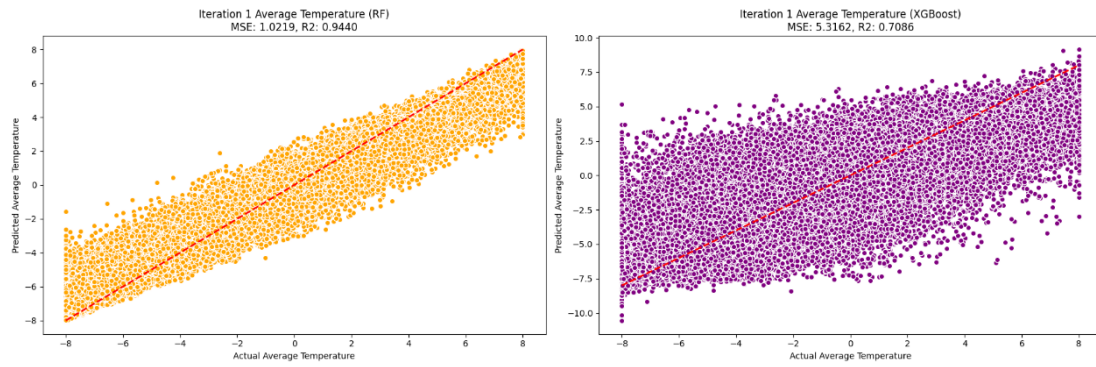


Figure 9: Temperature Prediction Based on Image Patch Size 100m, 100 Features

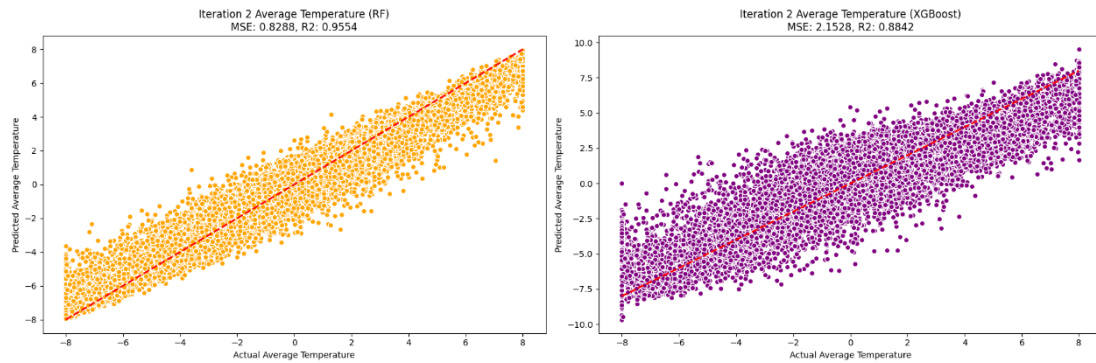


Figure 10: Temperature Prediction Based on Image Patch Size 200m, 100 Features

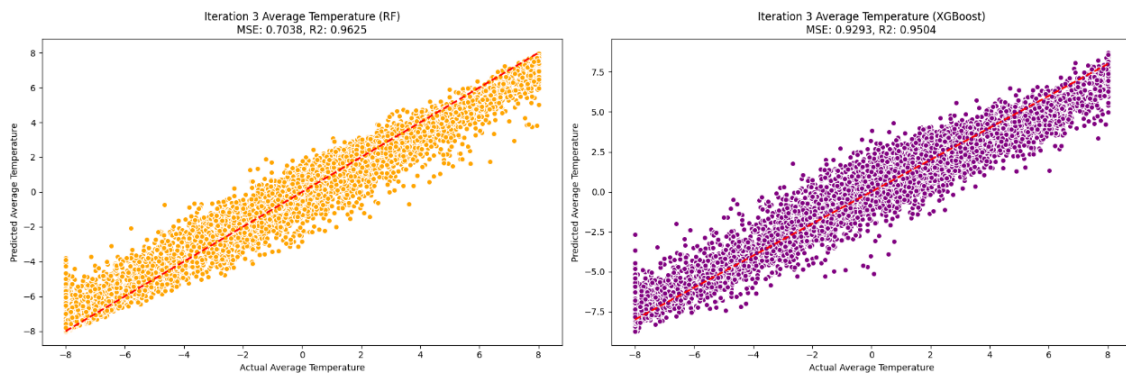


Figure 11: Temperature Prediction Based on Image Patch Size 300m, 100 Features

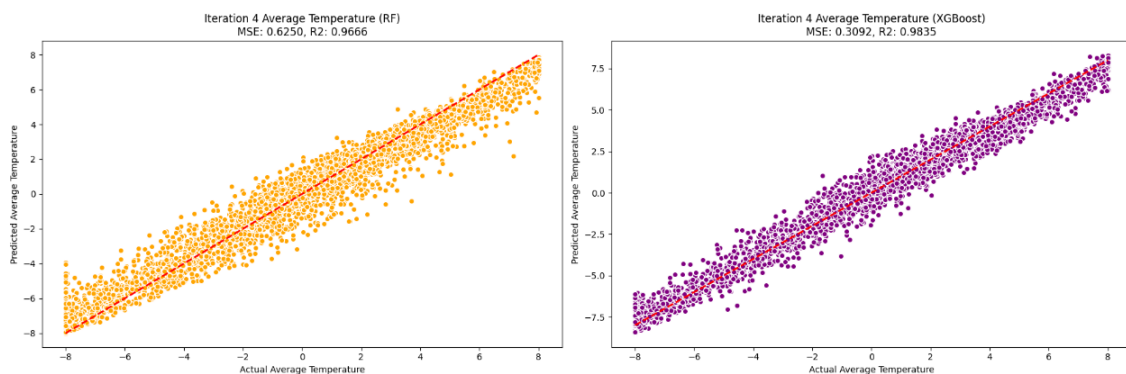


Figure 12: Temperature Prediction Based on Image Patch Size 400m, 100 Features

In terms of MSE and R-squared for all metrics: green space ratio, blue space ratio, building area ratio, and average temperature, the prediction with combined features outperforms the prediction with only 100 features. Visually, the combined features produce more aligned data points that are closer to the ideal prediction line, especially for the prediction of mean temperature. This improvement is particularly evident in the XGBoost model, where combining features significantly improves accuracy, reduces dispersion, and aligns predictions more closely with actual values.

This suggests that incorporating additional features into the model can improve predictive performance across different environmental and structural metrics.

Discussion

The feature extraction of the image covers the basic two-dimensional features of the urban orthophoto, and reliable prediction results can be obtained without a separate manual analysis of the features. However, there is a lack of attention to 3D features, which have an important impact on street valleys and, in turn, microclimate. Nonetheless, it also avoids redundant computation of 3D features, which usually takes a long time.

Conclusions

This study explores combining machine learning techniques with satellite imagery to strengthen the understanding of urban temperature changes, particularly in New York City. By using the MOSAIKS framework for feature extraction and combining it with advanced models such as random forests and XGBoost, we are able to predict temperatures with high accuracy in urban environments of different scales.

Firstly, the performance of MOSAIKS model was enhanced by modifying the process of temperature prediction from image patches. The new version of RCF model called RCF2, which includes an additional convolutional layer and dropout, demonstrates the superior ability to extract detailed features from satellite images, especially when combined with XGBoost for regression tasks. This finding aligns with what Andrew Ng introduced in the professional course of deep learning, increasing the depth of neural networks enables them to learn more complex hierarchical representations of data [33]. This view is supported by theoretical research, which shows that deep networks can represent certain functions with exponential efficiency compared to shallow networks [34-36]. This combination consistently performs better than the other methods, especially in temperature prediction for larger image patches ($\geq 300\text{m}$). While Random Forest is effective using smaller image patches ($< 300\text{m}$) and effectively captures generic features, the complexity and power of XGBoost appear to be more suitable for larger and more detailed image patches.

Second, the results demonstrates the importance of image block size. The validity of the prediction model is scale dependent. Smaller image patches benefit the random forest model, while larger image patches benefit from the novel RCF2 model for its feature extraction capability and XGBoost.

Again, the study highlights the impact of urban form on temperature, from which an improved workflow is found. The results show that urban green space, blue space and building footprint are the key factors affecting urban temperature change. Regression analysis of these features with image-derived features shows strong predictive power, underscoring the importance of integrating environmental and morphological indicators into urban climate studies. Combining key factors with image features improves the temperature prediction ability.

Finally, the study shows that the method is scalable and general. Feature extraction using RCF2 provides a scalable solution for temperature prediction in different urban environments. This approach is applicable not only to New York City but also to other urban areas, suggesting that similar approaches can be adapted to different geographical contexts.

In conclusion, this study introduces a robust and scalable method for predicting urban temperatures using satellite imagery and machine learning, providing valuable insights for urban planners and decision makers in the aim to mitigate the impact of urban heat islands.

Data Availability

The datasets and code supporting the findings of this study are openly available in the Zenodo repository at <https://doi.org/10.5281/zenodo.17281718>, under a CC-BY 4.0 license. The repository includes: the shapefile of New York City used for fetching the image patches, the surface temperature of the site relative to the average temperature of the city, The code including satellite image fetching, the Random Convolution Features (RCF), and RCF2 model for temperature prediction.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This work was supported by the Harvard Center for Geographic Analysis.

References

1. Rolf, E., Proctor, J., Carleton, T., Bolliger, I., Shankar, V., Ishihara, M., ... & Hsiang, S. (2021). A generalizable and accessible approach to machine learning with global satellite imagery. *Nature communications*, 12(1), 4392.
2. Qin, Z., Karnieli, A., & Berliner, P. (2001). A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *International journal of remote sensing*, 22(18), 3719-3746.
3. Jin, M., Li, J., Wang, C., & Shang, R. (2015). A practical split-window algorithm for retrieving land surface temperature from Landsat-8 data and a case study of an urban area in China. *Remote sensing*, 7(4), 4371-4390.
4. Adeniran, I. A., Nazeer, M., Wong, M. S., & Chan, P. W. (2024). An improved machine learning-based model for

- prediction of diurnal and spatially continuous near surface air temperature. *Scientific reports*, 14(1), 27342.
5. Tan, J., NourEldeen, N., Mao, K., Shi, J., Li, Z., Xu, T., & Yuan, Z. (2019). Deep learning convolutional neural network for the retrieval of land surface temperature from AMSR2 data in China. *Sensors*, 19(13), 2987.
 6. Yin, Z., Wu, P., Foody, G. M., Wu, Y., Liu, Z., Du, Y., & Ling, F. (2020). Spatiotemporal fusion of land surface temperature based on a convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2), 1808-1822.
 7. dos Santos, R. S. (2020). Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 88, 102066.
 8. Guo, Y., Unger, J., & Gál, T. (2024). Model Development for Estimating Sub-Daily Urban Air Temperature Patterns in China Using Land Surface Temperature and Auxiliary Data from 2013 to 2023. *Remote Sensing*, 16(24), 4675.
 9. United Nations Department of Economic and Social Affairs, Population Division. 2016. The World's Cities in 2016: Data Booklet. New York, NY: United Nations.
 10. Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly journal of the royal meteorological society*, 108(455), 1-24.
 11. Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote sensing of environment*, 86(3), 370-384.
 12. Asimakopoulous, D. N. (2001). *Energy and climate in the urban built environment*. Earthscan.
 13. Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. *Remote sensing of Environment*, 89(4), 467-483.
 14. Yu, S., Chen, Z., Yu, B., Wang, L., Wu, B., Wu, J., & Zhao, F. (2020). Exploring the relationship between 2D/3D landscape pattern and land surface temperature based on explainable eXtreme Gradient Boosting tree: A case study of Shanghai, China. *Science of the total environment*, 725, 138229.
 15. Yu, X., Guo, X., & Wu, Z. (2014). Land surface temperature retrieval from Landsat 8 TIRS—Comparison between radiative transfer equation-based method, split window algorithm and single channel method. *Remote sensing*, 6(10), 9829-9852.
 16. Rosenzweig, C., Solecki, W. D., Parshall, L., Lynn, B., Cox, J., Goldberg, R., ... & Watson, M. (2009). Mitigating New York City's heat island: Integrating stakeholder perspectives and scientific evaluation. *Bulletin of the American Meteorological Society*, 90(9), 1297-1312.
 17. Loose, T., & Bornstein, R. D. (1977). Observations of mesoscale effects on frontal movement through an urban area. *Monthly Weather Review*, 105(5), 563-571.
 18. Gu, X., Wu, Z., Liu, X., Qiao, R., & Jiang, Q. (2024). Exploring the nonlinear interplay between urban morphology and nighttime thermal environment. *Sustainable Cities and Society*, 101, 105176.
 19. Alexander, C. (2021). Influence of the proportion, height and proximity of vegetation and buildings on urban land surface temperature. *International Journal of Applied Earth Observation and Geoinformation*, 95, 102265.
 20. Wang, C., Liu, Z., Du, H., & Zhan, W. (2023). Regulation of urban morphology on thermal environment across global cities. *Sustainable cities and society*, 97, 104749.
 21. Verma, R., Zawadzka, J. E., Garg, P. K., & Corstanje, R. (2024). The relationship between spatial configuration of urban parks and neighbourhood cooling in a humid subtropical city. *Landscape Ecology*, 39(2), 34.
 22. Xue, Z., Hou, G., Zhang, Z., Lyu, X., Jiang, M., Zou, Y., ... & Liu, X. (2019). Quantifying the cooling-effects of urban and peri-urban wetlands using remote sensing data: Case study of cities of Northeast China. *Landscape and Urban Planning*, 182, 92-100.
 23. Yuan, B., Zhou, L., Dang, X., Sun, D., Hu, F., & Mu, H. (2021). Separate and combined effects of 3D building features and urban green space on land surface temperature. *Journal of Environmental Management*, 295, 113116.
 24. Buster, G., Benton, B. N., Glaws, A., & King, R. N. (2024). High-resolution meteorology with climate change impacts from global climate model data using generative machine learning. *Nature Energy*, 9(7), 894-906.
 25. Qi, Y., Chodron Drolma, S., Zhang, X., Liang, J., Jiang, H., Xu, J., & Ni, T. (2020). An investigation of the visual features of urban street vitality using a convolutional neural network. *Geo-spatial Information Science*, 23(4), 341-351.
 26. Petkova, E. P., Horton, R. M., Bader, D. A., & Kinney, P. L. (2013). Projected heat-related mortality in the US urban northeast. *International journal of environmental research and public health*, 10(12), 6734-6747.
 27. Zeng, P., Sun, F., Liu, Y., Tian, T., Wu, J., Dong, Q., ... & Che, Y. (2022). The influence of the landscape pattern on the urban land surface temperature varies with the ratio of land components: Insights from 2D/3D building/vegetation metrics. *Sustainable Cities and Society*, 78, 103599.
 28. Hu, Y., Dai, Z., & Guldmann, J. M. (2020). Modeling the impact of 2D/3D urban indicators on the urban heat island over different seasons: A boosted regression tree approach. *Journal of environmental management*, 266, 110424.
 29. Mo, Y., Bao, Y., Wang, Z., Wei, W., & Chen, X. (2024). Spatial coupling relationship between architectural landscape characteristics and urban heat island in different urban functional zones. *Building and Environment*, 257, 111545.
 30. Wei, L., & Sobrino, J. A. (2024). Surface urban heat island analysis based on local climate zones using ECOSTRESS and Landsat data: A case study of Valencia city (Spain). *International Journal of Applied Earth Observation and Geoinformation*, 130, 103875.

31. Rolf, E., Proctor, J., Carleton, T., Bolliger, I., Shankar, V., Ishihara, M., ... & Hsiang, S. (2021). A generalizable and accessible approach to machine learning with global satellite imagery. *Nature communications*, 12(1), 4392.
32. Kenyon, G. E., Arribas-Bel, D., & Robinson, C. (2024). Extracting Features from Satellite Imagery to Understand the Size and Scale of Housing Sub-Markets in Madrid. *Land*, 13(5), 575.
33. A. Ng, Deep Learning Specialization, Coursera / DeepLearning.AI, online course.
34. Eldan, R., & Shamir, O. (2016, June). The power of depth for feedforward neural networks. In *Conference on learning theory* (pp. 907-940). PMLR.
35. Telgarsky, M. (2016, June). Benefits of depth in neural networks. In *Conference on learning theory* (pp. 1517-1539). PMLR.
36. Raghu, M., Poole, B., Kleinberg, J., Ganguli, S., & Sohl-Dickstein, J. (2017, July). On the expressive power of deep neural networks. In *international conference on machine learning* (pp. 2847-2854). PMLR.
37. Busquets, J., D. Yang, and M. Keller. 2019. *Urban Grids: Handbook for Regular City Design*. Novato, CA: Oro Editions; Cambridge, MA.
38. Gaffin, S. R., Rosenzweig, C., Khanbilvardi, R., Parshall, L., Mahani, S., Glickman, H., ... & Hillel, D. (2008). Variations in New York city's urban heat island strength over time and space. *Theoretical and applied climatology*, 94(1), 1-11.
39. Chaurasia, K., Neeraj, B., Burle, D., & Mishra, V. K. (2020, September). Topographical feature extraction using machine learning techniques from sentinel-2A imagery. In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1659-1662). IEEE.
40. Li, J., Song, C., Cao, L., Zhu, F., Meng, X., & Wu, J. (2011). Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. *Remote sensing of environment*, 115(12), 3249-3263.
41. Zhao, J., Guo, F., Zhang, H., & Dong, J. (2024). Mechanisms of non-stationary influence of urban form on the diurnal thermal environment based on machine learning and MGWR analysis. *Sustainable Cities and Society*, 101, 105194.

During the preparation of this work the author(s) used ChatGPT in order to Correct Grammar and Expression. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.