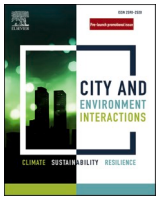
City and Environment Interactions 20 (2023) 100124

Contents lists available at ScienceDirect

City and Environment Interactions

journal homepage: www.sciencedirect.com/journal/city-and-environment-interactions



Insights into heat islands at the regional scale using a data-driven approach

Nicola Colaninno

Department of Architecture and Urban Studies, Politecnico di Milano, Milan, Italy
Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA

A R T I C L E I N F O

Keywords:
Regional heat islands
Regional planning
Urban climate
Resilient landscapes
Green infrastructures
Adaptation

A B S T R A C T

Urban heat island (UHI) phenomenon is crucial in the context of climate change. However, while substantial attention has been given to studying UHIs within cities, our understanding at the regional level still needs to be improved. This study delves into the intricate dynamics of the regional heat island (RHI) by examining its relationship with land use/land cover (LULC), vegetation, and elevation. The objective is to enhance our knowledge of RHI to inform effective mitigation strategies. The research employs a data-driven approach, leveraging satellite data and spatial modeling, examining surface and canopy-layer regional heat islands, and considering daytime and nighttime variations. To assess the impact of LULC, the study evaluates three main categories: anthropized (urbanized), agricultural, and wooded/semi-natural environments. Furthermore, it delves into the influence of vegetation on RHI and incorporates elevation data to understand its role in RHI intensity. The findings reveal meaningful variations in heat islands across different LULCs, providing essential insights. Although urbanized areas exhibit the highest RHI intensity, agricultural regions contribute notably to RHI due to land use changes and reduced vegetation cover. This emphasizes the significant impact of human activities. In contrast, wooded and semi-natural environments demonstrate potential for mitigating RHI, owing to their dense vegetation and shading effects. Elevation, while generally associated with reduced heat island, shows variations based on local conditions. Ultimately, this research underscores the complexity of the RHI phenomenon and the importance of considering factors such as different temperatures and their daily variation, landscape heterogeneity, and elevation. Additionally, the study emphasizes the significance of sustainable spatial planning and land management. Targeted efforts to increase vegetation in high daytime land surface temperature areas can reduce heat storage and mitigate RHI. Similarly, planning for agroforestry and green infrastructure in agricultural areas can significantly increase resilience to climate.

1. Introduction

Cities play a pivotal role in addressing the challenges posed by climate change, as they contribute to its causes and face its consequences. Global warming severely impacts cities, negatively affecting human health and well-being. Urban areas have unique characteristics influencing the urban climate, resulting in higher temperatures than surrounding rural and *peri*-urban regions. This phenomenon is known as the urban heat island (UHI). When studying the UHI, it is essential to consider three levels: the canopy layer UHI (CLUHI), which focuses on the air temperature between urban roughness components; the boundary layer UHI (BLUHI), which refers to the air temperature above the roofs; and the surface UHI (SUHI), which pertains to the warming of urban surfaces [27,33,41].

In recent decades, the urban heat island has gained significant attention in urban climatology and planning [32,35]. However, most studies have focused on the UHI phenomenon within cities, while there is still a need to explore its dynamics at the regional scale and in metropolitan areas. While UHI analysis studies have a history dating back to 1972, with increased momentum since 2010, studies explicitly focusing on the Regional Heat Island (RHI) have been more active since 2019 [10]. This research aims to reduce the notable gap in the body of literature regarding the dynamics of the heat island phenomenon at the regional scale.

The term regional heat island was proposed by Yu et al. [46,45] to describe significant variations in thermal conditions within urban agglomerations due to reduced distances between cities. They report that the heat island effect in cities can extend beyond city boundaries and form a larger regional heat island, particularly as cities merge into metropolitan systems. Also, they identify the regional heat island as an area that manifested a relative land surface temperature of more than 2 °C with respect to the average value of the whole area [11,46].

E-mail address: nicola.colaninno@polimi.it.

https://doi.org/10.1016/j.cacint.2023.100124
Received 20 June 2023; Received in revised form 24 September 2023; Accepted 10 October 2023
Available online 15 October 2023

However, because of the recent interest in the topic, a comprehensive exploration of how multifaceted aspects, including ecological, climate, and socioeconomic factors [4], contribute to the formation of RHI still needs to be represented. Indeed, Yu et al. [45,46] mainly based the study on land surface temperature, while Degefu et al. [10], albeit providing a relevant analysis that includes the effects of green space and land use/land cover on the urban thermal environment, focused the attention on a relatively limited area, considering an 8 km long transect based on a mobile traverse.

A relevant resource to broaden our knowledge of the phenomenon relies on remote sensing data and techniques, which have been extensively employed over the last decades to study heat islands [10]. In particular, land surface temperature has been largely assessed based on thermal infrared-derived imagery from satellites such as Landsat, ASTER, and MODIS [10,12,47]. MODIS data has been utilized to investigate the influences of various factors on the trends of RHI over different years [4]. The modeling of Surface UHI using MODIS LST has been approached through methods such as multiple linear regression at different times of day and night [43], geographically weighted regression [16], and support vector machine regression considering different features like the land cover, solar radiation, temperature, humidity, precipitation, wind speed, aerosol optical depth, and soil moisture to estimate nighttime SUHI (Lai et al., 2021).

Likewise, as satellite-based optical imaging enables vegetation analysis through spectral vegetation indices, numerous studies have explored the combined use of thermal and optical data, emphasizing the negative correlation between land surface temperature (LST) and the Normalized Difference Vegetation Index (NDVI) through linear regression analysis at different spatial resolutions [32,42] and during either daytime or nighttime [39]. However, although scientific research has revealed that LST and NDVI are among the most significant remote sensing variables for assessing air temperature [9], and vegetation is a fundamental element influencing the intensity of urban heat islands, the correlation between NDVI and urban temperatures should be carefully considered. As also emerges from this research, the correlation between LST and NDVI is strongly seasonal- and time-dependent, and vegetation's cooling effect improves effectiveness mostly during daytime than nighttime [7,36].

Despite the numerous heat island studies using a remote sensing approach, another aspect deserves attention. Previous research, including studies conducted by Huang et al. [16] and Zhang & Du [47], have analyzed solely the surface heat island or have neglected the interplay between air temperature and land surface temperature in the context of heat islands. On the other hand, it has been recognized that land surface temperature alone is insufficient for accurately assessing the heat island phenomenon and its spatial implications, particularly during the daytime [29]. Indeed, near-surface air temperature is pivotal to enhancing the knowledge of the heat islands.

To overcome this gap, modeling the near-surface air temperature (NSAT), approximately 2 m above the ground, is crucial yet challenging [16]. Various approaches have been investigated. A temperature-vegetation index (TVX) method is proposed to enhance daily maximum air temperature estimation using MODIS surface temperature [49]. Machine learning techniques, such as random forest algorithms [22,44,47] and deep learning approaches [34], have recently been explored for NSAT estimation. Geographically weighted regression has also proven to be effective in modeling near-surface air temperature, both during the day and at night, as it considers spatial non-stationarity [13,16,8].

However, NSAT estimation has predominantly revolved around monthly or daily average [22] or maximum air temperature [3,34]. Only some studies have reconstructed air temperature at higher temporal scales, such as sub-daily or hourly resolutions [47], albeit at the expense of spatial resolution. Hence, a need for high spatiotemporal resolution still exists. Although this study does not entirely solve the problem, as it does not provide complete hourly temperature modeling, it contributes to the topic by proposing an operational model to estimate instantaneous NSAT for specific day and night hours, with a spatial resolution of about 900 m.

Finally, recent research has increasingly focused on the relationship between land use/land cover transformation and regional climate. Several studies have employed MODIS or Landsat data, along with NDVI and land-use variables, to model the impact of different land uses on surface temperatures [5,6,11,18]. These studies have explored various aspects, such as the contribution of LULC to regional heat islands, the mitigation of urban heat islands through different types of LULC [48], and the simulation of land-use change effects on LST [5,18]. They have also analyzed temperature patterns associated with land-use configurations and changes [23], quantified the effects of urban and green areas on regional climate change using time-series analysis of land use and land cover [20], and assessed the intensity of the urban heat island effect in relation to LULC, NDVI, and LST [17].

Given that, in most cases, the data used for heat island studies is the surface temperature, this research raises the question of why not only LST but also near-surface air temperature should be considered. This standpoint is relevant because LST is not a good proxy for the heat island phenomenon, especially during daylight hours.

This research aims to comprehensively understand the regional heat island phenomenon, considering the interplay between land surface temperature, near-surface air temperature, vegetation, and land use/land cover. The study addresses two primary research questions: how to effectively model and estimate NSAT for both day and night-time at a scale suitable for regional studies and what drives the dynamics of RHI within metropolitan areas.

A practical model for estimating instantaneous NSAT is introduced, assessing the significance of the NDVI on the model's performance. The study further explores the spatial and temporal interactions between NSAT and LST, considering the influence of vegetation, land use/land cover, and different elevations on RHI.

The novelty of this work lies in its holistic analysis of RHI that involves the combined use of near-surface air temperature and land surface temperature to enhance our understanding of the phenomenon and examine the correlation between them in the context of heat islands, considering both day and night variations, and the impact of relevant environmental factors that can shape the RHI. A comprehensive understanding of the heat island phenomenon at a regional scale is essential for informing evidence-based policies, fostering sustainable urban development, and enhancing climate resilience.

2. Materials and methods

This section outlines the methodology employed in this study to investigate the regional heat islands phenomenon, including surface and canopy layer heat islands (SRHI and CLRHI).

Using remotely sensed data, namely land surface temperature and normalized difference vegetation index from MODIS sensors on Terra and Aqua satellites [19], a Shuttle Radar Topography Mission (SRTM)-derived Digital Elevation Model (DEM) and ground-based temperature measurements, a Geographically Weighted Regression model is designed to estimate day and night near-surface air temperature.

Hence, the analysis assesses the impact of different land use and land cover (LULC) types, specifically anthropized (urbanized), agricultural, and wooded/semi-natural environments, using LST and NSAT.

To evaluate the effect of vegetation on temperature, the NDVI is integrated. The analysis of vegetation's influence on surface and canopy layer heat islands considers NDVI's correlation with LST and NSAT under daytime and nighttime conditions.

Correlation analyses between NDVI and temperature at different elevations and within various LULC categories are also conducted to assess elevation's impact on NSAT and LST.

Finally, the intensity of the regional heat island phenomenon is assessed by comparing SRHI and CLRHI within different LULC

categories and examining temperature differences between LULCs and NDVI variations concerning elevation. Fig. 1 presents a methodology flowchart summarizing the key steps, offering a visual overview of the research process.

2.1. Case study, period under investigation, and data sources

The area under investigation is the Lombardy region, northwestern Italy, the fourth Italian region by extent, encompassing around 23,860 km^2, and the first in terms of the resident population, with around 10 million residents in 2020. The region, with cold winters, no dry season, and hot summers, is classified as Cfa (humid subtropical) by the Köppen climatic classification scheme [2,28]. In Lombardy, warming has accelerated significantly in the previous 30 years, resulting in an average air temperature anomaly of around +0.2–0.3 degrees Celsius when compared to the reference period of 1968–1996 [38]. This research considers a 4-day heatwave between the 29th of July and the 1st of August 2020. The heatwave is identified by comparing the 90th percentile of the daily maximum temperature over a reference period (1973–2019) with the daily maximum temperature of each day during July and August 2020. Because the objective addresses extreme events, the focus is on the hottest days during the heat waves, i.e., the 31st of July and the 1st of August 2020.

Employed data sources rely on the Moderate-resolution Imaging Spectroradiometer (MODIS), the Shuttle Radar Topography Mission Digital Elevation Model (SRTM-DEM), weather stations' measures, and the Land Use/Land Cover (LULC).

The MODIS sensor is on board the Terra and Aqua spacecraft as part of NASA's Earth Observing System (EOS). These satellites follow a sun-synchronous, near-polar circular orbit, allowing global coverage once or twice daily. MODIS version 6, with a 1-km spatial resolution, was used in this study. It provides daily per-pixel data and includes LST products (MOD11 and MYD11) and NDVI products (MOD13A2 and MYD13A2). The NDVI data are generated from the best available pixels over 16 days.

The SRTM-DEM, obtained from the Shuttle Radar Topography Mission, offers high-resolution elevation data across 80 % of the Earth's surface between 60° north and 56° south latitude. The data are available at 1-arc-second (30-meter) resolution for the United States and 3-arc-second (90-meter) resolution worldwide. A 30-meter resolution DEM, created through resampling the high-resolution dataset, is also accessible worldwide.

Weather data from the Regional Meteorological Service (SMR), established by the Regional Environmental Protection Agency (ARPA), is used for retrieving air temperature (T_a). The SMR operates a network of 250 automated stations, providing daily meteorological and climatological data, including temperature, humidity, radiation, wind speed, and precipitation. The data, recorded at 2 meters from the ground, are obtained in 10-minute intervals.

For land use/land cover, the research employs the DUSAF (uses of agricultural and forest land) database. DUSAF is the official LULC database in Lombardy. It is constructed from aerial orthophotos and satellite images and uses SPOT6 and 7 satellite images at a 1.5-meter resolution. The database classifies LULC into hierarchical levels, with the first level (Level 1) categorizing five main classes, i.e., Anthropized, Agricultural, Wooded Territories and Semi-natural Environments, Wetlands, and Water Bodies. Subsequent levels provide further detail. In this study, the Level 1 classification is used.

Fig. 2 illustrates the summer 2020 heat waves, showcasing the two hottest days, July 31st and August 1st (Fig. 2a). Additionally, it presents the LULC configuration (2b), a MOD11A1 LST image captured at 10.42 LT on July 31st, 2020 (2c), and the SRTM DEM (2d). The figure also includes the administrative boundaries of the Lombardy region (red line in Fig. 2c and d) and the spatial distribution of the ARPA stations (cyan dots).

2.2. The significance of the NDVI for modelling near-surface air temperature

The initial focus of the research was to evaluate the explanatory power of different variables in developing an operational Geographically Weighted Regression (GWR) model for estimating the near-surface air temperature. The variables assessed included Land Surface Temperature (LST), Digital Elevation Model (DEM), and Normalized Difference Vegetation Index (NDVI). Two NDVI products, MOD13A2 and

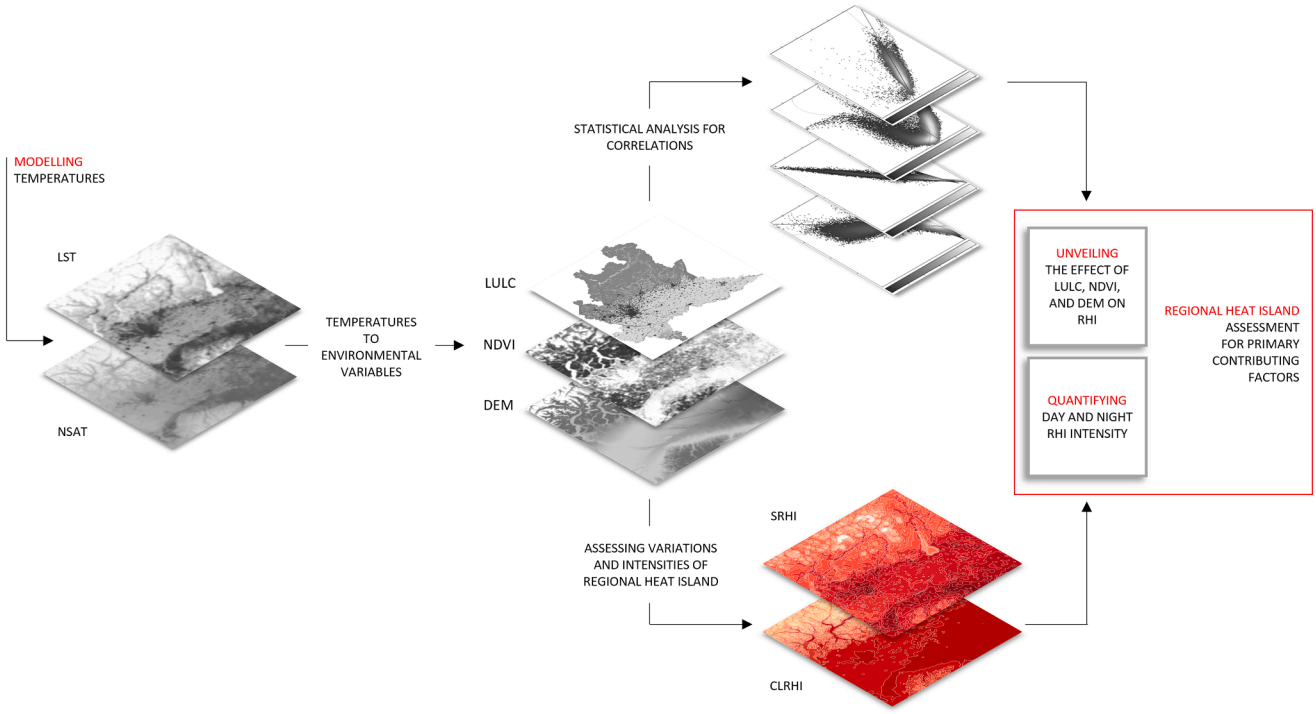


Fig. 1. Methodology Overview Flowchart.

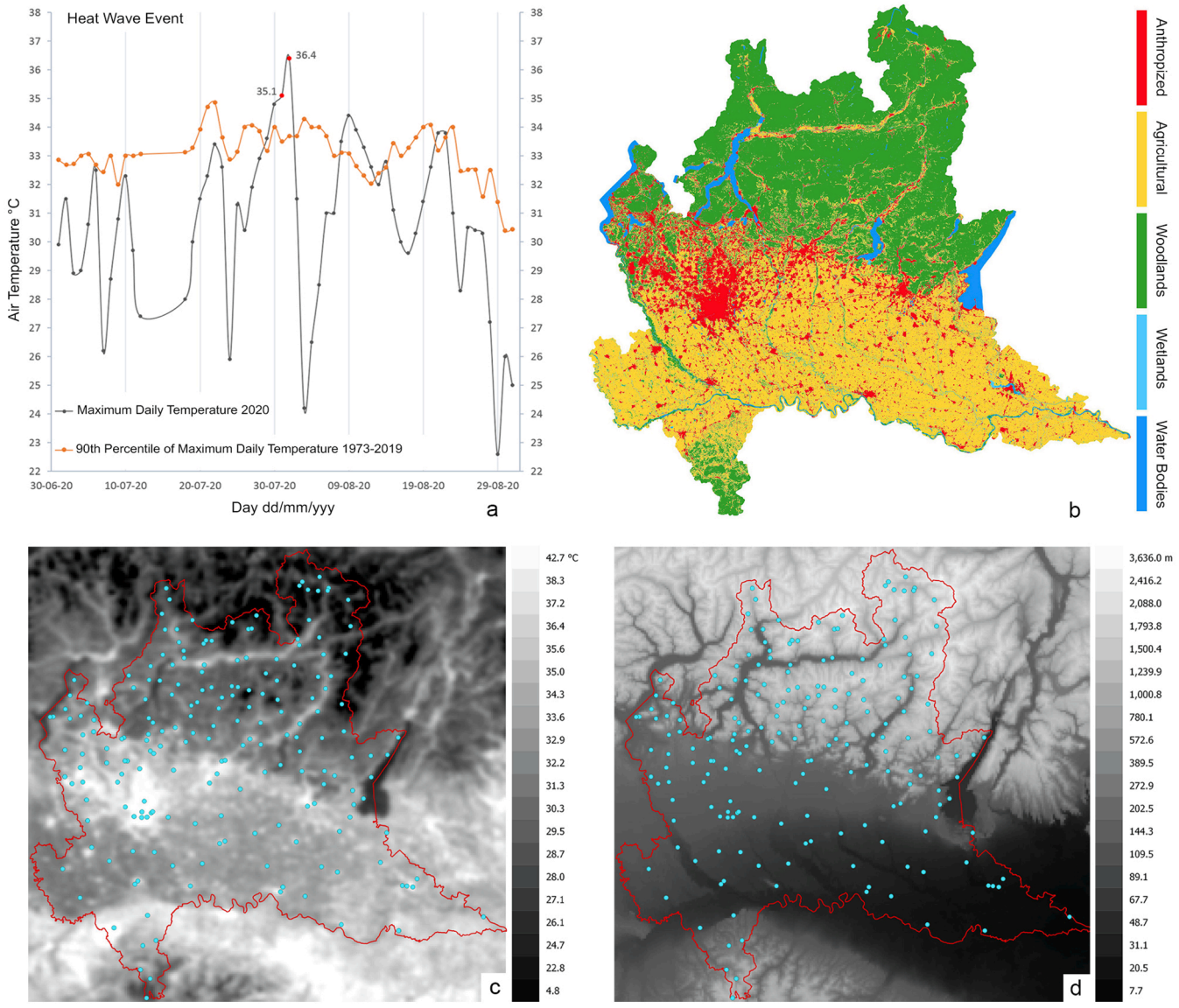


Fig. 2. a). Summer 2020 heat waves, spanning July 29th to August 1st and August 8th to 12th, along with the two hottest days on record, July 31st and August 1st; b). Land Use/Land Cover; c). MOD11A1 LST at 10.42 LT on the 31st of July 2020, and d). SRTM DEM, including administrative boundaries of the Lombardy region (red line) and ARPA weather stations (cyan dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

MYD13A2, obtained from Terra and Aqua satellites, respectively, were considered. Linear regression was conducted to assess the relationship between the weather station-derived air temperature (T_a) and the independent variables at various days and times. The evaluation included measures such as Pearson coefficient (r), coefficient of determination (R^2), and the F-test of significance, provided in Table 1, to

Table 1
Pearson correlation (r), R^2, and *F-test* significance between the dependent variable T_a and the tested predictors for each day and time, i.e., 31st of July 2020, at 10:40 and 21:50, and 1st of August 2020, at 02:00 and 13:00.

Predictors		Dependent variable - T_a - Weather stations											
		2020.07.31						2020.08.01					
		10:40			21:50			02:00			13:10		
		r	R^2	*F-test*	*r*	R^2	*F-test*	*r*	R^2	*F-test*	*r*	R^2	*F-test*
LST	10:42 - Day	**0.80**	0.65	352.2									
LST	21:48 - Night				**0.93**	0.86	1184.7						
LST	02:00 - Night							**0.91**	0.83	967.3			
LST	13:06 - Day										**0.77**	0.60	281.73
Elevation	DEM	**-0.93**	0.87	1256.9	**-0.90**	0.80	793.5	**-0.85**	0.72	518.8	**-0.93**	0.86	1240.1
NDVI	MOD13A2	-0.09	0.00	1.855	-0.27	0.07	15.253						
NDVI	MYD13A2							-0.26	0.06	14.638	-0.09	0.00	1.705

account for the correlation's direction, strength, and statistical significance.

A strong positive correlation was observed between LST and T_a, indicating that higher LST values are associated with higher T_a. Conversely, the DEM negatively correlated with T_a, with higher elevation values corresponding to lower T_a. Interestingly, the correlation between DEM and T_a was more robust in the daytime than LST, indicating that DEM significantly influences T_a during the day. The correlation between LST and T_a decreased at the peak temperature time (13:10). During the nighttime, the correlation between LST and T_a increased, suggesting that the two temperatures converged at night, with T_a being more strongly correlated with LST than DEM.

In contrast, both MODIS-derived NDVIs showed weak or almost no correlation with T_a. Although a slight negative correlation emerged during mostly nighttime, neither of the two NDVIs proved to be influential explanatory variables for T_a, whether during daytime or nighttime, for the specific days, times, and areas under investigation. Instead, the two NDVIs show a highly significant correlation with each other, with a Pearson of 0.96, slope of 1.03, and intercept of −0.03. Accordingly, only the MYD13A2 was used as the NDVI for subsequent analysis. Fig. 3 displays scatterplots of MYD13A2-NDVI against T_a at different time observations, indicating an absence of a discernible trend in the point cloud pattern. The scatterplots exhibit distinct patterns with two contrasting trends rather than a clear linear correlation. The distribution shows a positive association at lower temperatures, transitioning to a negative association as temperature increases. Ultimately, we point out an inconsistent correlation between NDVI and temperature in this experiment, and for the whole area under investigation, that prevents using NDVI for air temperature prediction.

2.3. GWR for modelling near-surface air temperature

Geographically Weighted Regression (GWR) is a widely discussed geospatial analytical technique in literature, particularly effective for examining non-stationary phenomena and investigating heterogeneity in data relationships. The GWR model calibrates coefficients and predictions locally using a neighboring region or bandwidth surrounding the target point. The local regression, described by Equation (1) [13,14,24], encompasses position i represented as a vector of coordinates in either a projected or geodetic coordinate system.

$$Y_i = \beta_{i0} + \sum_{n=1}^{m} \beta_{in} X_{in} + \epsilon_i \quad (1)$$

where Y_i is the dependent variable at location i (x_i, y_i), X_{in} is the n^{th} independent (explanatory) variable at location i, m is the number of independent variables, β_{i0} is the intercept parameter at location i, β_{in} is the local regression coefficient for the n^{th} independent variable at location i, and ε_i is the random error at location i. At each regression point i, the model's parameters (or coefficients β) are estimated locally by the weighted least squares. The weights, expressed in a matrix form, depend on the observed location with respect to the other observations in the dataset [21].

Using a GWR model, instantaneous NSAT is estimated, with MODIS-LST and DEM as the independent variables. The lack of a significant correlation between NDVIs and T_a led to the exclusion of NDVI as a predictor. The GWR model is designed according to Equation (2), where Y_i is the air temperature T_{ai} at location i (x_i, y_i), while X_{i1} and X_{i2} are LST and DEM, respectively. An exponential weighting scheme is employed, with a bandwidth of 40 Km automatically defined based on the spatial distribution of the points representing the weather stations.

$$T_{a_i} = \beta_{i0} + \beta_{i1} LST_{i1} + \beta_{i2} DEM_{i2} + \epsilon_i \quad (2)$$

Instantaneous NSAT is estimated day- and night-time, consistent with MODIS spatial and temporal resolution. MOD11A1 and MYD11A1 LST were obtained on the 31st of July and the 1st of August 2020. Four Terra/Aqua MODIS image scenes have been used, i.e., 10:42 local time (LT) and 21:48 LT for Terra on the 31st of July and 13:06 LT and 2:00 LT for Aqua on the 1st of August 2020. Due to cloud coverage, some images are severely limited, and not all days are practicable. A quick image reconstruction phase is undertaken to address small missing data due to cloud coverage in the selected images [8].

As there is a slight time difference because of the 10-minute temporal resolution discrepancy between T_a and MODIS-LST, to minimize this difference, T_a is selected at the nearest time to the MODIS local passing time. Specifically, T_a is taken at 10:40 and 21:50 on July 31, 2020, and at 02:00 and 13:10 on August 1, 2020. Hence, four models are computed. All 250 available weather stations were included in the analysis without imposing a preferred search distance to select observations. However, only 205 stations provided T_a values. Also, a few stations do not allow 10-minute observation but hourly data. Consequently, some stations were excluded at 10:40, 21:50, and 13:10. For the 13:10 model, 197 values were available. In contrast, for the 10:40 and 21:50 models, there were 195 values after excluding two stations that reported abnormal values (assigned −9999). For the 02:00 model, all 205 stations were used since it aligns with the top of the hour.

3. Results

3.1. Model assessment and estimated NSAT

The performance of the models is extensively discussed in Colaninno and Morello [8]. In this work, the effectiveness of the models is summarized in Table 2, employing a rigorous 2-fold cross-validation (CV) protocol. Three predictor combinations are examined, including LST

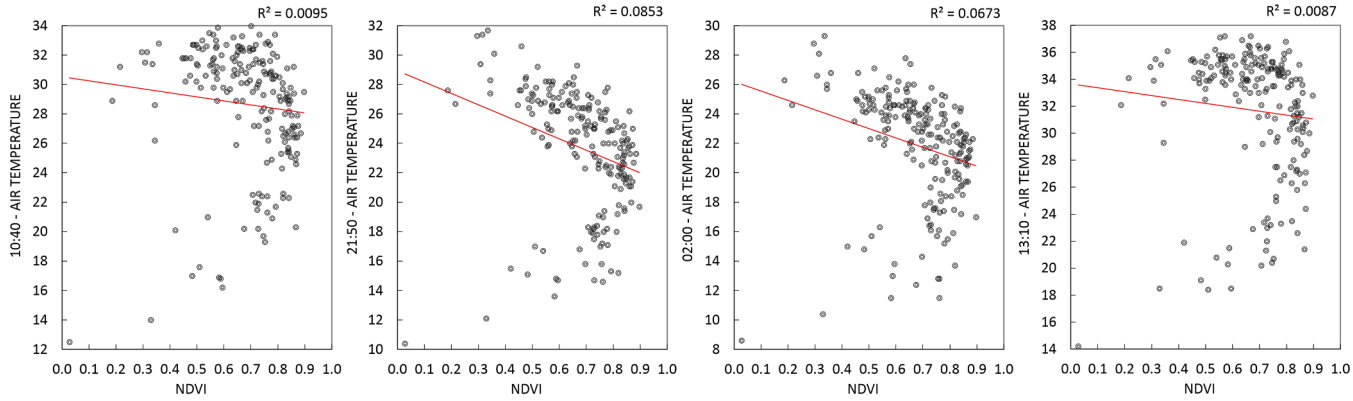


Fig. 3. Scatterplots of the correlation between T_a, measured by weather stations, and the MYD13A2-NDVI for each observation, i.e., 10:40 and 21:50 for the 31st of July, 02:00, and 13:10 for the 1st of August of 2020.

Table 2
2-Fold CV performance for each model, based on different combinations of predictors: LST and DEM, LST only, DEM only.

2-Fold Cross-Validation													
		LST and DEM				*LST*				*DEM*			
Time	n. Stations	Adj.R^2	RMSE	MAE	MBE	Adj.R^2	RMSE	MAE	MBE	Adj.R^2	RMSE	MAE	MBE
10:40 - Day	195	**0.88**	**1.50**	**1.15**	**−0.01**	0.70	2.37	1.85	−0.22	0.87	1.54	1.18	0.07
21:50 - Night	195	**0.87**	**1.45**	**1.15**	**0.05**	0.83	1.93	1.54	−0.89	0.83	1.75	1.39	0.12
02:00 - Night	205	**0.82**	**1.56**	**1.23**	**0.01**	0.82	1.57	1.22	0.00	0.73	1.93	1.52	0.08
13:10 - Day	197	**0.86**	**1.77**	**1.36**	**0.09**	0.64	2.87	2.28	−0.16	0.86	1.79	1.37	0.09

and DEM together, LST only, and DEM only. The 2-fold CV involves randomly splitting the testing dataset into two equal groups, with 50 % of the points used for training the model and the remaining 50 % for validation. The average error of the two tests is computed to evaluate model performance. Key performance metrics, such as the CV adjusted coefficient of determination (Adj.R2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Bias Error (MBE), are given to assess the models' accuracy and predictive capacity.

The model's predictive capacity experiences a substantial decline when considering only LST, particularly during daytime. This is evident from the increased errors and a significant decrease in the CV-Adj.R^2. Notably, at 13:10, the hottest time, the performance of all models is reduced. In contrast, the nighttime model exhibits higher performance due to the strong correlation between LST and T_a at night. Conversely, when incorporating DEM as a predictor, the situation is reversed. Despite elevation being widely acknowledged as a significant variable for T_a in previous studies [15,25,26,40], its effect on T_a is more pronounced during daytime and less influential at night. The poorest performance is observed at 02:00, the coldest hour.

Incorporating LST and DEM as predictors enhances temperature estimation accuracy, with daytime models influenced by elevation and nighttime models benefiting from the LST-temperature correlation. Time-specific modeling approaches are crucial for improved performance in temperature estimation. Fig. 4 shows, as an example, the result of the estimated instantaneous NSAT (4a) and the land surface temperature (4b) for the 31st of July 2020 at 10:40, along with a transect E-E[1] over the metropolitan area of Milan (Fig. 4a).

3.2. Land surface temperature and near-surface air temperature

Incorporating both near-surface air temperature and land surface temperature is essential in conducting comprehensive heat island studies, as they provide valuable insights into the phenomenon at various times of the day. Considering a 90-kilometer transect E-E[1], as depicted in Fig. 4, which traverses the city of Milan, it is observed that the profiles of NSAT and LST converge as temperatures decrease during the nighttime hours (Fig. 5). Additionally, land use/land cover characteristics influence the trends, with urbanized areas exhibiting higher LST and NSAT. Notably, while NSAT demonstrates a flattened pattern during the day, increased variability is observed for both temperatures during nighttime.

At 02:00, when temperatures are at their lowest, the profiles of the temperatures exhibit proximity, resulting in an almost perfect overlap in highly urbanized areas like the city of Milan. Consequently, LST can be a reliable proxy for urban heat island studies, specifically during nighttime. Notably, during nighttime, there is a tendency for the difference between LST and NSAT to be primarily negative, indicating that NSAT is slightly higher than LST. A reversal in the trend is observed at night. The inversion is observed at 21:50 and 02:00, indicating the differential thermic inertia between the land surface and air, with air temperature exhibiting a slower cooling rate.

The situation dramatically changes during the hottest hours, when the NSAT and LST curves diverge significantly, displaying distinct shapes. Analysis of the transect demonstrates reduced spatial variability in NSAT during the day. As temperatures increase, the NSAT curve becomes flatter. At 13:00, NSAT becomes dramatically flattened, indicating a more uniform air temperature distribution across the landscape.

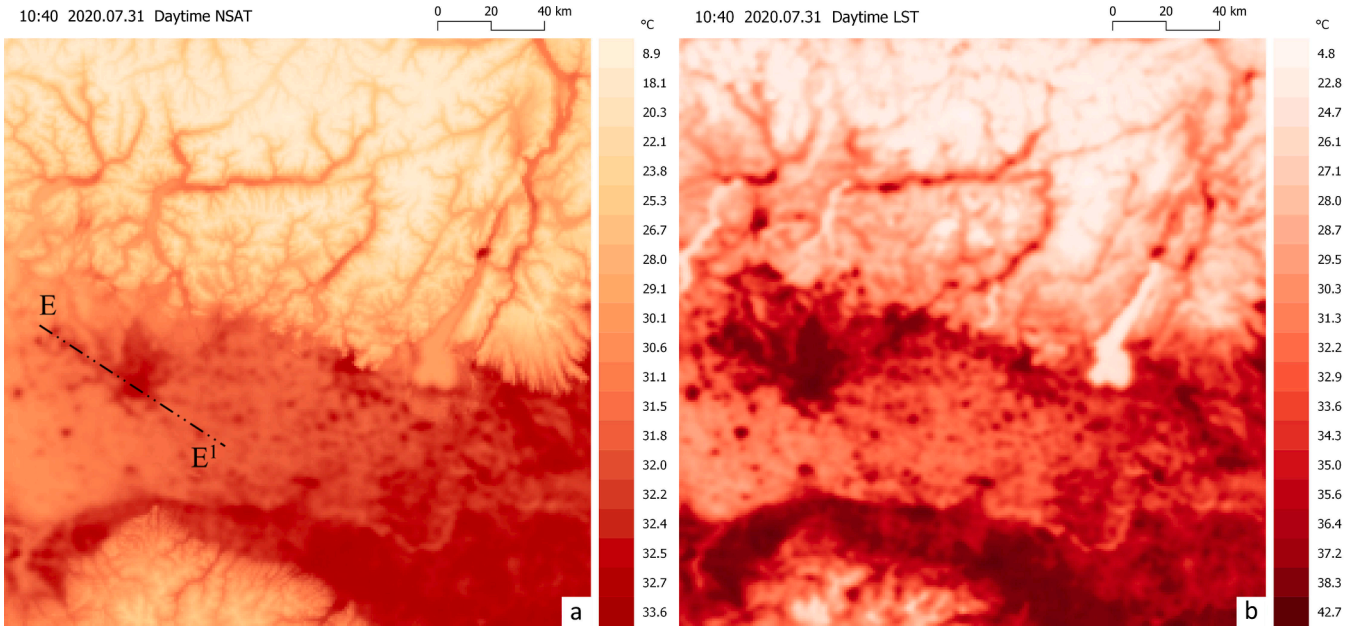


Fig. 4. a). Instantaneous nsat and transect e-e[1] over the metropolitan area of Milan, and b). Land Surface Temperature, for the 31st of July 2020 at 10:40.

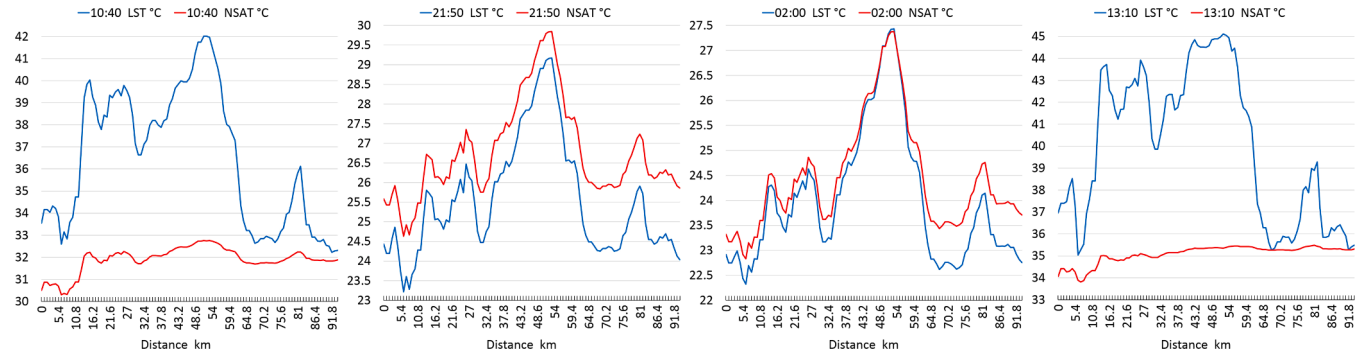


Fig. 5. Spatial behavior of NSAT and LST, along the transect E-E[1], for the different time observations.

In the daytime, LST can reach temperatures up to 10 degrees higher than the air temperature. Furthermore, LST exhibits higher values and significant variability across the territory during the hottest hours. The curves exhibit a slight convergence in non-urbanized areas, wherein different land uses and covers directly influence the variability of LST.

To account for the impact of different land use/land cover on the LST-NSAT interaction, three main LULC categories are considered: anthropized (urbanized), agricultural, and wooded and semi-natural environments. In the Lombardy region, the anthropized area is 15 % of the territory, the agricultural area covers 42 %, and the wooded and semi-natural areas cover 40 %. Around 4 % is covered by wetlands and water bodies. Including these specific categories is relevant for climate-resilient planning, as they hold different implications and significance. Fig. 6 analyzes the correlation between LST and NSAT for the entire study area and the different LULCs.

Although there is a highly positive correlation between NSAT and LST, the correlation patterns during extreme heat reveal a temperature-dependent relationship, with higher temperatures exhibiting reduced correlation. The lowest correlation is found at 13:10, with an R^2 value of approximately 78.2 %. During the morning, at 10:40, characterized by lower maximum temperatures, the correlation is relatively higher, reaching around 80.5 %. Conversely, at night, as temperatures decrease, NSAT and LST exhibit a strong positive correlation, approaching R^2 values of approximately 98 % at 21:50 and 99.6 % at 02:00.

When focused on anthropized and agricultural areas, the distinction between daytime and nighttime correlations becomes more evident. During the night, near-surface air and land surface temperatures exhibit a consistently high and relatively unchanged correlation, with R^2 values of approximately 94 % and 98 %. Conversely, the correlation is significantly reduced during the day. Notably, during the daytime, when temperatures are approximately below 30 degrees, a robust linear correlation is still evident. However, the increase in LST outpaces NSAT at higher temperatures, suggesting a saturation effect of the latter. This phenomenon underscores urban materials' substantial heat storage capacity during the daytime, making the ground considerably warmer than the air at higher temperatures.

Although agricultural areas generally show a slightly lower correlation, their behavior is similar to anthropized areas. Barren or sparsely vegetated lands can absorb solar radiation significantly, particularly during daylight hours.

Wooded and semi-natural environments exhibit distinct behavior. The correlation between LST and NSAT is notably strong both during the daytime, with an R^2 of approximately 80 %, and at nighttime, with an R^2 exceeding 99 %. The high linear correlation can be attributed to dense vegetation, which enhances the evapotranspiration effect and, more importantly, provides ample shading over the ground surface. For the wooded areas, if we consider the whole territory, very high elevations are reached in the mountains, where temperatures dramatically drop during the daytime.

3.3. The effect of vegetation on LST and NSAT

The impact of vegetation on heat islands was examined based on the NDVI. Although the relationship between temperatures and vegetation is widely acknowledged, diurnal variations (day/night) and the spatial heterogeneity of the landscape must be considered. A minimal correlation is observed between NDVI, LST, and NSAT across the entire region, as shown in Figs. 7 and 8. The points cloud distribution reveals a dual trend for the whole region instead. A positive correlation is evident at lower temperatures, while the expected negative correlation is observed as temperatures increase. These patterns persist throughout all observed periods and are consistent for NSAT and LST.

On the other hand, the analysis reveals an enhanced correlation between vegetation and daytime land surface temperature in anthropized and agricultural areas (Fig. 7). A clear negative trend is observed, indicating that higher levels of vegetation are associated with lower surface temperatures. This suggests that vegetation has a more significant and direct cooling effect on surfaces than on the air. It emphasizes the importance of green measures, particularly in man-made environments, as vegetation can effectively mitigate the intensity of heat islands by reducing the heat-storing capacity of urban materials during the day. Resembling the nighttime scenario observed in agricultural areas, the absence of correlation or a discernible linear trend is observed in woodland and semi-natural environments, where natural and densely vegetated surfaces inhibit the attainment of excessively high temperatures.

3.4. The impact of elevation on NSAT and LST

Although, generally, as elevation increases, the intensity of the heat island diminishes, the elevation at which heat islands arise can vary due to local climate conditions, air mass mixing, topography, and urban environment characteristics. As a general approximation, the observable effects of heat islands are typically concentrated within a few hundred meters above sea level. This study employs a threshold of 500 m to investigate the correlation between temperatures and vegetation across different elevations.

Table 3 shows the NDVI-LST and NDVI-NSAT correlations at various elevations for the three LULCs. The minimum elevation considered in this analysis is 200 m, as highly urbanized areas would be excluded at lower elevations. The analysis is given for the entire region and different elevation ranges, namely, below 200 m, between 200 and 500 m, below 500 m, and above 500 m. The table also provides the corresponding percentage of area covered for each extent. Approximately 95 % of anthropized and 94 % of agricultural areas fall below the 500-meter elevation threshold, while 80 % of the woodlands and semi-natural environments are over 500 m.

A high correlation is shown between NDVI and daytime LST for anthropized and agricultural areas. However, the correlation decreases when considering only the portion of the territory between 200 and 500

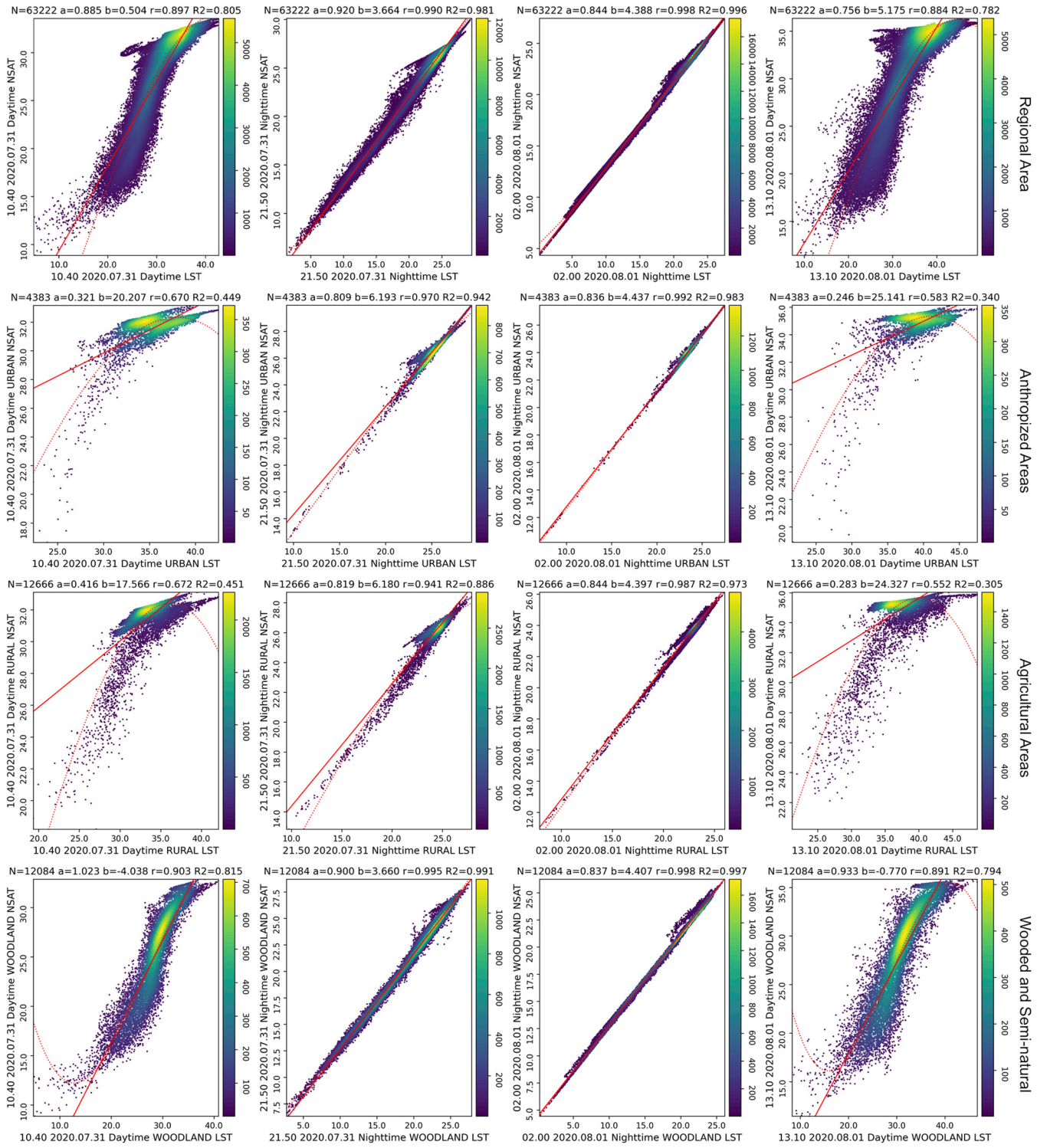


Fig. 6. Scatterplots, regression coefficients, *r,* and R-squared values for LST-NSAT correlation for each time observation in the whole region and the main LULCs: anthropized, agricultural, and wooded/semi-natural.

m. Although anthropized and agricultural areas exhibit similar trends, it is noteworthy that a stronger correlation between NDVI and LST is observed in agricultural areas. Enhancing vegetation health and density can benefit both classes, with a more pronounced impact on agricultural lands regarding LST. Some relevant correlation with daytime LST, of about −77 %, is also found for woodlands under 200 m, where the effect of elevation has less impact on the temperatures.

Although the correlation between NDVI and NSAT is generally weaker, it is worth noting that for areas below 500 m elevation, the correlation between NDVI and NSAT at night becomes more pronounced for anthropized areas, reaching values around −62, −68, or −69 %. This suggests that vegetation health and density can also contribute to cooling effects over the air temperature during nighttime hours in densely urbanized environments.

Overall, a negative correlation is observed in all cases below 500 m, indicating that higher vegetation health and density correspond to lower

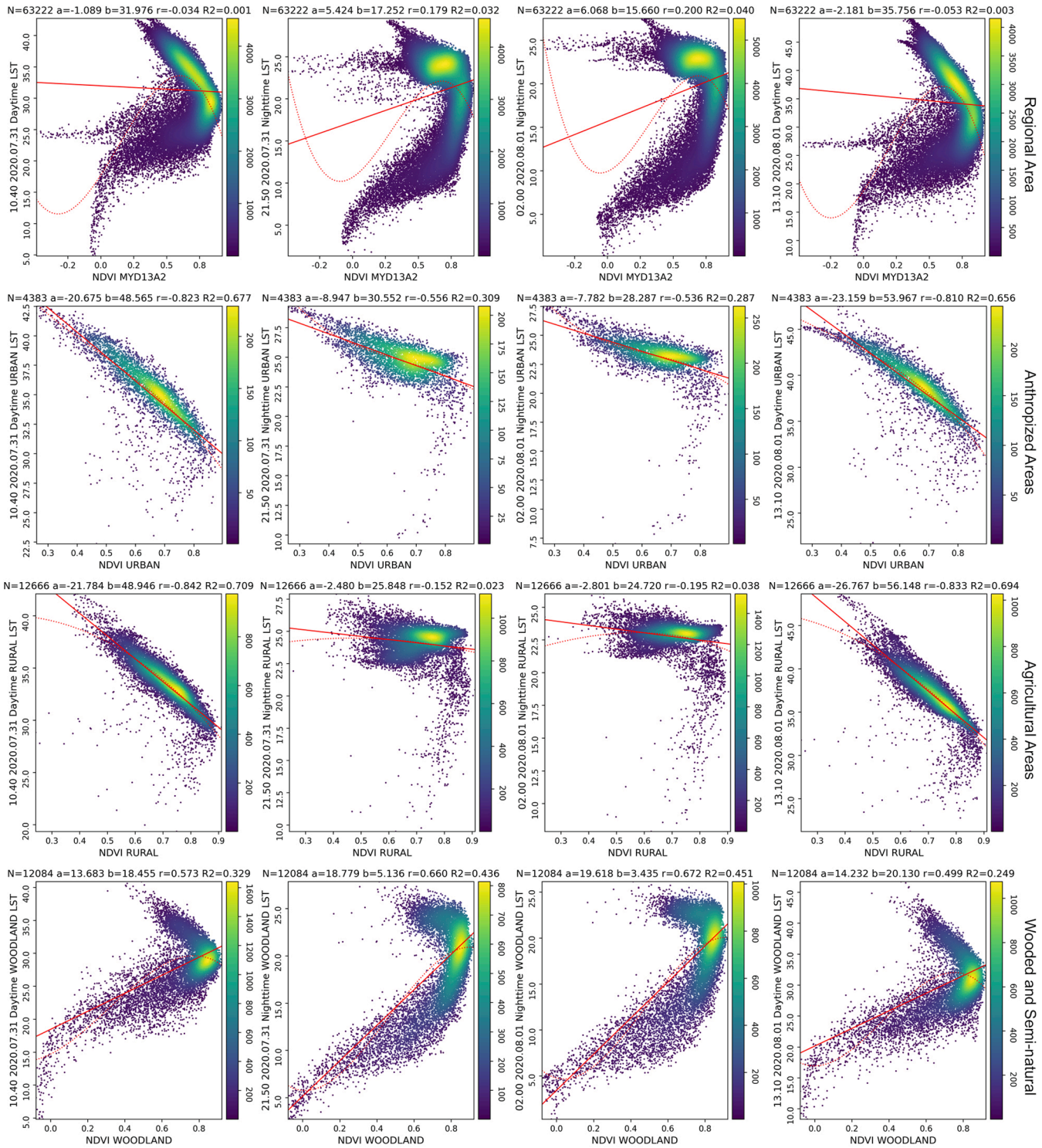


Fig. 7. Scatterplots, regression coefficients, *r,* and R-squared values for NDVI-LST correlation in the region and main LULCs: anthropized, agricultural, and wooded/semi-natural.

temperatures. This pattern, however, is not evident for woodlands over 500 m, where the correlation becomes positive. Notably, while the correlation over 500 m is not significant for anthropized and agricultural areas, it is relevant for woodlands. As previously mentioned, at high elevations, vegetation is not a key driver influencing temperatures, suggesting the presence of other influential factors that should be considered.

3.5. *About the intensity of the regional heat islands*

RHI intensity is examined through the main factors that cause its changes. As reported in Table 4, the surface RHI (SRHI) and canopy layer RHI (CLRHI) are studied and compared based on three elements: the intensity of RHI within different LULCs, heat islands determined by average temperature differences between LULCs, and the intensity of heat islands within LULCs based on changes in the NDVI. The analysis focuses on the region below 500 m.

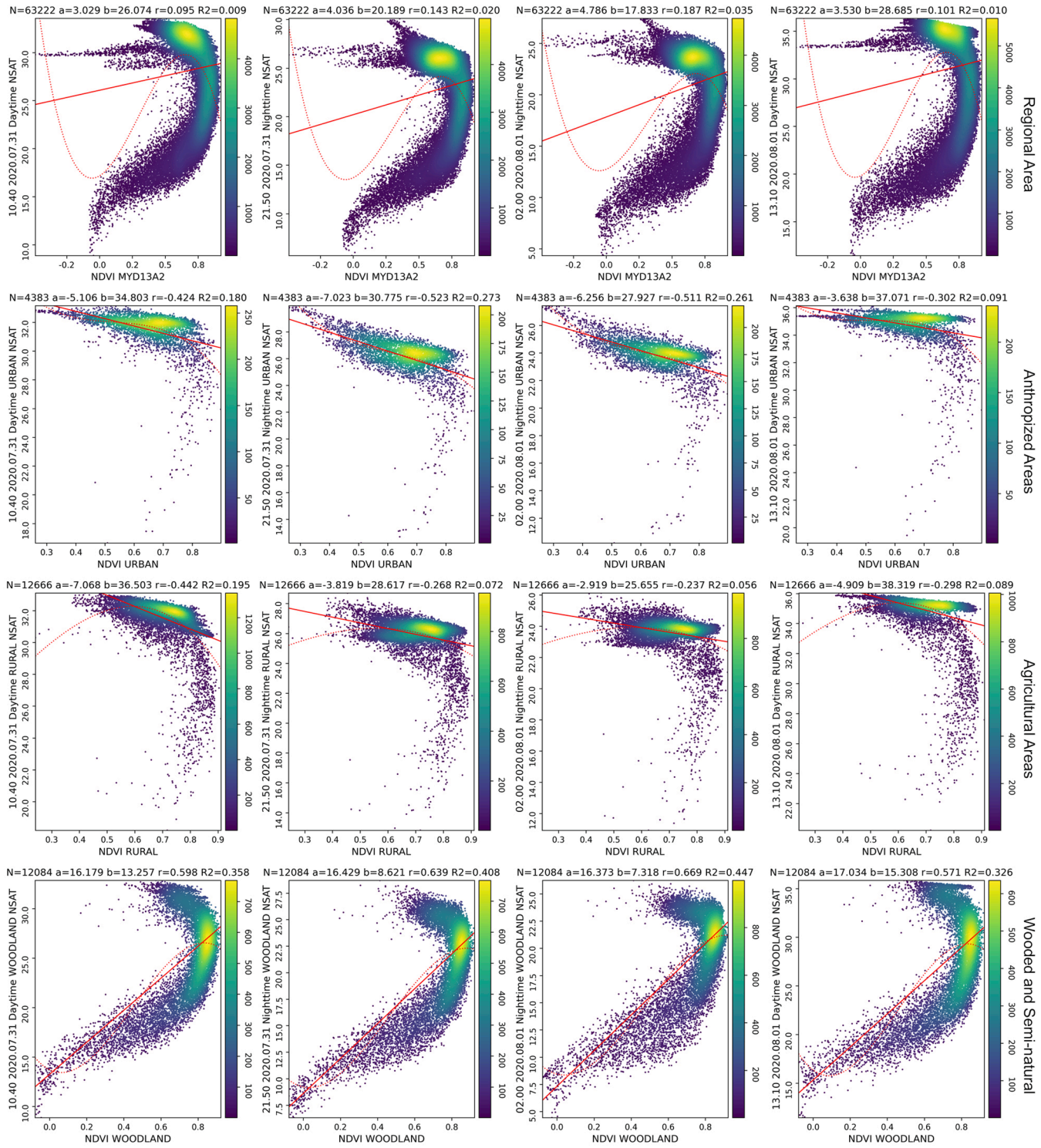


Fig. 8. Scatterplots, regression coefficients, *r,* and R-squared values for NDVI-NSAT correlation in the region and main LULCs: anthropized, agricultural, and wooded/semi-natural.

To calculate the relative RHI intensity within land use/land covers, the average temperatures of each LULC are compared to the minimum temperature value observed in the area under investigation. To prevent excessively low (outliers) values resulting from potentially biased sensor measurements, the 2nd percentile of temperatures is used to determine the minimum value.

Regarding the RHI intensity, the analysis reveals relevant variations across different LULCs. Anthropized areas exhibit the highest intensity, reaching maximum SRHI values of 7.0 at 13:10 (daytime) and CLRHI values of 2.4 at 02:00 (night). This suggests that urbanized regions experience more pronounced heat islands than other land cover types. Agricultural areas exhibit slightly lower values, with SRHI of 5.2 at 13:00 and CLRHI of 2.0 to 2.1 at each observation. Such values are alarming. It is reported that relative RHI intensities of LST ranging from 2 to 8 are recognized to be high-risk values [11,45]. Is it worth noting that agricultural areas show slightly higher CLRHI intensity than anthropized at the hottest hour, i.e., 13:00. Woodlands show lower values, with SRHI ranging from 2.3 at 02:00 to 4.6 at 13:00, and CLRHI

Table 3
Correlation of NDVI with LST and NSAT for the different LULCs, and at different elevations, for each time observation.

Pearson coefficient (*r*)				10:40 - Day		21:50 - Night		02:00 - Night		13:10 - Day	
	Elevation (m)	Area (%)		LST	NSAT	LST	NSAT	LST	NSAT	LST	NSAT
Anthropized Areas		100	NDVI	**-0.77**	-0.42	-0.54	-0.52	-0.52	-0.50	**-0.76**	-0.32
	< 200	62		**-0.86**	-0.56	-0.62	-0.69	-0.69	-0.68	**-0.85**	-0.17
	200-500	33		-0.67	-0.68	-0.62	-0.60	-0.59	-0.57	-0.67	-0.63
	< 500	95		**-0.79**	-0.51	-0.62	-0.62	-0.65	-0.62	**-0.79**	-0.31
	> 500	5		0.44	0.53	0.55	0.55	0.57	0.57	0.37	0.50
Agricultural Areas		100	NDVI	**-0.84**	-0.44	-0.15	-0.27	-0.19	-0.23	**-0.83**	-0.30
	< 200	83		**-0.90**	**-0.71**	0.08	-0.19	0.03	-0.09	**-0.89**	-0.58
	200-500	11		-0.69	-0.59	-0.44	-0.46	-0.48	-0.48	-0.69	-0.46
	< 500	94		**-0.88**	-0.55	0.02	-0.17	-0.04	-0.14	**-0.87**	-0.23
	> 500	6		-0.05	0.08	0.28	0.23	0.24	0.21	-0.16	0.06
Woodlands, Semi-natural		100	NDVI	0.57	0.59	0.66	0.64	0.67	0.66	0.50	0.57
	< 200	7		**-0.77**	-0.58	-0.19	-0.46	-0.44	-0.53	**-0.77**	-0.18
	200-500	14		-0.40	-0.49	-0.23	-0.31	-0.33	-0.35	-0.42	-0.51
	< 500	20		-0.53	-0.48	-0.26	-0.38	-0.37	-0.41	-0.55	-0.39
	> 500	80		**0.80**	**0.81**	**0.82**	**0.82**	**0.82**	**0.82**	**0.76**	**0.78**

Table 4
Comparative analysis of regional heat islands (RHI) intensity across different land use/land covers at different times, with temperature-based heat island calculation and NDVI-based heat island intensity variation.

Regional Heat Island	Elevation < 500 *m*	10:40 - Day		21:50 - Night		02:00 - Night		13:10 - Day	
		SRHI	CLRHI	SRHI	CLRHI	SRHI	CLRHI	SRHI	CLRHI
RHI Intensity	Anthropized	6.2	2.1	3.3	2.2	2.9	2.4	7.0	1.8
	Agricultural	4.6	2.1	2.7	2.0	2.4	2.1	5.2	2.0
	Woodlands	4.1	1.2	2.4	1.4	2.3	1.8	4.6	1.2
LULC-difference-based Intensity	Anthropized – Agricultural	1.7	−0.1	0.6	0.3	0.5	0.3	1.8	−0.2
	Anthropized – Woodlands	2.2	0.9	0.9	0.8	0.6	0.5	2.4	0.7
	Agricultural – Woodlands	0.5	0.9	0.3	0.8	0.1	0.3	0.6	0.9
NDVI-difference-based Intensity	Anthropized	4.0	0.6	1.9	1.4	1.6	1.3	4.2	0.3
	Agricultural	4.7	0.6	1.0	0.8	0.6	0.5	5.7	0.3
	Woodlands	1.7	0.5	0.8	0.9	0.9	0.8	1.0	0.3

from 1.2 to 1.8 being always below 2 degrees. Above all, for the CLRHI, there is a critical jump between the values of anthropized and agricultural areas compared to woodlands during the day.

The differences between average LULC temperatures also emphasize that anthropized areas exhibit significantly higher temperatures than woodland areas. While the difference between anthropized and agricultural areas reaches 1.7 and 1.8 for SRHI at night, the difference with woodlands reaches 2.2 and 2.4. However, it is worth pointing out that in terms of CLRHI, the difference between anthropized and agricultural areas is very low (roughly around zero), even showing an inversion, with negative values, during daytime. The anthropized and agricultural areas in the daytime reach similar intensity values regarding air heat islands. In contrast, the daytime difference between agricultural and woodlands is almost 1 degree.

Regarding the analysis of differences in average values among LULCs, it's crucial to emphasize that the average values, particularly for land surface temperature, can exhibit significant variations, particularly in agricultural areas. These variations are influenced by the variability of land cover types, including bare lands, densely vegetated areas, and sparsely vegetated areas within agricultural regions.

Analyzing heat island intensity within different LULC categories based on variations in NDVI yields further insights. NDVI values range from −1 to 1, where lower values indicate areas with less vegetation cover, while values around 0.4 represent healthy and dense vegetation. In this study, a threshold of 0.5 is applied. NDVI values below 0.5 indicate areas with lower vegetation presence, while values above 0.5 represent areas with higher vegetation density.

Anthropized areas with lower NDVI values display significant heat island intensity. The temperature difference between areas with low NDVI and high NDVI within anthropized areas ranges from 1.6 to 4.2 for SRHI and from 0.3 to 1.4 for CLRHI. Woodlands reach maximum values of 1.7 for the surface heat island and 0.9 for the CLRHI. This indicates that areas with lower vegetation cover experience a more pronounced heat island effect, thus emphasizing the substantial impact of human manipulation on the local climate. In the case of agricultural areas, and mainly for the SRHI, NDVI-based intensity is elevated. The daytime SRHI for agricultural areas ranges from 4.7 to 5.7, higher than the intensity found for the anthropized or woodland areas. This occurrence reports the critical presence of surface heat islands in agricultural regions. While increasing vegetation health and density by 0.5 NDVI can be highly beneficial.

4. Discussion

4.1. Unveiling RHI dynamics

4.1.1. On the intricate correlation between surface and air temperatures

The impact of different predictors on estimating near-surface temperature on a regional scale is first discussed. Through rigorous analysis, three predictor combinations have been examined: LST and DEM together, LST only, and DEM only. The results provide a clear picture of the influence of these predictors on air temperature estimation accuracy and support the hypothesis that incorporating LST and DEM would enhance accuracy. Daytime models are particularly sensitive to elevation, while nighttime models benefit from the strong correlation between LST and air temperature at night.

Once both land and near-surface temperatures are available, a pivotal question arises regarding how temperature varies across different times of the day based on the two temperatures. Indeed, although the difference between urban and rural temperatures is more prominent at night, causing nighttime heat island intensity to be higher, observed temperature values can reach concerning levels during the

day. When addressing heat illnesses, the focus should not be limited to nighttime heat islands alone. The findings support the concern of relying on both temperatures and considering different hours. Distinct temperature patterns during different times are revealed, particularly the divergence between LST and NSAT during daytime hours. According to the results, and in line with other studies [1,15,29,37], it is worth noting that while the thermal profiles of NSAT and LST tend to align at night, they significantly diverge during the day. During the hottest hours, LST could reach temperatures up to 10 degrees higher than NSAT. In summary, higher temperatures reduce the correlation, while lower temperatures produce a stronger positive correlation. Consequently, using LST as a proxy to evaluate the daytime heat island effect is inappropriate.

4.1.2. The significance of vegetation and LULC on temperature variations

Further, this research delves into the impact of land use/land cover and vegetation on temperature variations. The significance of vegetation in climate studies is undeniable. However, vegetation must be carefully considered within specific geographical and temporal contexts. This study reports significant variations in the correlation between vegetation and land surface temperature and near-surface air temperature under various temperature conditions.

In certain areas, such as densely urban regions, low-lying areas, or regions with limited land use/land cover variability, a notable linear (negative) correlation is often observed, particularly during high daytime temperatures [7,18,36]. Nevertheless, in the context of this research, while developing the model to estimate near-surface air temperature, a weaker correlation between normalized difference vegetation index and air temperature was found. This deviation from established patterns in previous research can be attributed to the unique conditions of the study area, including significant landscape heterogeneity, distinct geographical features, and specific land use/land cover patterns.

Regarding the mitigating effect of vegetation on temperatures, this study aligns with existing knowledge. By considering both land surface and air temperature and assessing various land use/land covers, the findings corroborate the existing understanding that emphasizes the more effective cooling effect of vegetation under high-temperature conditions [30]. Additionally, it underscores that vegetation's cooling impact is more pronounced during daytime than nighttime [7,36]. However, this primarily holds true for thermal comfort indices or land surface temperature.

When examining the impact of vegetation on air temperature within urbanized areas, as demonstrated in Table 4 (NDVI-difference-based intensity), vegetation's influence on air temperature is significantly high at night, as previously noted [31], when temperatures are lower.

Finally, the impact of vegetation on both LST and NSAT for different land use/land cover types was assessed, specifically anthropized (urbanized), agricultural, and wooded/semi-natural environments. The results highlight the importance of considering various LULCs when addressing heat islands. Temperatures are particularly sensitive in anthropized and agricultural areas, where the cooling effect of vegetation in mitigating heat islands is more significant than in wooded and semi-natural areas.

In anthropized and agricultural regions, daytime LST is influenced by heat storage in urban materials or barren and sparsely vegetated lands, causing the ground to be warmer than the air and leading to higher daytime surface regional heat island intensity for both anthropized and agricultural regions. Although evapotranspiration subsequently reduces the effect of heat islands in agricultural areas, mostly at nighttime, resulting in a divergent trend and a cooling effect, the patterns of Surface RHI and air RHI (CLRHI) are similar. During the daytime, the intensity of SRHI is approximately double that of CLRHI, as shown in Table 4. In any case, the air RHI intensity remains around 2 degrees, which is identified as a critical value [45]. In contrast, wooded and semi-natural areas exhibit lower RHI intensity due to dense vegetation and shading. As highlighted by Yu et al. [45] and supported by this research, wooded and semi-natural regions demonstrate exceptional potential for heat mitigation, surpassing that of grassland.

4.1.3. Why should elevation be considered?

Elevation consistently emerges as a pivotal factor impacting temperature patterns throughout the regional heat island [4,20]. Accordingly, this study deals with the distinctive temperature dynamics that arise over a regional extent at different elevations. Spatial and temporal interactions between near-surface air and land surface temperatures are addressed, revealing how elevation uniquely shapes these interactions across heterogeneous geographic areas. The findings underscore the intricate interplay of elevation, land use/land cover, and vegetation in molding temperature patterns.

The influence of elevation is particularly noticeable in heat island effects, which tend to concentrate within a few hundred meters above sea level. While a strong correlation was found between vegetation and daytime land surface temperature in anthropized and agricultural regions, this correlation weakens when focusing on elevations between 200 and 500 m. Below 500 m, increased vegetation is associated with cooler temperatures, but this pattern reverses for woodlands situated above 500 m, where the temperature correlation becomes positive.

In summary, although vegetation plays a crucial role, especially in anthropized and agricultural areas, other influential factors become more significant at higher elevations. This underscores the need for a comprehensive understanding of local conditions to accurately model temperature variations on a regional scale.

4.2. Spatial planning and policy implications and recommendations

This work provides valuable insights for policymakers, urban planners, and environmental practitioners. First and foremost, it emphasizes the importance of investigating the urban heat island phenomenon on multiple scales, particularly in metropolitan systems that extend beyond city boundaries. The research elucidates how urban heat islands can extend their influence into agricultural zones, underscoring the interconnectedness between urban and rural climates. This highlights the necessity for holistic climate studies not only focusing on the urban environment.

Also, the results highlight the pivotal role of land use/land cover in shaping temperature variations, emphasizing its significance in climate-proof planning.

Vegetation emerges as a compelling tool for mitigating the regional heat islands, being effective not only in densely populated but also in agricultural areas. Indeed, the correlation patterns found between vegetation (NDVI) and temperature support the pivotal role of greenery in cooling urban and rural environments. It is worth noting that agricultural areas, although less urbanized than heavily developed regions, exhibit discernible impacts on the local climate due to human activities and modifications associated with farming activities. The conversion of natural vegetation to agricultural land and the removal of tree cover can diminish the cooling effect of evapotranspiration, resulting in higher temperatures and an intensified heat island effect. As the cooling effect of woodlands is higher than grass or sparsely vegetated areas, providing the most significant potential for mitigating heat islands in areas characterized by high levels of human activities, it is essential to safeguard and enhance wooded and natural environments.

To address these implications and engage in sustainable spatial planning, some actions could be considered, such as:

- Advocate for data-driven decision-making and investment in data infrastructure and analytical capabilities to effectively address heat islands phenomenon.
- Shaping zoning regulations to optimize the equilibrium between urban development and the preservation of natural green spaces.

- Encouraging the establishment and maintenance of green infrastructures, including urban afforestation, green corridors, parks, and community gardens, within urban and *peri*-urban regions to mitigate daytime temperature extremes.
- Promoting agroforestry and sustainable agricultural practices, incorporating tree planting and other vegetation strategies in agricultural landscapes to moderate temperature variations and enhance resilience. This involves advocating sustainable farming practices such as shade trees, cover crops, short rotation forestry, and precision irrigation systems to mitigate temperature impacts on crop production.

4.3. Limitations and future research directions

Some limitations must be pointed out, delineating potential directions for future research. First, the spatial resolution, reliant on around 1-kilometer pixel raster data, may result in a loss of spatial heterogeneity. Future research could prioritize enhancing the spatial resolution of temperature modeling.

Similarly, the climate dynamics within broadly generalized land uses may still need to be fully represented. For instance, the generalization obscures variations within anthropized regions, such as industrial settlements or large urban parks, as well as within agricultural areas, which can yield more nuanced temperature behaviors depending on different vegetation covers. Hence, future research could delve into specific land uses, exploring internal factors contributing to temperature variations, such as vegetation density, building materials, and local topography.

The analysis incorporates land surface and air temperatures and explores correlation patterns between them and LULC, vegetation, and elevation. However, other factors like humidity remain unaccounted for, which could enhance our understanding of the phenomenon. Furthermore, while the NDVI is widely accepted as a proxy for vegetation density and health, it has limitations in differentiating between various vegetation types, such as grass and tree cover.

Additionally, it is worth mentioning that the analysis is limited by its specific period, which may not fully capture climate variations over time. Although the study focuses on heat stress during the hot season, providing a reasonable estimate of the primary concerns for an extreme heat event, the findings could be tested during different heat waves, for instance. Likewise, the research could extend its temporal scope to assess the impacts of climate change on temperature dynamics over multiple years. This long-term perspective would offer valuable insights into the evolving nature of heat islands.

Lastly, the findings are specific to the study region and may not directly apply to other geographic areas with different climate conditions, land cover characteristics, or urbanization patterns. Comparative studies across other regions will be planned. This would enable researchers to identify common trends and unique regional factors influencing the RHI dynamics. Currently, caution should be exercised when extrapolating these insights to other regions.

5. Conclusions

This research employs a data-driven approach leveraging remote sensing and spatial statistics to study how vegetation and land use/land cover affect the heat island phenomenon at the regional scale. It comprehensively analyzes the regional heat island, considering surface and canopy layer heat islands. Thus, the study emphasizes the importance of correlating land surface and air temperature while considering factors like elevation, landscape composition, and configuration.

Based on the correlation analysis between land surface temperature and near-surface air temperature, the analysis first suggests that targeting areas with high daytime temperatures can boost thermal resilience. Increasing vegetation in these areas reduces daytime heat storage and lessens heat island intensity. Moreover, the research highlights the impact of all human-induced land use changes on the climate. It is worth noting that although heavily anthropized areas significantly contribute to the heat islands, remarkably, agriculture also impacts the heat island intensity on a regional scale. While only highly vegetated and natural areas demonstrate potential for heat mitigation.

Hence, from sustainable spatial planning and land management perspective, besides adopting specific vegetation strategies for different land cover types, sustainable practices like agroforestry and green infrastructure should be increasingly promoted, especially in agricultural areas, as they can reduce heat island effects.

Ultimately, this research deepens our understanding of the regional heat island phenomenon and emphasizes the need for sustainable spatial planning to combat heat islands and enhance climate resilience, providing valuable insights for evidence-based policies and informed urban development.

CRediT authorship contribution statement

Nicola Colaninno: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This research is part of the project '*MultiCAST - Multiscale Thermal-related Urban Climate Analysis and Simulation Tool*,' which has received funding from the European Union's Horizon 2020 (H2020) Research and Innovation program under the Marie Skłodowska-Curie Action - Individual Fellowship | Global Fellowship (MSCA-IF-GF), with grant agreement number 101028035. The area under investigation is among the MultiCAST case studies.

References

[1] Alvi U, Suomi J, Käyhkö J. A cost-effective method for producing spatially continuous high-resolution air temperature information in urban environments. Urban Clim 2022;42(September 2021). https://doi.org/10.1016/j.uclim.2022.101123.
[2] Belda M, Holtanová E, Halenka T, Kalvová J. Climate classification revisited: from Köppen to Trewartha. Climate Res 2014;59(1):1–13. https://doi.org/10.3354/cr01204.
[3] Chen F, Liu Y, Liu Q, Qin F. A statistical method based on remote sensing for the estimation of air temperature in China. Int J Climatol 2015;35(8):2131–43. https://doi.org/10.1002/joc.4113.
[4] Chen S, Yu Z, Liu M, Da L, Faiz ul Hassan M. Trends of the contributions of biophysical (climate) and socioeconomic elements to regional heat islands. Scientific Reports 2021;11(1). https://doi.org/10.1038/s41598-021-92271-3.
[5] Chen XL, Zhao HM, Li PX, Yin ZY. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens Environ 2006;104(2):133–46. https://doi.org/10.1016/j.rse.2005.11.016.
[6] Chen Y, Xie M, Chen B, Wang H, Teng Y. Surface regional heat (cool) island effect and its diurnal differences in arid and semiarid resource-based urban agglomerations. Chin Geogr Sci 2023;33(1):131–43. https://doi.org/10.1007/s11769-022-1324-y.
[7] Chun B, Guhathakurta S. Daytime and nighttime urban heat islands statistical models for Atlanta. Environ Plan B: Urban Anal City Sci 2017;44(2):308–27. https://doi.org/10.1177/0265813515624685.
[8] Colaninno N, Morello E. Towards an operational model for estimating day and night instantaneous near-surface air temperature for urban heat island studies: outline and assessment. Urban Clim 2022;46:101320. https://doi.org/10.1016/J.UCLIM.2022.101320.
[9] Cristóbal J, Ninyerola M, Pons X. Modeling air temperature through a combination of remote sensing and GIS data. J Geophys Res 2008;113(D13):D13106. https://doi.org/10.1029/2007JD009318.

[10] Degefu MA, Argaw M, Feyisa GL, Degefa S. Regional and urban heat island studies in megacities: a systematic analysis of research methodology. In: Indoor and Built Environment, 31. SAGE Publications Ltd.; 2022. p. 1775–86. https://doi.org/10.1177/1420326X211061491.
[11] Dutta I, Das A. Exploring the Spatio-temporal pattern of regional heat island (RHI) in an urban agglomeration of secondary cities in Eastern India. Urban Clim 2020; 34. https://doi.org/10.1016/j.uclim.2020.100679.
[12] Elmes A, Healy M, Geron N, Andrews MM, Rogan J, Martin DG, et al. Mapping spatiotemporal variability of the urban heat island across an urban gradient in Worcester, Massachusetts using in-situ Thermochrons and Landsat-8 Thermal Infrared Sensor (TIRS) data. GIScience Remote Sens 2020;57(7):845–64. https://doi.org/10.1080/15481603.2020.1818950.
[13] Fotheringham AS, Charlton ME, Brunsdon C. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ Plan A 1998;30(11):1905–27. https://doi.org/10.1068/a301905.
[14] Fotheringham A. Stewart, Brunsdon Chris, Charlton, Martin, 2002. Geographically weighted regression : the analysis of spatially varying relationships. 269.
[15] Good EJ, Ghent DJ, Bulgin CE, Remedios JJ. A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series. J Geophys Res Atmos 2017;122(17):9185–210. https://doi.org/10.1002/2017JD026880.
[16] Huang Y, Yuan M, Lu Y. Spatially varying relationships between surface urban heat islands and driving factors across cities in China. Environ Plan B: Urb Anal City Sci 2019;46(2):377–94. https://doi.org/10.1177/2399808317716935.
[17] Karakuş CB. The impact of land use/land cover (LULC) changes on land surface temperature in Sivas city center and its surroundings and assessment of urban heat Island. Asia-Pac J Atmos Sci 2019;55(4):669–84. https://doi.org/10.1007/s13143-019-00109-w.
[18] Kim JP, Guldmann JM. Land-use planning and the urban heat island. Environ Plann B Plann Des 2014;41(6):1077–99. https://doi.org/10.1068/b130091p.
[19] King MD, Platnick S. 1.02 The Earth Observing System (EOS). In: Liang S, editor. Comprehensive Remote Sensing, Vol. 1. Elsevier; 2018. p. 7–26. https://doi.org/10.1016/B978-0-12-409548-9.10312-4.
[20] Lee C. Quantifying effects of spatiotemporal changes of urban and green areas on regional climate change: South Korean cities from the 1980s to the 2010s. Urban For Urban Green 2021;64. https://doi.org/10.1016/j.ufug.2021.127286.
[21] Leung Y, Mei CL, Zhang WX. Statistical tests for spatial nonstationarity based on the geographically weighted regression model. Environ Plan A 2000;32(1):9–32. https://doi.org/10.1068/a3162.
[22] Li L, Zha Y. Estimating monthly average temperature by remote sensing in China. Adv Space Res 2019;63(8):2345–57. https://doi.org/10.1016/j.asr.2018.12.039.
[23] Lougeay R, Brazel A, Hubble M. Monitoring Intraurban temperature patterns and associated land cover in phoenix, Arizona using Landsat thermal data. Geocarto Int 1996;11(4):79–90. https://doi.org/10.1080/10106049609354564.
[24] Lu B, Charlton M, Harris P, Fotheringham AS. Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. Int J Geogr Inf Sci 2014;28(4):660–81. https://doi.org/10.1080/13658816.2013.865739.
[25] Lussana C, Salvati M, Pellegrini U, Uboldi F. Efficient high-resolution 3-D interpolation of meteorological variables for operational use. Adv Sci Res 2009: 105–12. http://www.adv-sci-res.net/3/105/2009/asr-3-105-2009.html.
[26] Mutiibwa D, Strachan S, Albright T. Land surface temperature and surface air temperature in complex terrain. IEEE J Sel Top Appl Earth Obs Remote Sens 2015; 8(10):4762–74. https://doi.org/10.1109/JSTARS.2015.2468594.
[27] Oke TR. The distinction between canopy and boundary layer urban heat islands. Atmos 1976;14(July 2015):268–77. https://doi.org/10.1080/00046973.1976.9648422.
[28] Peel MC, Finlayson BL, McMahon TA. Updated world map of the Köppen-Geiger climate classification. Hydrol Earth Syst Sci 2007;11(5):1633–44. https://doi.org/10.5194/hess-11-1633-2007.
[29] Pepin NC, Maeda EE, Williams R. Use of remotely sensed land surface temperature as a proxy for air temperatures at high elevations: findings from a 5000m elevational transect across Kilimanjaro. J Geophys Res Atmos 2016;121(17): 9998–10015. https://doi.org/10.1038/175238c0.
[30] Perini K, Magliocco A. Effects of vegetation, urban density, building height, and atmospheric conditions on local temperatures and thermal comfort. Urban For Urban Green 2014;13(3):495–506. https://doi.org/10.1016/j.ufug.2014.03.003.
[31] Qiu GY, Zou Z, Li X, Li H, Guo Q, Yan C, et al. Experimental studies on the effects of green space and evapotranspiration on urban heat island in a subtropical megacity in China. Habitat Int 2017;68:30–42. https://doi.org/10.1016/j.habitatint.2017.07.009.
[32] Rasul A, Balzter H, Smith C, Remedios J, Adamu B, Sobrino J, et al. A review on remote sensing of urban heat and cool islands. Land 2017;6(2):38. https://doi.org/10.3390/land6020038.
[33] Roth M, Oke TR, Emery WJ. Satellite-derived urban heat islands from 3 coastal cities and the utilization of such data in urban climatology. Int J Remote Sens 1989;10(11):1699–720. https://doi.org/10.1080/01431168908904002.
[34] Shen H, Jiang Y, Li T, Cheng Q, Zeng C, Zhang L. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens Environ 2020;240(April 2019):111692. https://doi.org/10.1016/j.rse.2020.111692.
[35] Souch C, Grimmond S. Applied climatology: urban climate. Prog Phys Geogr 2006; 30(2):270–9. https://doi.org/10.1191/0309133306 pp484pr.
[36] Sun D, Kafatos M. Note on the NDVI-LST relationship and the use of temperature-related drought indices over North America. Geophys Res Lett 2007;34(24). https://doi.org/10.1029/2007GL031485.
[37] Sun T, Sun R, Chen L. The trend inconsistency between land surface temperature and near surface air temperature in assessing Urban heat island effects. Remote Sens (Basel) 2020;12(8). https://doi.org/10.3390/RS12081271.
[38] Terrádez Mas J, Fraschini F, Rossetto M. Rapporto di sintesi della Strategia Regionale di adattamento ai cambiamenti climatici; 2016.
[39] Tiangco M, Lagmay AMF, Argete J. ASTER-based study of the night-time urban heat island effect in Metro Manila. Int J Remote Sens 2008;29(10):2799–818. https://doi.org/10.1080/01431160701408360.
[40] Uboldi F, Lussana C, Salvati M. Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks. Meteorol Appl 2008;15(3):331–45. https://doi.org/10.1002/met.76.
[41] Voogt JA. Image representations of complete urban surface temperatures. Geocarto Int 2000;15(3):21–32. https://doi.org/10.1080/10106040008542160.
[42] Weng Q, Lu D, Schubring J. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. Remote Sens Environ 2004; 89(4):467–83. https://doi.org/10.1016/j.rse.2003.11.005.
[43] Xing Z, Li ZL, Duan SB, Liu X, Zheng X, Leng P, et al. Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime MODIS instantaneous observations. ISPRS J Photogramm Remote Sens 2021;178(May): 51–67. https://doi.org/10.1016/j.isprsjprs.2021.05.017.
[44] Yoo C, Im J, Park S, Quackenbush LJ. Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data. ISPRS J Photogramm Remote Sens 2018;137:149–62. https://doi.org/10.1016/j.isprsjprs.2018.01.018.
[45] Yu Z, Yao Y, Yang G, Wang X, Vejre H. Spatiotemporal patterns and characteristics of remotely sensed region heat islands during the rapid urbanization (1995–2015) of Southern China. Sci Total Environ 2019;674:242–54. https://doi.org/10.1016/j.scitotenv.2019.04.088.
[46] Yu Z, Yao Y, Yang G, Wang X, Vejre H. Strong contribution of rapid urbanization and urban agglomeration development to regional thermal environment dynamics and evolution. For Ecol Manage 2019;446:214–25. https://doi.org/10.1016/j.foreco.2019.05.046.
[47] Zhang Z, Du Q. Hourly mapping of surface air temperature by blending geostationary datasets from the two-satellite system of GOES-R series. ISPRS J Photogramm Remote Sens 2022;183(June 2021):111–28. https://doi.org/10.1016/j.isprsjprs.2021.10.022.
[48] Zhou X, Chen H. Impact of urbanization-related land use land cover changes and urban morphology changes on the urban heat island phenomenon. Sci Total Environ 2018;635:1467–76. https://doi.org/10.1016/j.scitotenv.2018.04.091.
[49] Zhu W, Lu A, Jia S. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. Remote Sens Environ 2013;130: 62–73. https://doi.org/10.1016/j.rse.2012.10.034.