The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIX-B3-2026
XXV ISPRS Congress 2026 "From Imagery to Understanding", Commission III, 4–11 July 2026, Toronto, Canada

Multi-Sensor Random Forest Downscaling for 10 m LST Mapping and Urban Heat Island Monitoring in a Small-Sized City

Nicola Colaninno

Department of Architecture and Urban Studies (DAStU), Politecnico di Milano, via Bonardi 3, 20133 Milano, Italy - nicola.colaninno@polimi.it

Keywords: SUHI, Random Forest Regression, Data Downscaling, Multi-Sensor Data Fusion, Urban Climate, Adaptation

Abstract

Urban heat islands (UHIs) present a critical challenge to sustainable urban development, demanding high-resolution monitoring tools for effective climate adaptation. We address this need by implementing a machine learning framework for downscaling Land Surface Temperature (LST) data, demonstrating its ability to capture fine-scale thermal variations. The methodology leverages multi-sensor remote sensing data fusion, integrating high-resolution optical observations from Sentinel-2 with thermal imagery from Landsat 9 (daytime LST reference) and ASTER (nighttime LST reference). Random Forest (RF) regression is employed, utilizing Sentinel-2 multispectral bands, derived spectral indices (e.g., NDVI, NDBI) to characterize land cover, and a Digital Elevation Model (DEM) to account for topographic effects. The RF model was rigorously trained, and its hyperparameters optimized via randomized cross-validation to predict LST at a 10-meter resolution. Results demonstrate robust performance, achieving a high R^2 of 0.75 (Mean Absolute Error, MAE: 1.7°C) for daytime LST and R^2 of 0.50 (MAE: 0.6°C) for nighttime LST. The resulting downscaled maps delineate pronounced heat accumulation in dense built-up areas, notably its historic center and large commercial zones, contrasting sharply with cooler vegetated areas and green urban corridors. A comparative assessment against bilinear interpolation, TsHARP thermal sharpening, and linear regression confirms that the RF framework achieves the best balance between predictive accuracy, spatial coherence with the source thermal data, and meaningful sub-pixel detail, effectively preserving the critical fine-scale thermal patterns. Ultimately, this study advances UHI monitoring by enabling the precise identification of heat-vulnerable areas, thereby supporting targeted mitigation strategies even in small and medium-sized cities.

1. Introduction

Urban heat islands (UHIs) are a critical consequence of urbanization, characterized by elevated land surface and air temperatures in densely built environments relative to their rural surroundings. Effective mitigation requires monitoring at spatial resolutions that can capture the fine heterogeneity of thermal landscapes, differentiating, for example, the different thermal behaviors of streets, building blocks, and vegetated corridors. However, most satellite thermal sensors operate at coarse spatial resolutions (e.g., MODIS at 1 km, Landsat TIRS at 100 m, ASTER at 90 m), which are insufficient for neighborhood-scale planning (Agam et al., 2007; Bechtel et al., 2012).
The integration of optical and thermal remote sensing data has shown promise in UHI studies, yet limitations persist not only in terms of spatial but also temporal resolution (Uhrin & Onačillová, 2025). This limitation has motivated the development of downscaling methods to derive fine-resolution Land Surface Temperature (LST) from coarser thermal inputs, including day and night observations.
Early approaches were dominated by index-based sharpening, particularly those leveraging the relationship between vegetation cover and surface temperature (Agam et al., 2007; Essa et al., 2012; Kustas et al., 2003). While computationally efficient, these methods typically performed poorly in heterogeneous urban settings where impervious surfaces dominate thermal behavior (Andriambololonaharisoamalala et al., 2025; Xu et al., 2020). The availability of high-resolution optical data, such as Sentinel-2, has opened new frontiers for thermal data downscaling. At the same time, Machine Learning (ML) algorithms emerged as a popular choice do to the ability to handle nonlinear relationships. ML algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), require fewer prior statistical assumptions and are adept to model multivariate, complex relationships (Ebrahimy & Azadbakht, 2019; Roy et al., 2025; Yang et al., 2017).
Different studies have demonstrated that RF models, when trained with multispectral bands, spectral indices, and auxiliary layers such as Digital Elevation Models (DEM), can downscale LST to very-high spatial resolution with strong accuracy (Xu et al., 2021, 2020). A hybrid "geospatial machine learning" (GeoML) model, combining random forest and kriging downscaling with high-resolution Sentinel-2 data, has been shown to significantly improve LST estimates from Landsat 8/9, with a strong match with ground measurements (Andriambololonaharisoamalala et al., 2025).
However, despite these advances, important gaps remain. High computational demand is pointed out as a limiting factor for practicality when handling large-scale or time-sensitive applications. Moreover, nighttime LST downscaling remains underexplored (Arellano & Roca, 2021) and typically achieves weaker accuracy. Local Linear Forest and Random Forest have been applied to downscale nighttime MODIS LST to a 250 m resolution, producing models with strong goodness-of-fit and low error (Yoo et al., 2022). While these results are promising, the achieved resolution remains too coarse for detailed neighborhood-level UHI analysis.
Furthermore, comparative day and night LST analysis over the same location, as well as extending such methods to small and medium-sized cities, remains an important gap for advancing UHI monitoring (Uhrin & Onačillová, 2025).
We present an integrated, low-computational framework to generate high-resolution daytime and nighttime LST maps by fusing multi-sensor remote sensing data with RF regression. We apply the framework to examine surface urban heat islands (SUHIs) in Salò, a city in Brescia Province, Lombardy, northern Italy, situated on the western shore of Lake Garda.

The primary objective is to combine Sentinel-2 optical data, Landsat 9 daytime thermal imagery, and ASTER nighttime thermal imagery to generate 10 m resolution LST maps for both diurnal conditions over the same study area.

The specific contributions of this work are: (i) the systematic integration of three satellite platforms (Sentinel-2, Landsat 9, and ASTER) within a unified downscaling pipeline, enabling both daytime and nighttime LST mapping at 10 m resolution; (ii) the application and validation of the framework in a small-sized city, a context underrepresented in the literature, typically focused on large metropolitan areas; and (iii) the development of a computationally lightweight and transferable methodology, designed to be accessible for operational use by local authorities with limited technical resources.

The high-resolution outputs reveal micro-scale heat patterns, empowering the City of Salò to implement targeted climate adaptation strategies such as urban greening and depavement. Furthermore, this framework can be adapted for other small and medium-sized cities facing similar climate challenges, advancing wider efforts toward sustainable urban development.

The remainder of this paper is organized as follows: Section 2 details the methodology, including data preprocessing, feature extraction, and model training. Section 3 presents the results of LST downscaling and hotspot analysis. Section 4 discusses the implications for urban planning. Section 5 concludes with recommendations for future research and policy applications.

2. Methodology

The methodology integrates multi-sensor remote sensing data with machine learning techniques to enhance the spatial resolution of land surface temperature (LST) observations. The framework consists of four main components: data acquisition and preprocessing, feature extraction and selection, model development and optimization, and spatial prediction. The methodology builds established techniques in spectral index derivation (e.g., NDVI, NDBI) and machine learning optimization (Uhrin & Onačillová, 2025), while incorporating topographic corrections through a DEM.

2.1 Data Acquisition and Preprocessing

The study utilizes optical and thermal imagery from Sentinel-2, Landsat 9, and ASTER, harmonized to a common spatial resolution of 10 m. Sentinel-2 Level-2A data provide multispectral reflectance bands (B02, B03, B04, B08 at 10 m; B11, B12 at 20 m resampled to 10 m using bilinear interpolation), while Landsat 9 Thermal Infrared Sensor (TIRS) and ASTER thermal bands supply daytime and nighttime LST references at 30 m and 90 m, respectively, with Landsat 9 TIRS data resampled from its native 100 m resolution to 30 m.

Two particularly hot days were selected: July 31st, 2024, and August 23rd, 2023. The choice of these dates was influenced by the availability of sufficiently clear satellite images (some images were excluded due to severe cloud coverage), and by the need to ensure temporal compatibility between the various datasets. Specifically, for each day, Sentinel-2 images (July 31st, 2024, and August 23rd, 2023) were paired with Landsat 9 images (July 31st, 2024) and ASTER images (August 23rd, 2023), ensuring temporal consistency. The July 31st image supports daytime LST analysis through Landsat 9 TIRS, whereas the August 23rd image enables nighttime LST analysis through ASTER. The data acquisition times and air temperatures (T_a) were recorded at the Brescia Montichiari meteorological station and are summarized in Table 1. The dual-date configuration captures both daytime and nighttime thermal conditions under comparably hot summer regimes.

Parameter	July 31st, 2024	August 23rd, 2023
Time	10:00 AM	8:49 PM
Data Combination	Sentinel-2 + Landsat 9	Sentinel-2 + ASTER
Average T_a	30°C	31°C
Minimum T_a	25°C	25°C
Maximum T_a	34°C	37°C

Table 1. Data acquisition times and air temperature statistics recorded at the Brescia Montichiari meteorological station.

All datasets were radiometrically calibrated, geometrically corrected, and projected into a common coordinate system. Cloud masking was applied using quality assessment bands. Temporal alignment ensured consistency across acquisition dates.

A 5 m resolution DEM, resampled to 10 m, was incorporated to account for topographic effects on temperature patterns. The DEM was particularly valuable for nighttime modeling when elevation-driven cooling effects become pronounced.

Spectral indices were derived from Sentinel-2 reflectance data to characterize land surface properties influencing thermal behavior. The Normalized Difference Vegetation Index (NDVI) and Modified Soil-Adjusted Vegetation Index (MSAVI) quantified vegetation density, while the Normalized Difference Built-up Index (NDBI) and Built-up Surface Index (BSI) identified urbanized areas. Additional indices included the Perpendicular Vegetation Index (PVI) for vegetation structure, Leaf Area Index (LAI) for canopy density, the Normalized Difference Water Index (NDWI) to account for water, and broadband albedo (Liang, 2001) for surface reflectivity.

2.2 Dimensionality Reduction and Regression Modeling

2.2.1 Predictors Selection and Principal Component Analysis: To ensure that the predictors included in the RF model carried meaningful explanatory power for LST, a multi-step variable selection and dimensionality reduction procedure was implemented. Sentinel-2 bands, derived spectral indices, and DEM represent the set of explanatory variables (predictors), while L9- and ASTER-derived LST are the dependent variables. Spearman's rank correlation analysis was conducted between each candidate predictor and LST. This non-parametric measure was chosen because it captures monotonic relationships without assuming linearity, which is particularly appropriate for ecohydrological and land surface processes that may exhibit non-linear dynamics.

Predictors with negligible or weak correlation with LST were excluded, thereby retaining only indicators with a demonstrable statistical association with the response variable. The selected predictors were then standardized (zero mean, unit variance) to address multicollinearity, a common issue among vegetation indices and biophysical parameters that are often derived from overlapping spectral information.

Principal Component Analysis (PCA) was then applied to the standardized selected predictors. The transformation produced a set of orthogonal PCs, each representing an uncorrelated linear combination of the original predictors. By construction, PCs capture decreasing proportions of total variance in the predictor set, ordered from most to least explanatory.

Figure 1 shows the correlation matrix heatmap between dependent and independent variables, while Table 2 summarizes the PCA results for both L9 and ASTER LST models, listing the standardized Sentinel-2-based indicators retained for analysis, their loadings on the selected principal components, and the proportion of variance explained by each component relative to the original predictor set.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-355-2026

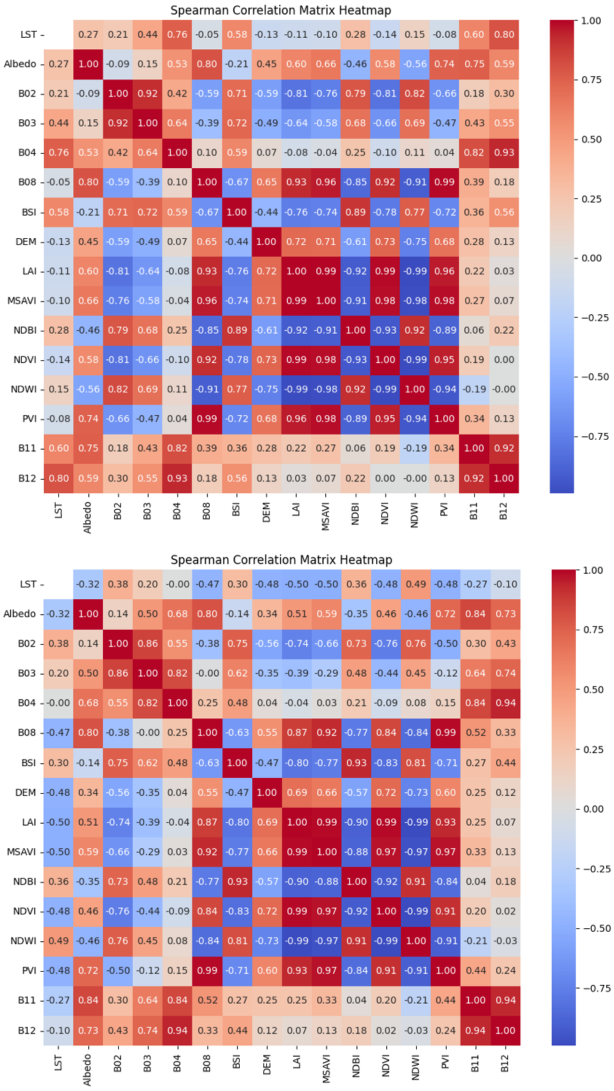

Figure 1. Correlation between daytime Landsat 9 LST (top image) and nighttime ASTER LST (bottom image) with Sentinel-2 spectral bands, derived vegetation and built-up indices, and elevation data (DEM).

Dataset	Indicator	PC1	PC2	PC3	PC4	PC5	PC6	PC7
S2-L9	Variance	**0.71**	**0.20**	**0.08**	0.00	0.00	–	–
	B03	0.80	-0.46	-0.36	-0.10	-0.00	–	–
	B04	0.96	-0.16	-0.14	0.15	-0.02	–	–
	BSI	0.77	-0.38	0.49	-0.03	-0.03	–	–
	B11	0.71	0.70	-0.02	-0.04	-0.08	–	–
	B12	0.92	0.35	0.08	-0.01	0.11	–	–
S2-ASTER	Variance	**0.89**	**0.08**	**0.02**	0.00	0.00	0.00	0.00
	B08	0.96	-0.15	0.24	-0.01	-0.02	0.01	-0.00
	DEM	0.69	0.72	0.05	-0.00	0.00	0.00	0.00
	LAI	0.98	-0.07	-0.11	-0.09	0.01	0.03	0.00
	MSAVI	0.99	-0.09	-0.03	0.01	0.06	-0.01	-0.00
	NDVI	0.97	-0.01	-0.19	0.11	-0.02	0.01	-0.00
	NDWI	-0.99	0.04	0.12	0.07	0.03	0.02	0.00
	PVI	0.98	-0.13	0.17	0.04	0.00	-0.00	0.00

Table 2. Principal Component Analysis results showing standardized indicators, component loadings, and variance explained.

Only the leading PCs (i.e., those cumulatively explaining the majority of variance) were retained for subsequent modeling. This step simultaneously reduced dimensionality and eliminated redundancy, while preserving the dominant structure of the predictor space, and it also decreased computational load by restricting RF training to a smaller, more informative set of predictors.

While RF is inherently robust to multicollinear predictors, PCA was applied primarily to stabilize variable importance estimates and to reduce computational load, consistent with the objective of developing a lightweight, transferable framework. This choice entails a trade-off, i.e., direct interpretability of individual predictors is reduced, though PC loadings (Table 2) preserve insight into the dominant spectral drivers.

The first three components have been selected for the Landsat LST model, and the first four components for the ASTER LST model. Selected PCs have been used as input features in the RF regression model. This approach ensures that the model operates on a compact, non-collinear set of predictors, thereby improving computational efficiency and stabilizing variable importance measures.

2.2.2 Random Forest Regression for Downscaling

Random Forest Regression for Downscaling: The core of the downscaling methodology employed the RF regressor, an ensemble machine learning technique widely implemented in remote sensing for its robustness in modeling complex, non-linear spatial relationships (Hutengs & Vohland, 2016; Mao et al., 2022; Yang et al., 2017; Zhu et al., 2025). This technique was accessed using the RandomForestRegressor class from the scikit-learn (sklearn) Python package, which provides an efficient and standardized framework for machine learning applications.

The strong predictive performance of the Random Forest model arises from two core ensemble mechanisms enabled by default in the sklearn implementation. Bagging, or bootstrap aggregating, involves training each decision tree on a unique bootstrap sample of the training data, introducing variation between trees. Feature randomness is applied by considering only a random subset of predictor variables at each split node, which reduces correlation between trees, lowers prediction variance, and enhances generalization.

The prepared and vectorized dataset was initially divided into two distinct, non-overlapping subsets: a training set (80%) used exclusively for model calibration, and a testing set (20%) reserved for independent validation and accuracy assessment. Hyperparameter optimization employed randomized search with 5-fold cross-validation across 50 parameter combinations, targeting the coefficient of determination (R^2) as the optimization metric.

Key parameters were not fixed a priori but defined as search ranges, allowing the cross-validation procedure to select the optimal values independently for each model. The search space included the number of trees (n_estimators: 50, 100, 150, 200), maximum tree depth (None, 10, 20, 30), minimum samples per split (2, 5, 10), minimum samples per leaf node (1, 2, 4), bootstrap sample fraction (0.6, 0.8, 1.0), and the number of features considered at each split (None, sqrt, log2). A fixed random state of 42 was applied to guarantee reproducibility. The model was trained by fitting the selected suite of 10-m predictors (X) to the corresponding smooth LST values (Y) within the training dataset.

Once the ensemble was trained, it was used to execute the final downscaling step. The fitted RF model was applied across the entire spatial domain of the 10-m predictor rasters. For any given 10-m pixel, the final downscaled LST prediction was

computed as the average of the individual LST predictions outputted by all trees in the optimized ensemble.
This final prediction yielded a continuous, spatially explicit 10-m resolution downscaled LST product that integrates the thermal characteristics of the coarse L9 and ASTER data with the fine spatial detail captured by the high-resolution Sentinel-2 and auxiliary features.

2.2.3 Comparison with Alternative Downscaling Methods: To contextualize the performance of the RF-based framework, three alternative downscaling approaches were implemented on the same datasets and evaluated under a unified assessment protocol: (a) bilinear interpolation, representing the simplest spatial resampling baseline with no predictor information; (b) TsHARP (Thermal Sharpening), a classical index-based method that exploits the NDVI-LST relationship through a second-order polynomial fitted at the native thermal resolution and applied to fine-resolution NDVI, with bilinear-resampled residual correction (Agam et al., 2007; Kustas et al., 2003); and (c) ordinary least squares (OLS) linear regression, using the same PCA-transformed predictors and the same 80/20 train-test partition (random_state = 42) as the RF model, thereby isolating the contribution of non-linear modeling.
Two complementary evaluation strategies were adopted depending on the nature of each method. For predictor-based models (RF and linear regression), which infer LST from independent spectral features, standard point-based accuracy metrics (R^2, MAE, RMSE, MBE) were computed on the held-out 20% test set.
For methods that derive their output directly from the coarse LST signal (bilinear interpolation and TsHARP), point-based evaluation at native-resolution pixel centroids would yield artificially inflated accuracy due to circular dependency between input and output values.
Therefore, all four methods were additionally evaluated through (i) an upscaling consistency check, in which each 10 m product was spatially aggregated back to the native thermal resolution and compared with the original LST, and (ii) a sub-pixel variance analysis, quantifying the within-pixel thermal heterogeneity introduced at 10 m within each coarse pixel.
For Landsat 9, the upscaling consistency check was performed at 30 m, corresponding to the spatial resolution of the distributed Level-2 product (native TIRS resolution: 100 m, resampled to 30 m by USGS). For ASTER, the check was performed at the native sensor resolution (~100 m). Together, these two approaches assess both the fidelity of the downscaled products to the source thermal data and their capacity to resolve genuine sub-pixel thermal heterogeneity.

3. Results

3.1 RF Downscaling Performance

The trained model was applied to the entire study area through pixel-wise prediction, generating continuous 10 m LST maps. Post-processing included Gaussian smoothing (σ=0.8, kernel radius=2) to reduce high-frequency noise while preserving thermal gradients.
Figure 2 displays the RF feature importance scores for the retained principal components, for both models. In the case of Landsat 9 model, PC1 emerges as the dominant predictor, while PC2 and PC3 contribute only marginally, highlighting that most of the predictive signal is concentrated in the first component. On the other hand, the ASTER model PC1 and PC2 carry the largest contributions, while PC3 and PC4 provide smaller but non-negligible predictive value, suggesting that model performance benefits from multiple components beyond the first one alone.

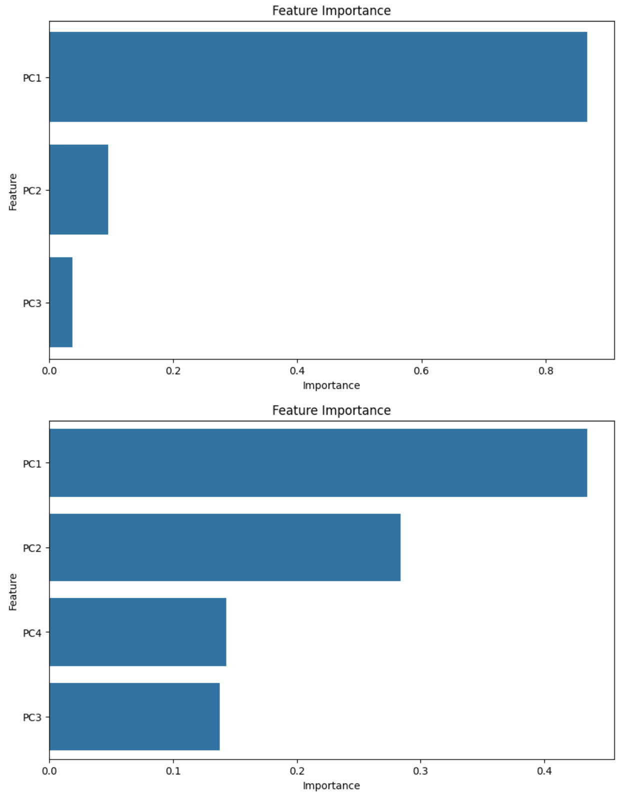

Figure 2. Feature importance of principal components in the Random Forest models, using Landsat 9 predictors (top) and ASTER predictors (bottom).

Model performance was evaluated based on a set of measures. Table 3 reports the evaluation metrics for the RF models trained on L9 and ASTER LST data. The results include goodness-of-fit (R^2), error metrics (MAE, RMSE, MBE), and their normalized counterparts, allowing direct comparison of predictive accuracy across datasets.

Models Evaluation Report	L9 LST	ASTER LST
Best R^2 Score	0.75	0.48
Training R^2	0.77	0.68
Testing R^2	0.75	0.50
MAE	1.7	0.6
RMSE	2.3	0.8
MBE	-0.03	0.00
n-MAE	0.06	0.02
n-RMSE	0.07	0.03
n-MBE	-0.00	0.00

Table 3. Evaluation metrics for the RF models trained on Landsat 9 (daytime) and ASTER (nighttime) LST data.

The application of the RF regression successfully produced a high-resolution downscaled LST product at a 10-meter spatial resolution for both daytime (Landsat 9) and nighttime (ASTER) acquisitions. This enhanced resolution provides a detailed view of the thermal heterogeneity across the study area. The results, visualized in Figure 3, clearly illustrate the stark diurnal thermal variation driven by surface energy balance and land cover type.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-355-2026 | © Author(s) 2026. CC BY 4.0 License.

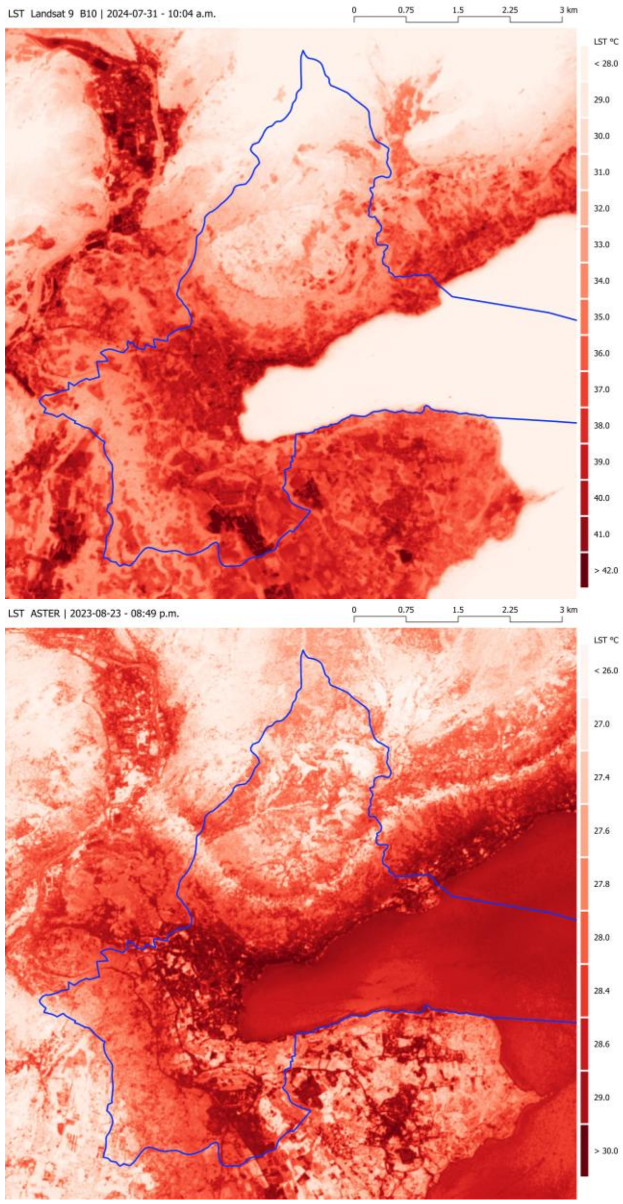

Figure 3. Comparison of the downscaled LST product at 10-meter resolution over the study area: (Top) Daytime LST (e.g., July 31st, 2024, 10:00 AM local time; (Down) Nighttime LST (e.g., August 23rd, 2023, 08:49 PM local time. Blue line shows the administrative boundaries of Salò.

During the day, the LST exhibits a wide range, reflecting intense solar heating. Areas characterized by low vegetation or impervious surfaces (e.g., urban centers, bare soil) show significantly higher temperatures, demonstrating the SUHI effect and the impact of low albedo surfaces. Conversely, vegetated areas and water bodies maintain lower, more stable temperatures due to evapotranspiration and higher thermal inertia, respectively.
At night, as the sun's heating stops and surfaces lose energy primarily through thermal radiation, the intense temperature contrasts seen during the day fade, causing the overall LST range to narrow. However, areas with high thermal inertia, such as dense urban cores and rock formations, cool more slowly and remain warmer than the surrounding suburban or rural areas (Arellano & Roca, 2022). This highlights the importance of the downscaling methodology in capturing the microclimatic variations necessary for accurate studies of thermal comfort, nocturnal energy consumption, and ecosystem heat stress. The high resolution of the downscaled product allows for precise correlation between specific features (e.g., individual parks, building footprints, street networks) and localized thermal anomalies.

3.2 Comparative Assessment of Downscaling Methods

Table 4 summarizes the comparison of the four downscaling methods for the Landsat 9 daytime model, and Table 5 for the ASTER nighttime model.

Model	Pt. R^2	Pt. MAE	Up. R^2	Up. MAE	Up. RMSE	Up. MBE	Sub-px.Var.
Bilinear	--	--	0.99	0.08	0.46	-0.04	0.053
TsHARP	--	--	0.98	0.40	0.57	-0.24	1.245
LR	0.36	2.86	0.40	2.77	3.54	-0.02	0.798
RF	**0.75**	**1.7**	**0.93**	**0.87**	**1.23**	**-0.00**	**0.261**

Table 4. Comparison of downscaling methods for Landsat 9 daytime LST. Pt. = point-based; Up. = upscaling consistency (10 m aggregated to 30 m).

Model	Pt. R^2	Pt. MAE	Up. R^2	Up. MAE	Up. RMSE	Up. MBE	Sub-px.Var.
Bilinear	--	--	0.97	0.16	0.22	-0.01	0.037
TsHARP	--	--	0.97	0.16	0.21	-0.03	0.105
LR	0.22	0.77	0.24	0.80	1.03	0.01	0.050
RF	**0.50**	**0.60**	**0.35**	**0.73**	**0.95**	**0.01**	**0.049**

Table 5. Comparison of downscaling methods for ASTER nighttime LST. Pt. = point-based; Up. = upscaling consistency (10 m aggregated to ~100 m).

For the daytime model (Table 4), the RF approach achieves the highest point-based predictive accuracy (R^2 = 0.75, MAE = 1.7°C), substantially outperforming the linear regression baseline (R^2 = 0.36, MAE = 2.86°C), which uses the same PCA-transformed predictors. This difference underscores the capacity of nonlinear modeling to capture the complex relationships between spectral predictors and surface temperature. The upscaling consistency check confirms that the RF output remains well-aligned with the original 30 m Landsat 9 LST (R^2 = 0.93, MBE = -0.0004°C), indicating a practically unbiased downscaled product. Bilinear interpolation achieves near-perfect upscaling consistency (R^2 = 0.99) but introduces negligible sub-pixel variance (0.053 °C²), confirming that it merely redistributes the coarse thermal signal without adding spatial detail. TsHARP, despite a reasonable upscaling R^2 of 0.98, exhibits the highest sub-pixel variance among all methods (1.245 °C²), more than four times the RF value (0.261 °C²), indicating excessive spatial noise rather than meaningful thermal heterogeneity. This is consistent with the limited explanatory power of the NDVI-LST relationship in heterogeneous urban environments where impervious surfaces dominate.
For the nighttime model (Table 5), the overall accuracy of all methods decreases, reflecting the inherently weaker relationship between daytime-acquired optical predictors and nocturnal surface temperatures. The RF model retains the best point-based performance (R^2 = 0.50, MAE = 0.6°C) compared to linear regression (R^2 = 0.22, MAE = 0.77°C). The TsHARP regression at ~100 m yields R^2 = 0.24, confirming that NDVI alone explains less than a quarter of the nighttime LST variance. In terms of sub-pixel variance, nighttime differences between methods are less pronounced than during the daytime, consistent with the narrower nocturnal thermal range. The RF sub-pixel variance (0.049 °C²) is comparable to that of linear regression (0.050 °C²) and bilinear interpolation (0.037 °C²),

while TsHARP again introduces the highest variance (0.105 °C²). The RF upscaling consistency ($R^2 = 0.35$) is lower than that of bilinear (0.965) and TsHARP (0.97), which is expected given that the RF predicts LST from independent spectral predictors rather than directly transforming the coarse thermal signal. Despite this, the RF nighttime MAE of 0.73°C and near-zero MBE (0.01°C) confirm that the downscaled product remains practically useful for relative hotspot identification.

4. Discussion

4.1 Key Findings and Relevance for Urban Planning

The findings of this study carry significant implications for both theoretical advancements in urban climate science and practical applications in urban planning. The successful integration of Sentinel-2, Landsat 9, and ASTER data through a RF regression model demonstrates that machine learning can effectively bridge the resolution gap between optical and thermal remote sensing data. This capability can be particularly relevant for small and medium-sized cities like Salò, where thermal patterns operate at scales finer than traditional thermal data.

The effectiveness of the implemented methodology is visually demonstrated in Figure 4, which compares the original Landsat 9 LST data with the final 10-meter downscaled product against a high-resolution, true-color reference image. This comparison provides compelling evidence of the method's ability to resolve fine-scale thermal heterogeneity that is obscured in the coarse input data.

The downscaled results distinctly reveal a strong concentration of heat within high-density built-up areas, with critical thermal hotspots consistently identified across the historic center, large parking facilities, and commercial zones. Within these specific environments, the intensity of the SUHI reaches its most significant peak: approximately 14-15°C during the daytime and still registers a notable 4-5°C intensity at night.

Such large-magnitude temperature variation gradients between adjacent urban and vegetated surfaces are physically unresolvable by the native resolutions of the Landsat 9 and ASTER thermal bands. The coarse data from these original sensors mainly capture averaged temperature value over a much larger area, which effectively masks critical micro-scale variations.

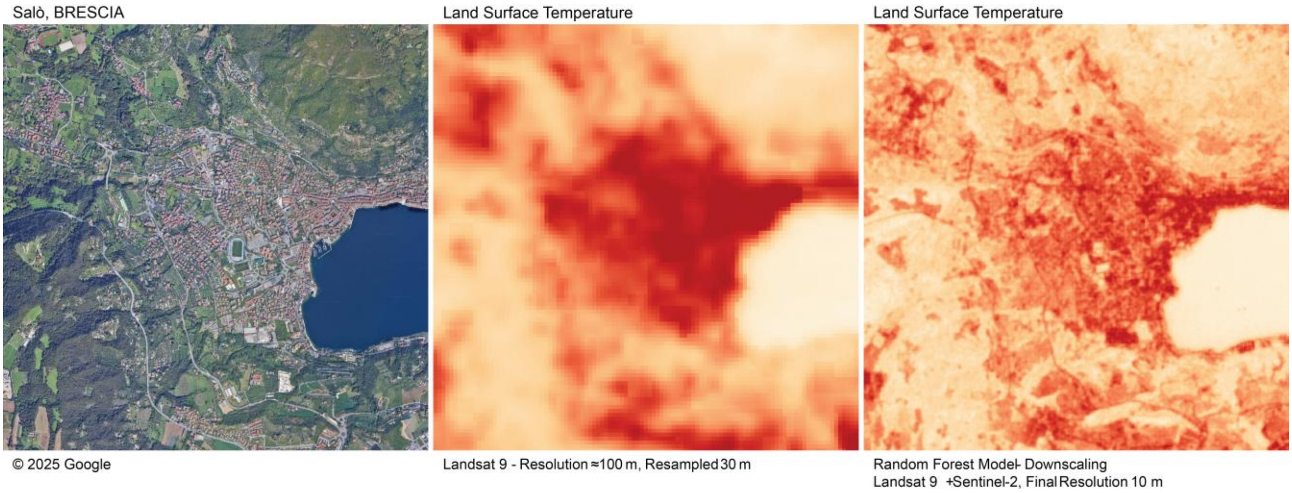

Figure 4. Visual comparison of the spatial enhancement achieved by RF downscaling. (left) High-resolution reference image (Google Satellite View), showing true-color land cover features; (middle) Original Landsat 9 LST, illustrating the coarse 100-meter (resampled to 30-meter) resolution and blocky thermal patterns resulting from the sensor's Instantaneous Field of View (IFOV); (right) Downscaled LST product (10-meter resolution), generated by the RF regression method.

While available LST products capture broad thermal patterns, the spatial detail is insufficient to correlate LST directly with individual features such as buildings, roads, or small green spaces visible in the reference image. In sharp contrast, the output of the RF model, the 10-meter downscaled LST product reveals distinct thermal boundaries that align precisely with the land cover features seen in the high-resolution reference image. This visual enhancement confirms that the RF model effectively learned the non-linear relationship between the biophysical variables and the thermal signal, thereby providing the necessary spatial detail for accurate urban microclimate and heat stress analyses.

From a practical standpoint, the 10 m resolution LST maps provide urban planners and policymakers with actionable insights for heat mitigation. The identification of specific hotspots, such as the historic center and large commercial areas, enables targeted interventions like cool pavement installation or strategic tree planting.

Illustrative examples of this thermal amplification include the expansive parking area ("*Parcheggione*") situated near the *Gli Ulivi* retirement home and the adjacent Salò 2 shopping center. Correspondingly high SUHI values are also evident across industrial districts. This pervasive accumulation of heat in structurally dense and low-albedo areas strongly underscores the urgent need for targeted urban planning and mitigation strategies designed to counteract the intensification of the heat island phenomenon in specific urban sectors.

Conversely, temperatures are demonstrably lower in areas characterized by rich green infrastructure and extensive tree canopy cover, confirming the essential role of vegetation in buffering and mitigating urban overheating. This cooling effect is particularly pronounced within the "*Canottieri Garda*" sports complex, where the vast sports park registers the lowest temperatures within the domain. Here, the presence of expansive lawns effectively reduces the SUHI intensity to 10-12°C during the day and 2-3°C at night. This observed 2°C reduction in heat island intensity is a vital finding, translating directly into tangible benefits for urban thermal comfort and overall public health.

Municipal authorities could use high-resolution thermal maps to prioritize retrofitting projects in areas with persistent thermal anomalies, thereby maximizing the cost-effectiveness of climate adaptation measures.

The methodological framework presented here extends beyond UHI monitoring, with potential applications in related domains. The high-resolution LST data could enhance urban energy demand modeling by providing precise inputs for building cooling load simulations (Kolokotroni et al., 2012). Similarly, the approach could be adapted to monitor thermal pollution in water bodies or assess the microclimatic impacts of green infrastructure projects, thereby supporting broader environmental management objectives (Palmer et al., 2015).

While this study focused on a single case study, the framework is transferable to other cities with similar climatic characteristics. The consistent performance across daytime and nighttime conditions—despite differences in underlying physical processes—suggests that the RF-based downscaling approach is robust enough for application in diverse settings.

4.2 Study Limitations and Future Directions

The integration of multi-sensor data through machine learning opens promising possibilities for urban climate research and planning but also raises important questions about data harmonization and computational scalability. The successful fusion of Sentinel-2, Landsat 9, and ASTER data in this study demonstrates that cross-platform compatibility can be achieved through careful preprocessing.

Nevertheless, a number of methodological limitations must be acknowledged when interpreting these results. The reliance on cloud-free satellite imagery and the need for temporal alignment between optical and thermal data from different sensors, restricted the analysis to two specific dates, potentially overlooking seasonal variations in SUHI dynamics. It should be noted that the daytime and nighttime analyses correspond to different dates and years (July 2024 and August 2023, respectively), driven by the constraint of acquiring cloud-free, temporally aligned optical-thermal images. Consequently, the daytime and nighttime results should not be interpreted as a paired diurnal cycle analysis of a single event, but rather as two independent demonstrations of the downscaling framework under distinct hot-summer conditions. Future applications would benefit from same-day or near-contemporaneous acquisitions to enable direct diurnal SUHI characterization. While the selected dates represent hot summer conditions, the absence of multi-temporal validation limits the generalizability of the model’s performance across different meteorological scenarios.

The comparatively lower predictive accuracy of the nighttime model (R^2=0.50) reflects the inherently more complex thermal dynamics governing nocturnal surface temperatures. Unlike daytime LST, which is strongly driven by solar irradiance and surface optical properties well captured by Sentinel-2 spectral indices, nighttime LST is primarily governed by thermal inertia, anthropogenic heat release, and three-dimensional urban canyon geometry (Yuan et al., 2020). At night, in the absence of solar radiation, the urban thermal environment is shaped by reduced radiative cooling due to low sky view factors, the release of heat stored in urban materials, and the proportionally greater influence of anthropogenic heat emissions, factors that two-dimensional spectral predictors alone cannot capture. Indeed, nighttime LST downscaling remains comparatively underexplored in the literature, partly due to these inherent complexities and the limited availability of fine-resolution nighttime thermal data (Arellano & Roca, 2021; Yoo et al., 2022). Despite this limitation, the nighttime MAE of 0.6°C and near-zero MBE indicate that the model remains practically useful for identifying relative thermal hotspots, even if absolute accuracy is reduced.

Additionally, the 10 m resolution, while a substantial improvement over native thermal data, may still be insufficient to resolve micro-scale thermal features such as individual building facades or narrow alleys that contribute to pedestrian-level heat exposure. Consequently, these results should be interpreted as most relevant to city-scale planning rather than micro-scale assessments. The integration of higher-resolution optical data from platforms like PlanetScope (3 m) could further enhance LST mapping precision, particularly for capturing thermal gradients along streetscapes and building edges. Incorporating 3D urban morphology metrics such as the sky view factor and building height-to-width ratios could improve nighttime LST predictions by better accounting for urban canyon effects (Wu et al., 2022).

The application of PCA prior to RF training, while beneficial for computational efficiency and importance measure stability, may have constrained the model’s predictive capacity. A direct comparison of RF performance with and without PCA-transformed inputs would help quantify this trade-off and is recommended for future investigations.

Future research should address these limitations while building upon the current findings. There is a need for longitudinal studies that apply the RF downscaling approach across multiple seasons and different geographical areas to assess its robustness under varying climatic conditions (Faraji et al., 2025; Wen et al., 2025). The moderate computational requirements of the RF model suggest feasibility for operational use. However, the development of real-time or near-real-time downscaling pipelines would significantly increase the operational utility of this approach for heat emergency response and urban climate monitoring.

5. Conclusions

This study demonstrates that integrating multi-sensor remote sensing data with machine learning can significantly enhance the spatial resolution of land surface temperature monitoring in urban environments. Emphasis was placed on the assessment of the SUHI phenomenon for medium and small-sized cities.

The Random Forest-based downscaling approach successfully bridged the resolution gap between Sentinel-2 optical data and Landsat 9/ASTER thermal observations, achieving high predictive accuracy while resolving critical micro-scale thermal patterns at 10 m resolution. The findings confirm that spectral indices and land cover characteristics serve as robust proxies for LST estimation, particularly in heterogeneous urban landscapes, as confirmed by the comparative assessment presented in this study.

The methodological framework presented here provides a scalable solution for cities seeking to develop precision heat mitigation strategies, with potential applications extending to urban energy modeling and green infrastructure performance assessment. By transforming coarse thermal data into actionable high-resolution insights, this approach advances the integration of climate science into urban planning practice, offering a pathway toward more resilient cities in an era of escalating heat risks.

Ultimately, the findings contribute to the growing body of research on urban climate adaptation, particularly for Mediterranean cities facing increasing heat stress under climate change. The ability to identify specific heat-vulnerable areas at fine spatial scales represents a critical step towards implementing the European Climate Adaptation Strategy at the local level. By providing planners with spatially explicit evidence of UHI effects, this approach can help bridge the gap between climate science and urban design practice, ensuring that mitigation measures are both scientifically informed and contextually appropriate.

Acknowledgements

This research was carried out with the support of the City of Salò, the Department of Urban Planning, Private Construction, and Public Works. The authors gratefully acknowledge their valuable contribution to the development of this study.

References

Agam, N., Kustas, W.P., Anderson, M.C., Li, F., Neale, C.M.U., 2007. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sensing of Environment* 107(4), 545–558.

Andriambololonaharisoamalala, R.R., Helmholz, P., Bulatov, D., Ivanova, I., Song, Y., Soon, S., Jones, E., 2025. Downscaling of Urban Land Surface Temperatures Using Geospatial Machine Learning with Landsat 8/9 and Sentinel-2 Imagery. *Remote Sensing* 17(14), 2392.

Arellano, B., Roca, J., 2021. Remote sensing and night time urban heat island. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XLIII-B3-2021(B3-2021), 15–22.

Arellano, B., Roca, J., 2022. Effects of urban greenery on health. A study from remote sensing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XLIII-B3-2022(B3-2022), 17–24.

Bechtel, B., Zakšek, K., Hoshyaripour, G., 2012. Downscaling Land Surface Temperature in an Urban Area: A Case Study for Hamburg, Germany. *Remote Sensing*, 4(10), 3184-3200.

Ebrahimy, H., Azadbakht, M., 2019. Downscaling MODIS land surface temperature over a heterogeneous area: An investigation of machine learning techniques, feature selection, and impacts of mixed pixels. *Computers & Geosciences*, 124, 93–102.

Essa, W., Verbeiren, B., van der Kwast, J., Van de Voorde, T., Batelaan, O., 2012. Evaluation of the DisTrad thermal sharpening methodology for urban areas. *International Journal of Applied Earth Observation and Geoinformation*, 19(1), 163–172.

Faraji, Z., Kaviani, A., Khosravi, L., 2025. Advanced spatiotemporal downscaling of MODIS land surface temperature: utilizing Sentinel-1 and Sentinel-2 data with machine learning technique in Qazvin Province, Iran. *Environmental Monitoring and Assessment*, 197(8), 1–17.

Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, 178, 127–141.

Kolokotroni, M., Ren, X., Davies, M., Mavrogianni, A., 2012. London’s urban heat island: Impact on current and future energy consumption in office buildings. *Energy and Buildings*, 47, 302–311.

Kustas, W.P., Norman, J.M., Anderson, M.C., French, A.N., 2003. Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. *Remote Sensing of Environment*, 85(4), 429–440.

Liang, S., 2001. Narrowband to broadband conversions of land surface albedo I: Algorithms. *Remote Sensing of Environment*, 76(2), 213–238.

Mao, T., Shangguan, W., Li, Q., Li, L., Zhang, Y., Huang, F., Li, J., Liu, W., Zhang, R., 2022. A Spatial Downscaling Method for Remote Sensing Soil Moisture Based on Random Forest Considering Soil Moisture Memory and Mass Conservation. *Remote Sensing*, 14(16), 3858.

Palmer, S.C.J., Kutser, T., Hunter, P.D., 2015. Remote sensing of inland waters: Challenges, progress and future directions. *Remote Sensing of Environment*, 157, 1–8.

Roy, D., Das, B., Singh, P., Santra, P., Deb, S., Bhattacharya, B.K., Govind, A., Jatav, R., Sethi, D., Ghosh, T., Mukherjee, J., Sehgal, V.K., Jha, P.K., Goroshi, S., Prasad, P.V.V., Chakraborty, D., 2025. Assessing the accuracy of multi-model approaches for downscaling land surface temperature across diverse agroclimatic zones. *Scientific Reports*, 15(1), 1–15.

Uhrin, A., Onačillová, K., 2025. Spatiotemporal analysis of land surface temperature and land cover changes in Prešov city using downscaling approach and machine learning algorithms. *Environmental Monitoring and Assessment*, 197(2), 1–25.

Wen, J., He, Y., Yang, L., Wan, P., Gu, Z., & Wang, Y., 2025. A Two-Step Downscaling Model for MODIS Land Surface Temperature Based on Random Forests. *Atmosphere*, 16(4), 424.

Wu, W. Ben, Yu, Z. W., Ma, J., Zhao, B., 2022. Quantifying the influence of 2D and 3D urban morphology on the thermal environment across climatic zones. *Landscape and Urban Planning*, 226, 104499.

Xu, J., Zhang, F., Jiang, H., Hu, H., Zhong, K., Jing, W., Yang, J., Jia, B., 2020. Downscaling Aster Land Surface Temperature over Urban Areas with Machine Learning-Based Area-To-Point Regression Kriging. *Remote Sensing*, 12(7), 1082.

Xu, J., Zhang, F., Ruan, H., Hu, H., Liu, Y., Zhong, K., Jing, W., Yang, J., Liu, X., 2021. Hybrid modelling of random forests and kriging with sentinel-2A multispectral imagery to determine urban brightness temperatures with high resolution. *International Journal of Remote Sensing*, 42(6), 2174–2202.

Yang, Y., Cao, C., Pan, X., Li, X., Zhu, X., 2017. Downscaling Land Surface Temperature in an Arid Area by Using Multiple Remote Sensing Indices with Random Forest Regression. *Remote Sensing*, 9(8), 789.

Yoo, C., Im, J., Cho, D., Lee, Y., Bae, D., Sismanidis, P., 2022. Downscaling MODIS nighttime land surface temperatures in urban areas using ASTER thermal data through local linear forest. *International Journal of Applied Earth Observation and Geoinformation*, 110, 102827.

Yuan, C., Adelia, A.S., Mei, S., He, W., Li, X.X., Norford, L., 2020. Mitigating intensity of urban heat island by better understanding on urban morphology and anthropogenic heat dispersion. *Building and Environment*, 176, 106876.

Zhu, H., Zhou, Q., Krisp, J.M., 2025. Exploring machine learning approaches for precipitation downscaling. *Geo-Spatial Information Science*, 1-17.