



**Politecnico
di Torino**

Politecnico di Torino

Planning for the Global Urban Agenda

A.a. 2022/2023

Sessione di Laurea di settembre 2023

Analysing Urban Heat Island Dynamics in Turin:
A Machine Learning-based Investigation of Climate
Changes and Intervention Effects

Relatrice:

Guglielmina Mutani

Candidati:

Alessandro Scalise

Xhoana Sufa

Abstract

This thesis aims to explore the intricate dynamics of urban heat islands (UHIs), a critical concern exacerbated by the combined forces of climate change and urbanization.

The primary objective is to develop an *urban-scale model*, leveraging advanced *machine learning algorithms*, capable of incorporating a wide array of variables. The study also endeavours to probe the *influence of climate change on UHIs* through a meticulous examination of historical climate data, seeking to identify any discernible trends that might be linked to the exacerbation of the phenomenon.

Furthermore, the research sets out to assess the *efficacy of substantial urban regeneration projects* in *mitigating the adverse effects of UHIs*. A key aspect of this evaluation involves a comparative analysis of data sets collected both prior to and after these interventions, focusing specifically on the years 2001 (pre-intervention) and 2018 (post-intervention).

Through these multifaceted approaches, this research seeks to yield valuable insights into UHIs, their interaction with climate change, and the potential of urban revitalization initiatives to mitigate their detrimental impacts, thereby offering valuable guidance for more sustainable urban planning strategies.

Acknowledgements

Alla nostra relatrice.

Desideriamo esprimere la nostra sincera gratitudine alla nostra relatrice, la **Professoressa Guglielmina Mutani**, per il sostegno che ci ha fornito durante questo lavoro di ricerca, per la pazienza e la sua disponibilità nel rispondere alle nostre domande. La sua dedizione all'insegnamento e la sua competenza hanno costantemente supportato il nostro lavoro, al fine di stabilire una solida base teorica e metodologica.

Le siamo grati professoressa per averci guidato in questo percorso e per averci fornito ogni volta "la chiave" per superare le difficoltà che ci si sono presentate.

- *Alessandro Scalise & Xhoana Sufa*

A noi (Alessandro e Xhoana).

A noi che abbiamo affrontato questo viaggio sapendo di poter contare l'uno sull'altro, a noi che, in questi lunghi anni di percorso universitario, abbiamo trovato supporto reciproco e ci siamo avventurati in ogni esperienza con grande entusiasmo e tanta, tanta voglia di fare.

Un ringraziamento alle nostre famiglie.

Desideriamo esprimere la nostra profonda gratitudine alle nostre famiglie, che hanno costantemente sostenuto le nostre scelte, ci hanno accompagnato nei momenti difficili e ci hanno incoraggiato a dare sempre il massimo!

Un sentito ringraziamento a tutti coloro che ci hanno ispirato, insegnato e valorizzato. Non facciamo nomi perché rischieremmo di non citare tutti. Nei momenti di difficoltà sappiate che abbiamo trovato grande conforto nelle vostre parole e rassicurazioni. Molti di voi hanno tolto tempo alle proprie attività per poterlo dedicare a noi.

Speriamo di ritornarvi quello che ci avete dato.

Alessandro SCALISE

Xhoana SUFA

*Climate change is not a mere hypothesis; it's an
undeniable scientific truth, and the very cities we
inhabit stand as prime architects fueling this
planetary crisis.*

Alessandro Scalise & Xhoana Sufa

Table of contents

Abstract.....	1
Acknowledgements.....	2
Table of contents.....	4
1. Literature Review.....	7
1.1 General overview of the analysis	7
1.2 Introduction to Urban Heat Islands	8
1.3 Urban Canopy, Urban Canyon, and Urban Boundary Layer	10
1.4 Consequences of UHIs	12
1.5 Evaluate and mitigate UHIs	15
1.6 Environmental protocols for reducing UHI effect	29
2. Studying urban heat islands: the case of Turin	34
2.1 Main purposes of the UHI analysis in Turin	35
2.2.1 Visualizing Thesis Methodology: Flowchart Overview.....	36
2.2 Understanding Climate Context: A Brief Overview	37
2.3 Transforming the Urban Landscape: Large-Scale Interventions in Turin	53
3. Pre-modelling: Assessment of Environmental Variables	57
3.1 Data collection and processing.....	57
3.2 Candidate variables	61
3.2.1 Land cover variables.....	64
3.2.2 Urban geo-morphology variables	81
3.2.3 Weather stations variables	91

3.2.4	UHI evaluation.....	112
4	Modelling: Variables Influencing Urban Heat Island Effect in Turin	119
4.1	Hyperparameters tuning	119
4.2	Machine Learning algorithm [RF].....	127
4.3	Training & Prediction model of the UHI in the case study	130
5	Urban regeneration & UHI	155
5.1	UHI magnitude before and after the great intervention.....	156
5.2	Potential insights: Combating UHIs through Urban Planning for Climate-Resilient Futures	160
6	Conclusions: Integrated Reading and Interpretation of Results	168
6.1	Assessing the Impact of Climate Change: Analysis of Variables Across Seasons in 2001 and 2018	169
6.2	Planning documents and future directions for UHIs research	181
7	Bibliography.....	184
8	Sitography	190
9	Table of figures	192
APPENDIX.....		194
A.	Python scripts (Variables calculation)	194
B.	Hyper parameters tuning (codes and results)	204
C.	Literature Table	0

1. Literature Review

1.1 General overview of the analysis

Urban heat islands (UHIs) represent a significant and growing environmental challenge, as the world's urban population continues to expand. UHIs refer to the phenomenon of cities, and urban areas, experiencing higher temperatures than their rural surroundings, due to increased heat absorption and retention by materials in the built-up environment.

This report aims to provide a comprehensive literature review of the latest (and most relevant) research on UHIs, examining the negative impacts of UHIs and interventions that can be taken into consideration to mitigate their effects. Through this review, we aim to enhance the comprehension of UHIs and help in detect tools to spot this phenomenon.

Remote sensing technologies, like satellite imagery and aerial photography, can provide an extensive overview of the urban landscape and temperature distribution; ground-based measurements, such as temperature sensors and thermal cameras, can provide more detailed data at specific locations; additionally, computer models can simulate temperature patterns and identify potential UHIs.

The combination of these tools can provide a comprehensive understanding of UHIs and aid in devising strategies to mitigate their effects.

In the following paragraphs, we will explore several methodologies that can be used to study and address the challenges posed by urban heat islands, but first, we need to set a general overview of the phenomenon.

1.2 Introduction to Urban Heat Islands

Luke Howard (1772 - 1864), was an English chemist and meteorologist of the early 19th century¹, he was the first to provide evidence that urban areas impact local climates by revealing higher air temperatures compared to adjacent rural regions (Howard, 1833 [2]; Oke, 1982 [3]). He concentrated his analysis (of the London urban extension) on examining two **interrelated yet distinct subjects: (1) the impact of urbanization on the quality of meteorological data and (2) the factors contributing to the urban effect.**

Howard's analysis stated that *“the temperature of the city is not to be considered as that of the climate”*, in fact, he talks about an *“artificial warmth”* that partakes to its construction, dense population, and high fuel consumption.

Howard (1833) identified, four main causes for these differences in air temperature:

- ***Human activities*** that generate heat contribute to the warming of the atmosphere, especially during the winter season.
- ***The geometry of urban surfaces*** which ‘traps’ radiation and obstructs ‘free radiation to the sky.
- ***The unevenness of urban surfaces*** can hinder the movement of the “gentle” summer winds.
- ***The availability of moisture*** for evaporation in the country.

Some years later, another important scholar has been concentrated on urban climate and micrometeorology studies, T.R. Oke [4] a Canadian geographer, and climatologist who is known for his research on urban climate and micrometeorology.

According to Oke's study (1982), UHIs can be generalised, initially, by the illustration of common features of heat island morphology, describing the *“near-surface”* of the heat

[1] <https://www.scienceandsociety.co.uk/index.asp>

[2] Howard, L. (1833). *The Climate of London*. London, UK: Harvey and Darton.

[3] Oke, T. R. (1982). *The energetic basis of the urban heat island*. *Quarterly Journal of the Royal Meteorological Society*, 108(455), 1-24.

[4] <https://geog.ubc.ca/profile/tim-oke/>

island, moreover, he considers irrelevant the topography of the surroundings and puts the accents to the urban/rural *boundaries* that follows the outline of the built-up area. In fact, from the *Figure 1*, it is noticeable that the city centre displays higher temperature readings in comparison to the surrounding region.

Additionally, it is crucial to bring attention to the fact that, most of the urban area, exhibits a more gradual horizontal thermal gradient, but occasional, warm, or cool *spots* can be found, correlated with specific areas of unusually high or low building density.

It is possible for areas such as parks or lakes to have a cooler temperature, whereas more industrialized areas, apartments, shopping districts, or the downtown core may be warmer.

In the presence of a light wind, these temperature discrepancies can be observed as they move slightly downwind from their source locations.

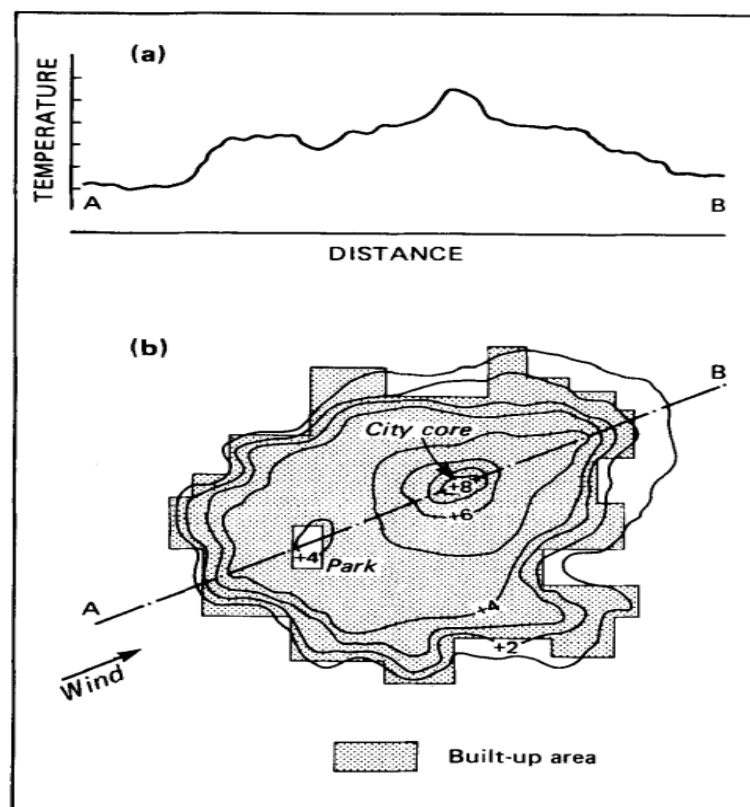


Figure 1 - Representation by Oke T.R. The energetic basis of the UHIs (Oke,1982)

Oke took inspiration, for his representation, by Howard (1833) *Figure 2* where he illustrated the increase magnitude of UHI towards the core of the settlement (of London urban area), in correspondence with greatest building density.

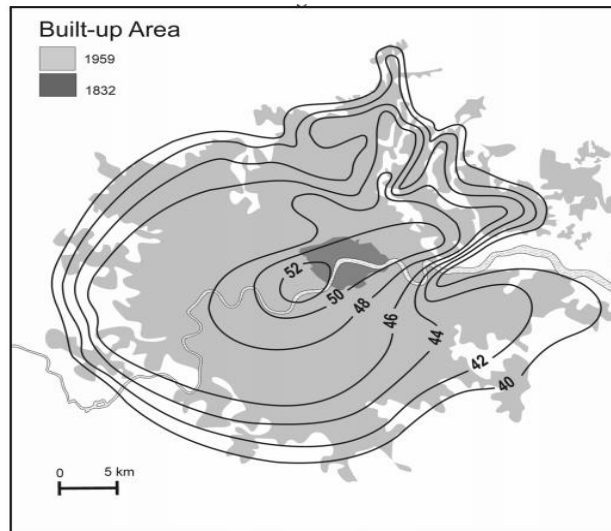


Figure 2 - Howard's examination of the urban effect in London (1833)

1.3 Urban Canopy, Urban Canyon, and Urban Boundary Layer

So far, we've been talking about how urban heat islands (UHIs) can spread out *horizontally*, but it's also important to consider their vertical components. According to Oke, there are two distinct layers of UHIs that can form from urbanization *Figure 3*:

- *the Urban Canopy*
- *the Urban Boundary Layer.*

This distinction is crucial because each layer has unique characteristics that contribute to the overall UHI effect. Oke's research on this topic can be found in his 1976 study [5].

Where he defined the *Urban Canopy Layer* (UCL) *Figure 3*, as the layer that lies between the ground and the average height of buildings and trees. Its features are mainly influenced by the transfer of energy and matter between the different patches of land surfaces and the air in the *urban canyon*. The *urban canyon* *Figure 4* refers to the narrow

[5] Oke T.R. (1976) *The distinction between canopy and boundary-layer urban heat islands*, *Atmosphere*, 14:4, 268-277, DOI: 10.1080/00046973.1976.9648422

and tall structures that create an enclosed space in an urban area, often resembling the shape of a canyon. These structures can be buildings, skyscrapers, and other urban features that line up along the streets, blocking natural light and airflow.

Instead, the **urban boundary layer** *Figure 3* refers to the layer of the atmosphere that is directly influenced by the characteristics and activities of urban areas. It typically extends from the surface up to a height of a few hundred meters and is characterized by intense turbulence and mixing due to the presence of buildings, roads, and other urban infrastructure.

In other words, the urban boundary layer encompasses the entire portion of the atmosphere that is influenced by urban areas, while the urban canopy layer is a subset of that layer that focuses specifically on the interactions between the air and the surfaces of buildings and other structures.

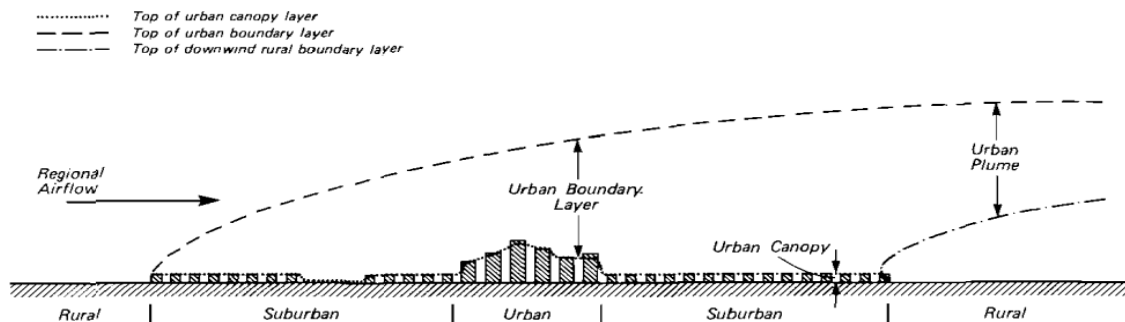


Figure 3 - Schematic representation of the urban atmosphere illustrating a two-layer classification of thermal modification (Illustration by T.R. Oke, 1976)

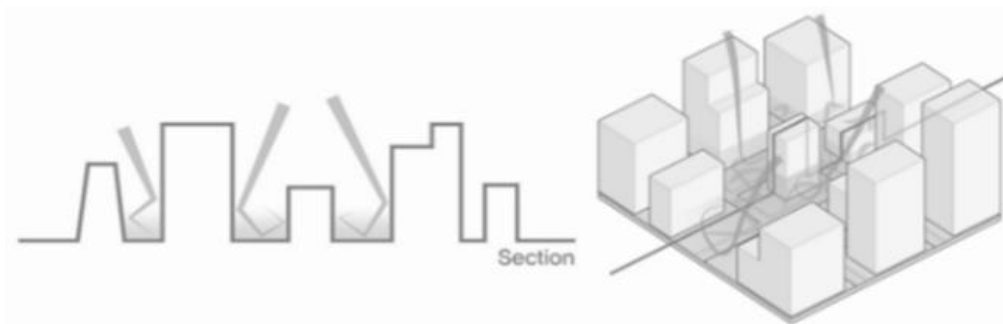


Figure 4 - Urban Canyon (Illustration by Y. Choi, S.Lee, H.Moon, 2018)

1.4 Consequences of UHIs

The causes of UHI are multifaceted, including factors such as urbanization, air pollution, and changes in land use. Numerous studies have shown that ***UHI has significant impacts on the environment and public health***, making it an important area of research. However, the continuous expansion of urban areas has caused a rise in surface temperature, which has a significant impact on the distribution of resources and energy in urban environments. Rising temperatures in urban areas can lead to an increase in energy usage, which places a heavier strain on power resources. Furthermore, the ***urban heat island effect intensifies the concentration of pollutants and worsens atmospheric conditions within cities***. Overall, the UHI effect negatively impacts the various aspects of social-ecological systems in local areas, this change also affects the way urban ecological systems work, which could ultimately harm severely the health of city dwellers (Howard, 1833; Yang et al., 2016 [6]).

Climate change is further exacerbating this problem, leading to more frequent, severe, and longer heat waves. Pollution has been reported to affect health with hard outcomes: research studies have suggested that higher temperatures and pollution in the air could elevate the incidence of infectious ailments such as malaria and long-term diseases including obesity, hypertension, diabetes, asthma, respiratory and cardiovascular disorders (Wang et al., 2023) [7].

Urbanization can be considered a significant driver of environmental problems: the four main aspects introduced by Howard in 1833, and later elaborated by Oke in his studies on UHIs in 1982 (conveyed in this report in the previous paragraph at page 3), led to the assumption that these *four factors* can be modified by urban expansion.

[6] Yang, L., Qian, F., Song, D.-X., & Zheng, K.-J. (2016). Research on Urban Heat-Island Effect. *Procedia Engineering* 169 (2016) 11 – 18. Doi: 10.1016/j.proeng.2016.10.002

[7] Wang S., W. Cai, Y. Tao, Q. Chayn Sun, P. Pui Yun Wong, X.Huang, Y. Liu, (2023). Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models. *Science of The Total Environment*. Volume 871. 162005. ISSN 0048-9697.

Although, in the research chased by *Vujovic et al. (2021)* [8], the **development of urban areas has significant impacts on the environment**; specifically, on the natural hydrology and energy systems, this is due to the replacement of pervious and semi-pervious surfaces with impervious surfaces (such as asphalt and concrete pavements). These modifications directly impact watershed hydrology by (1) reducing infiltration capacity, (2) increasing surface runoff, (3) shortening time of concentration, and (4) decreasing recharge of groundwater. If surface runoff drains faster, then there won't be as much water for evapotranspiration. This affects the way that energy balances out on the surface in urban areas. When there are more impervious surfaces, it causes the land surface temperature to go up, this is because the changes of the local energy balance, by:

- **Albedo**, defined by Britannica (2022) [9] as the fraction of incoming solar energy that is reflected into space from the Earth's surface. It is a measure of the reflective properties of a surface, with values ranging from 0 to 1. Surfaces with high albedo reflect more solar energy and absorb less, whereas those with low albedo absorb more solar energy and reflect less.
- **Specific heat capacity**, defined by Britannica (2022) [10] as the amount of heat energy required to raise the temperature of a unit mass of a substance by one degree Celsius. It is a measure of the ability of a substance to store heat energy.
- **Thermal conductivity**, defined by Britannica (2022) [11] as the property of a material that determines how quickly it can transfer heat energy. It is a measure of the ability of a material to conduct heat.

[8] Vujovic, S., Haddad, B., Karaky, H., Sebaibi, N., & Boutouil, M. (2020). Urban Heat Island: Causes, Consequences, and Mitigation Measures with Emphasis on Reflective and Permeable Pavements. *Sustainability*, 12(23), 9836. doi: 10.3390/su12239836

[9] Britannica, The Editors of Encyclopaedia. "albedo". *Encyclopedia Britannica*, 29 Dec. 2022, <https://www.britannica.com/science/albedo>. Accessed 27 April 2023.

[10] Britannica, T. Editors of Encyclopaedia (2022) specific heat. *Encyclopedia Britannica*. <https://www.britannica.com/science/specific-heat>

[11] Stewart, K. (2023). thermal conductivity. *Encyclopedia Britannica*. <https://www.britannica.com/science/thermal-conductivity>

Xu et al. (2013) [12] conducted a 20-year study revealing a significant positive exponential relationship between impervious surface and land surface temperature. The study highlights the *need for urban planners and civil engineers to consider the environmental impacts of urban development and to implement sustainable practices that mitigate these impacts*, as also retrieved by the study conducted by Vujovic et al. (2021) [13].

[12] Xu, H.; Lin, D.; Tang, F. The impact of impervious surface development on land surface temperature in a subtropical city: Xiamen, China. *Int. J. Climatol.* 2013, 33, 1873–1883.

[13] *Ibid.* [7]

1.5 Evaluate and mitigate UHIs

The scientific literature is consulted with the purpose of comprehending and extracting the methodologies utilized for evaluating the urban heat island phenomenon in diverse case studies. The fact that various approaches were presented in the literature suggests that *there is no one-size-fits-all method for UHI assessment*, and the optimal approach may depend on specific research objectives and spatial scales.

By summarizing the varied approaches presented in the literature, and identifying the most intriguing ones, the analysis may facilitate the development of more effective UHI assessment methods.

Various analysis has been developed to assess the UHI effect, ranging from remote sensing and geospatial analysis to in-situ measurements and numerical models. However, the choice of the most appropriate method is often dependent on factors such as:

- *Data availability*
- *Spatial resolution*
- *Research objectives.*

Therefore, a comprehensive review of the scientific literature on UHI assessment is crucial for selecting the most appropriate model for a specific study.

This paragraph delves into an examination of a variety of research studies and papers intended *to identify the methodologies that have been utilized to assess urban heat island occurrences*. As a general overview of the analysis, the most used methodologies to deal with urban heat islands, in these papers are: 1. *GIS analysis*; 2. *Machine learning algorithms*. Specifically, *GIS* was used to: *a.)* map the spatial distribution of UHI, *b.)* assesses urban morphology, *c.)* identifies ventilation corridors, and *d.)* predict UHI using artificial neural networks. *Machine learning algorithms* were used to: *a.)* predict land use land cover changes and surface UHI phenomena, *b.)* model outdoor thermal comfort, *c.)* simulate and mitigate UHI effects, and *d.)* conduct scenario analysis to study the future of UHI effects.

Some commonly formulas for quantifying UHIs are:

- **Delta T formula** (Scott et al, 2019; Oke, 1987) [14] [15]: This formula calculates the temperature difference between an urban area and a nearby rural area. It is expressed as:

$$\Delta T = T_u - T_r$$

where:

ΔT = temperature difference (in °C)

T_u = temperature in the urban area (in °C)

T_r = temperature in the rural area (in °C)

The Delta T formula can be corrected using coefficients to account for different factors that may affect the urban heat island intensity. Here are some examples of **correction factors** that can be included in the *Delta T* formula:

- Wind speed correction: Wind speed can affect the heat transfer between the urban and rural areas, with higher wind speeds leading to more efficient mixing and lower urban heat island intensity. The wind speed correction coefficient can be calculated based on the ratio of the wind speed at the urban location to the wind speed at the rural location.

$$WS_coeff = (WS_urban - WS_rural) / WS_rural$$

- Surface cover correction: Different land cover types can have different thermal properties and affect the surface energy balance differently. The surface cover correction coefficient can be calculated based on the

[14] Scott, A. A., Waugh, D. W., & Zaitchik, B. F. (2019). Reduced urban heat island intensity under warmer conditions. *Environmental Research Letters*, 14(9), 094011.

[15] Oke, T. R. (1987). *Boundary Layer Climates* (2nd ed.). Methuen.

fractional coverage of impervious surfaces, vegetation, and water bodies in the urban and rural areas.

- Solar radiation correction: Solar radiation can affect the surface temperature and energy balance, with higher solar radiation leading to higher urban heat island intensity. The solar radiation correction coefficient can be calculated based on the incoming solar radiation at the urban and rural locations.
- Time of day correction: The urban heat island intensity can vary throughout the day and night, with the largest differences typically occurring during the late afternoon and early evening. The time-of-day correction coefficient can be calculated based on the time of day when the temperature measurements are taken.
- **Energy Balance Method:** involves estimating the energy balance at the surface of an urban area, which involves accounting for the different heat fluxes (incoming solar radiation, outgoing longwave radiation, sensible heat flux, latent heat flux, and ground heat flux).

The UHI intensity is then calculated as the difference between the surface temperature in the urban area and the surface temperature in a nearby rural area.

The formula used for this method is:

$$\text{UHI} = T_{\text{s_urban}} - T_{\text{s_rural}}$$

where:

UHI = Urban Heat Island intensity

$T_{\text{s_urban}}$ = Surface temperature in the urban area

$T_{\text{s_rural}}$ = Surface temperature in the rural area

The surface temperature (T_s) can be calculated using the energy balance equation, which involves accounting for the different heat fluxes at the surface. The basic formula for surface temperature is:

$$T_s = (R_n - G - H) / \sigma \varepsilon$$

where:

R_n = net radiation at the surface

G = soil heat flux

H = sensible heat flux

σ = Stefan-Boltzmann constant ($5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$)

ε = surface emissivity

Net radiation (R_n) is the difference between the incoming solar radiation and the outgoing longwave radiation. It can be calculated as:

$$R_n = R_{si} * (1 - \alpha) - R_l$$

where:

R_{si} = incoming solar radiation

α = surface albedo (reflectivity)

R_l = outgoing longwave radiation

Soil heat flux (G) is the energy transferred into or out of the soil and can be estimated using soil temperature measurements and heat transfer equations.

Sensible heat flux (H) is the energy transferred from the surface to the atmosphere due to temperature differences and can be estimated using atmospheric measurements and heat transfer equations.

Surface emissivity (ϵ) is a measure of the efficiency with which the surface emits longwave radiation and can be estimated based on the surface material and its temperature.

It's important to note that the energy balance equation is a complex equation that requires accurate measurements of the different heat fluxes and properties such as albedo and emissivity. In practice, there are different methods and models that can be used to estimate surface temperature, depending on the data availability and level of accuracy required.

Additionally, the choice of method may also depend on the specific characteristics of the urban area being studied.

Some commonly used factors for measuring UHIs include:

- **Land surface temperature (LST)** - This is the temperature of the Earth's surface as measured by satellite imagery or thermal infrared imaging. Urban heat islands can be identified by comparing LSTs.
- **Air temperature** - This can be measured using weather stations or mobile sensors located in different urban areas.
- **Emissivity** - This is a measure of how efficiently a surface emits thermal radiation. Urban areas typically have lower emissivity due to their built environment, which can lead to higher temperatures.
- **Vegetation cover** - The amount of vegetation cover in an urban area can impact the urban heat island effect. Studies have shown that areas with a higher percentage of green space tend to have lower temperatures.
- **Wind speed measurements** - Urban areas tend to have lower wind speeds due to the presence of buildings and other structures which create obstacles to the flow of air. This can result in higher temperatures in urban areas due to reduced ventilation.
- **Urban density** - refers to the degree of compactness of urban development within a given area; higher urban density can exacerbate the phenomenon by increasing the number of heat-absorbing surfaces and reducing green spaces, while lower urban density with more green spaces can help mitigate it.

The subsequent section will provide an exposition of the scientific sources we compiled during our inquiry into measurement methodologies for the urban heat island effect. This literature review forms the groundwork for determining the most appropriate techniques for our research.

The first of the research papers is the study conducted by *Y. Zheng et al. (2023)* [16] in Hong Kong, that focuses on the impact of ***urban morphology and geographical conditions on local thermal environments*** and associated heat exposure risks in high-density cities.

The spatial variability of urban forms and geographical conditions has been quantified using the Local Climate Zone (LCZ) framework, which consists of 17 basic classes that standardize the observation and report of UHI (Urban Heat Island) studies. In Hong Kong, the LCZ classification maps have been used to differentiate local thermal variation, and UHI intensity has been found to be maximized in deep night or early morning. The method includes the use of TESTO 480 temperature and humidity probe and data logger for measurement, bivariate and multivariate analysis, and the establishment of quantitative connections between LCZ parameters and UHI intensity.

The study suggests the need for increased urban openness and permeability, and the promotion of radiative and convective cooling through wider setbacks of buildings and the use of green spaces with high pervious surface cover.

In research by *K. Liao et al. (2021)* [17] it is considered the ***impact of spatial heterogeneity in urban morphology*** on surface urban heat islands (SUHIs). It analyses the relationship between SUHIs and urban morphology factors such as building density, height, and materials used for construction.

[16] Sun, T., Gao, J., Yu, Q., & Wang, X. (2021). Mapping the spatial distribution of nocturnal urban heat island based on Local Climate Zone framework. *Sustainable Cities and Society*, 72, 103072. <https://doi.org/10.1016/j.scs.2021.103072>

[17] Liao, K., Hong, Y., & Heo, J. (2018). The effect of spatial heterogeneity in urban morphology on surface urban heat islands. *Environmental Research Letters*, 13(6), 064014. doi: 10.1088/1748-9326/aac2d8

The research found that urban morphology plays a significant role in the formation of SUHIs, with factors such as high building density and low vegetation cover contributing to higher temperatures. The study also emphasizes the importance of considering urban morphology in the design and planning of sustainable cities.

The data analysis between these two articles (Y. Zheng et al. (2023); K. Liao et al. (2021)) uses **spatial statistical analysis** to map the spatial distribution of nocturnal UHIs, while the second article uses a **regression analysis** to examine the relationship between urban morphology and SUHIs.

C. Reis, et al. (2021) [18] presents maps of Urban Heat Island (UHI) for the Lisbon Metropolitan Area (LMA) based on local weather patterns. The maps were created by analysing **hourly air temperature data**. An **urban mask was created from Local Climate Zones (LCZ) classification** (using the same approach of Y. Zheng et al. (2023)), and UHI intensity was estimated using an R script. The maps were divided by thermal seasons and local weather types (LWT) with average UHI intensity estimated for each LWT.

The study found that UHI effects are more pronounced in areas with less vegetation and more impermeable surfaces, particularly during periods of high temperature and low wind speed.

Research conducted by X. Liu, et al. (2018) [19], discusses the various parameters that are used to calculate the roughness of urban surfaces, primarily focusing on the frontal area index (FAI) as the most reliable indicator of roughness. It goes on to discuss the **planning of urban ventilation corridors**, which must consider factors such as wind environment, ventilation potential, and heat island intensity. The passage explains how various indicators are used to assess the wind environment, including DEM, vegetated space, open space, building density, building height, frontal area density (FAD), road density, road connectivity, and examines the correlations between FAD and these factors

[18] Reis, C., Lopes, A., & Santos Nouri, A. (2021). Urban heat island data by local weather types in Lisbon metropolitan area based on Copernicus climate variables dataset for European cities. *Sustainable Cities and Society*, 65, 102618.

[19] Liu, X., Huang, B., Li, R., Zhang, J., Gou, Q., Zhou, T., & Huang, Z. (2018). Wind environment assessment and planning of urban natural ventilation corridors using GIS: Shenzhen as a case study. *Building and Environment*, 129, 100-113.

using Pearson's correlation analysis. The authors suggest that this approach could be applied to other cities facing similar issues with poor air quality and ventilation.

The study by *Wang W., et al. (2022)* [20] proposes a novel method that integrates land surface temperature retrieval, GIS spatial analysis, and weather data to identify functional and compensatory spaces, assess ventilation potential, and determine wind direction and speed.

The findings of the study demonstrate the potential of **constructing primary and secondary ventilation corridors** in Hangzhou, China, utilizing water bodies, green spaces, and roads. The article presents a cost-effective research model for UVC planning, using remote sensing images and basic urban and meteorological data.

The article by *V. Equere, et al. (2021)* [21] illustrates research that aimed to improve the **prediction of urban heat islands (UHI) by incorporating topographic variations** using a new terrain factor (TF) integrated with other morphological parameters. The study used Landsat images and LiDAR data to derive vertical and horizontal morphological parameters, which were trained in an artificial neural network (ANN) to predict the surface UHI indicated by land surface temperature (LST) in two areas in Illinois, USA with high and low topographic variations. Results showed that the model trained with TF performed better in predicting UHI in areas with complex surface elevation. The study also found that NDBI and NDVI were the most significant parameters influencing UHI in urban areas, while TF was a significant factor in areas with high topographic variations. The study concluded that the inclusion of TF is essential for explaining UHI formation in areas with significant topographic variations and suggested that future studies should consider the impact of atmospheric and climatic factors.

[20] Wang, W., Wang, D., Chen, H., Wang, B., & Chen, X. (2022). Identifying urban ventilation corridors through quantitative analysis of ventilation potential and wind characteristics. *Building and Environment*, 208, 108759.

[21] Equere, V., Mirzaei, P. A., Riffat, S., & Wang, Y. (2021). Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using GIS and ANN. *Sustainable Cities and Society*, 69, 102825. <https://doi.org/10.1016/j.scs.2021.102825>

C.M. Nakata-Osaki, et al (2018) [22] created a simulation model called THIS, which is integrated into a GIS and designed to *estimate the maximum intensity of urban heat islands by utilizing urban geometry data*. The researchers used a theoretical-numerical approach (*Oke model*) and tested the model on two Brazilian cities. The main purpose of the tool is to simulate *how urban geometry affects the maximum intensity of nocturnal heat islands*. The study found that different urban settings had varying trends in the UHI_{\max} value for the H/W ratio and roughness length. Specifically, *urban canyons* with greater roughness showed UHI_{\max} values that were roughly half of the value of canyons with less roughness for the same H/W ratio.

A. Oliveira, et al (2021) [23] discusses a *machine learning approach to predict the synthetic nocturnal surface Urban Heat Island (SUHI) during a heatwave event* in Naples. The authors calculate urban energy balance components for diurnal Landsat-8 imagery and Meteosat's Spinning Enhanced Visible and InfraRed Imager (MSG-SEVIRI) to understand the relation with the nocturnal SUHI. They test a data fusion approach to develop a sub-kilometric synthetic nocturnal LST prediction model based on diurnal high-resolution heat flux components and nocturnal kilometric thermal imagery. The Random Forest algorithm is used for machine learning model development. The study explores an energy balance-based machine learning approach to predict the nocturnal SUHI during an HW event, ensuring mean square errors inferior to 0.1 K. The results support the adoption of the energy balance approach, and the study's method and results are a precursor to upscale an energy balance-based data fusion approach for SUHI prediction, which would provide an efficient urban climate monitoring tool whenever in-situ observations are not readily available.

[22] Nakata-Osaki, C. M., Souza, L. C. L., & Rodrigues, D. S. (2017). Tool for Heat Island Simulation: A GIS extension model to calculate urban heat island intensity based on urban geometry. *Computers, Environment and Urban Systems*, 65, 104-117. doi: 10.1016/j.compenvurbsys.2017.04.002

[23] Oliveira, A., Lopes, A., Niza, S., & Soares, A. (2021). An urban energy balance-guided machine learning approach for synthetic nocturnal surface Urban Heat Island prediction: A heatwave event in Naples. *Building and Environment*, 195, 107749. <https://doi.org/10.1016/j.buildenv.2021.107749>

In the article by G. Mutani and S. Beltramino (2022) [24], The authors describe the ***development and validation of a geospatial model for assessing outdoor thermal comfort in urban areas, using data from field measurements, satellite imagery, and numerical modelling***. They also present a case study of the model's application in the city of Turin, Italy, to evaluate the thermal comfort conditions of different urban typologies and land uses.

The model (SOLWEIG) uses high-quality Digital Surface Models and weather data to simulate the effects of radiant flux densities and mean radiant temperature. The study compares SOLWEIG with ENVI-met for thermal comfort analyses in various urban spaces in hot and cold weather conditions. The results show that SOLWEIG is a more suitable tool for assessment and analyses at the urban scale, while ENVI-met is more useful for feasibility studies with high spatial and temporal resolution or for the pre-design phase of little neighbourhoods. The study highlights the importance of evaluating outdoor thermal comfort across the whole city to prioritize interventions and define the urban characteristics for more liveable outdoor spaces.

V. Todeschi and G. Mutani, et al. (2021) [25] in their study describe the development and implementation of a ***multi-disciplinary approach to urban rooftop transformation***, which involves stakeholders from different sectors and disciplines, including architecture, engineering, and social sciences. They also provide an ***overview of the smart solutions and technologies that are used to improve the sustainability and functionality of urban rooftops***, such as green roofs, solar panels, and rainwater harvesting systems.

The study finds that the Re-Coding project has the potential to make significant contributions to the sustainability and liveability of cities, by enhancing the quality of urban spaces and reducing the environmental impact of buildings. The authors suggest that the multi-disciplinary approach and smart solutions employed in the project can serve as a model for other urban development initiatives.

[24] Mutani, G., & Beltramino, S. (2022). *Geospatial assessment and modeling of outdoor thermal comfort at urban scale*. *Sustainable Cities and Society*, 77, 103245. <https://doi.org/10.1016/j.scs.2021.103245>

[25] Todeschi, V., Mutani, G., Baima, L., Nigra, M., & Robiglio, M. (2021). *Smart Solutions for Sustainable Cities—The Re-Coding Experience for Harnessing the Potential of Urban Rooftops*. *Sustainability*, 13(16), 8954. <https://doi.org/10.3390/su13168954>

The paper "Urban Heat Island Mitigation: A GIS-based Model for Hiroshima" by G. Mutani, et al. (2019) [26] presents a GIS-based model that evaluates the urban heat island (UHI) effect in the city of Hiroshima and proposes mitigation measures to reduce its impact. The authors used ***GIS analysis to calculate the surface temperature of the city and identify areas with the highest UHI intensity***. They also conducted surveys to gather data on the thermal comfort of residents and the urban morphology of the city. The methodology used for evaluating air temperature variations and UHI effects on the city is described, and variables such as satellite images, weather station (WS) data, and indicators are used to construct the models. The main variables that influence air temperature were identified, including altitude, vegetation, and water, which reduce air temperature. Linear regression models of air temperature are presented, with the best results obtained from models including all non-normalized variables or normalized variables without LST. The models show that air temperature is higher in urban areas than in the peripheral plain and mountain areas, which are mitigated by altitude, vegetation, and lower building density. The UHI intensity (UHII) is used as an indicator to measure the hourly and daily amplitude and temperature gradient of the air between the urban and surrounding rural areas. The model results suggest that interventions on variables such as buildings density and relative building height can improve microclimatic conditions and mitigate air temperature.

In the next part, of the literature review, our goal was to identify articles that implemented machine learning techniques in the evaluation of urban heat islands (UHIs). This involved a targeted search of academic databases and other relevant sources using specific keywords and search terms related to machine learning and UHIs.

Research by D. Espino, et al. (2022) [27] the article presents ***a new GIS add-in called ArcUHI that allows for the automated modelling of the Urban Heat Island (UHI) effect***

[26] Mutani, G., Todeschi, V., & Matsuo, K. (2019). Urban Heat Island Mitigation: A GIS-based Model for Hiroshima. *Sustainability*, 11(8), 2369. <https://doi.org/10.3390/su11082369>

[27] Espino, D. J., Manchado, C., Valcarce, A. R., & Moscardò, V. (2022). ArcUHI: A GIS add-in for automated modelling of the Urban Heat Island effect through machine learning. *Environmental Modelling & Software*, 146, 105176. doi: 10.1016/j.envsoft.2021.105176

using machine learning techniques. The add-in is designed to work with ArcGIS software and allows users to quickly and easily generate UHI models using a variety of machine learning algorithms. The authors provide a detailed description of the ArcUHI add-in and its capabilities, and present results from a case study in the city of Oviedo, Spain, where the add-in was used to model the UHI effect. The results demonstrate that ArcUHI can generate accurate UHI models using machine learning techniques and has the potential to be a valuable tool for urban planning and management. The article concludes with a discussion of the potential applications and limitations of the ArcUHI add-in, as well as directions for future research. The article explains the process of using satellite data to extract surface temperature information, and the creation of Digital Surface Models (DSM) from building rasterizations. Nonlinear machine learning algorithms such as Support Vector Regression (SVR) and Random Forest Regression (RFR) were used to model the UHI effect. The results showed the suitability of using ΔT (the difference between land surface temperature and air temperature) as a representative for UHI. The add-in can be used to propose retrofitting measures for SUHI mitigation, such as roof and wall greening, planting trees, and replacing dark asphalt roads with "cool" pavements.

The article by *J. Lin, et al. (2022)* [28] *investigates the relationship between the spatial pattern of green spaces and urban heat island (UHI) intensity using machine learning methods.* The study found that UHI intensity was negatively related to the density of green space and positively related to the density of built-up areas, population density, and mean building height. Additionally, the study found that the morphological characteristics of green space could exert a substantial influence on UHI intensity. The random forest model suggested that the density of core green space can substantially contribute to the mitigation of UHI effect in the study area. The study suggests that improving the morphological spatial pattern of green space through land use planning would be more helpful in mitigating the UHI effect.

[28] Lin, J., Qiu, S., Tan, X., & Zhuang, Y. (2022). *Measuring the relationship between morphological spatial pattern of green space and urban heat island using machine learning methods. Sustainable Cities and Society*, 95, 103877. doi: 10.1016/j.scs.2021.103877

The paper by P. Mohammad, et al. (2022) [29] describes a study that uses **machine learning algorithms to predict land use land cover (LULC) and land surface temperature (LST) changes** to characterize the surface urban heat island (SUHI) phenomena in Ahmedabad, India. The study uses remote sensing data from Landsat 8 satellite images to extract LULC and LST data for the period from 2013 to 2020. The authors then use machine learning algorithms, specifically random forest and support vector regression, to model the relationships between LULC and LST and to predict LST changes based on LULC changes.

The study conducted by S. Liu, et al. (2021) [30] used a combination of remote sensing data, meteorological data, and field measurements to build a model that predicted UHI intensity in the study area. The model was then used to simulate **different mitigation scenarios, such as the addition of green roofs and cool roofs to buildings, to determine their effectiveness in reducing UHI intensity.**

The results of the study showed that the machine learning model was able to accurately predict UHI intensity in the study area, and that the addition of green roofs and cool roofs had a significant impact in reducing UHI intensity. The authors suggest that the use of machine learning models in UHI mitigation planning can help to identify the most effective strategies for reducing UHI effects in urban areas.

Quite similar research has been conducted by G. Mutani, et al. (2020) [31] this study looks at **how green surfaces integrated with building envelopes and urban morphology can mitigate the urban heat island (UHI)**, improve thermal comfort, and save energy. Using data from seven weather stations in Turin, the study finds that covering roofs with vegetation and creating green urban areas can have several benefits, including increasing biodiversity, decreasing pollution and storm water runoff, and improving indoor and

[29] Mohammad, P., Goswami, A., Chauhan, S., & Nayak, S. (2022). Machine learning algorithm-based prediction of land use land cover and land surface temperature changes to characterize the surface urban heat island phenomena over Ahmedabad city, India. *Sustainable Cities and Society*, 96, 103952. doi: 10.1016/j.scs.2021.103952

[30] Liu, S., Zhang, J., Li, J., Li, Y., Zhang, J., & Wu, X. (2021). Simulating and mitigating extreme urban heat island effects in a factory area based on machine learning. *Sustainable Cities and Society*, 70, 102977. doi: 10.1016/j.scs.2021.102977.

[31] Mutani, G., & Todeschi, V. (2020). The Effects of Green Roofs on Outdoor Thermal Comfort, Urban Heat Island Mitigation and Energy Savings. *Sustainability*, 12(2), 616. doi: 10.3390/su12020616.

outdoor thermal comfort. The study also shows that outdoor and indoor thermal comfort increases with vegetation and green roofs, higher values of albedo, and lower building density, while UHI discomfort increases with higher values of canyon height-to-width ratio and building density. The study proposes requalifying critical areas in Turin with green roofs, parks, and rows of trees to improve thermal comfort, reduce energy consumption, and mitigate the UHI effect.

The authors, *S. Wang et al., (2023)* [32] conducted a ***study that explores the relationship between health and exposure to heat and air quality in different urban areas*** across Australia. The authors use global and local machine learning models to analyse how these factors affect health outcomes, and they examine how this relationship varies depending on specific geographic locations. The study finds that heat and air quality are significant predictors of health outcomes, but the effects vary across different urban areas. The authors suggest that policymakers should use these findings to inform targeted interventions to mitigate the negative health impacts of urban heat and air pollution.

The last of the researchers analysed is by *T. Lan, et al. (2023)* [33], the study uses a ***hybrid approach that combines global climate models and regional climate models to predict the future UHI effect in China under different climatic and socioeconomic scenarios***. The results show that the UHI effect is expected to increase in the future under all scenarios, but the magnitude of the increase varies depending on the policies implemented. The study also highlights the potential benefits of implementing sustainable policies to mitigate the UHI effect, such as green infrastructure and energy-efficient buildings. Overall, the study provides insights into the future of UHI in China and the potential role of policies in mitigating its impact.

[32] Wang, S., Cai, W., Tao, Y., Sun, Q. C., Wong, P. P. Y., Huang, X., & Liu, Y. (2023). Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models. *Environmental Research*, 204, 112012. <https://doi.org/10.1016/j.envres.2022.112012>

[33] Lan, T., Peng, J., Liu, Y., Zhao, Y., Dong, J., Jiang, S., Cheng, X., & Corcoran, J. (2023). The future of China's urban heat island effects: A machine learning based scenario analysis on climatic-socioeconomic policies. *Science of The Total Environment*, 153114. <https://doi.org/10.1016/j.scitotenv.2022.153114>

1.6 Environmental protocols for reducing UHI effect

Environmental protocols such as ITACA and LEED play a crucial role in the planning and development of new neighbourhoods with a focus on sustainability. These protocols provide a framework for assessing and promoting sustainable development in the built environment, considering various environmental factors and impacts.

When planning a new neighbourhood, environmental protocols can be used to guide decisions related to building design, energy efficiency, water management, waste reduction, and more. For example, protocols may recommend strategies for reducing energy consumption, such as the use of renewable energy sources, high-efficiency insulation, and low-energy lighting. They may also promote the use of sustainable materials, such as recycled or locally sourced building materials.

In addition to promoting sustainable development, environmental protocols can also help improve the quality of life for residents. For example, by promoting green spaces and vegetation, protocols can help mitigate the *effects of urban heat islands* and improve air quality. By reducing water consumption and managing stormwater runoff, protocols can help prevent flooding and protect water resources.

Overall, environmental protocols provide a roadmap for creating sustainable, liveable neighbourhoods that balance environmental, economic, and social considerations. By following these protocols, planners and developers can help create communities that are more resilient, healthy, and equitable for all residents.

This chapter will provide a technical examination of the environmental protocols ITACA and LEED, specifically with regards to the *methods used to calculate urban heat island* (UHI). The UHI calculation is an essential component of these protocols as it helps to identify the factors that contribute to the urban heat island effect, such as building density, materials, and vegetation cover. By examining the UHI calculation methods in ITACA and LEED, we can gain a deeper understanding of how these protocols help promote sustainable development in the built environment and mitigate the environmental and health impacts of urban heat islands.

ITACA (Integrated Assessment for the sustainable Built Environment) is an environmental protocol developed in Italy to assess the sustainability of buildings and

neighbourhoods. It considers various aspects, such as energy efficiency, water consumption, materials used, and indoor comfort. The aim is to promote sustainable development in the built environment.

LEED (Leadership in Energy and Environmental Design) is a similar protocol developed in the United States. It provides a framework for designing and constructing green buildings and neighbourhoods, with a focus on energy efficiency, water conservation, and sustainable materials.

In urban areas, the phenomenon of urban heat islands can have significant environmental and health impacts. To address this issue, various methods are used to calculate the urban heat island (UHI) in different neighbourhoods. This involves measuring the temperature differences between urban and rural areas, and analysing factors that contribute to the urban heat island effect, such as building density, materials, and vegetation cover.

The subsequent section outlines through a table, the factors considered for the identification and calculation of urban heat island (UHI) in the two environmental protocols.

	Solar Reflec tance index (SRI)	Urban Morpho logy (H/D ratio, SVF Sky view factor)	Soil permea bility	Green roof and vegetati on	Waste heat from urban anthropoge nic activities (e.g., air conditioner, transports)	Presen ce of water (artifici al and natural)	Materi als
ITACA	✓	✓	✓	✓	✓	✓	✓
LEED	✓	✗	✗	✓	✗	✗	✓

Table 1 - Factors of UHIs in Environmental Protocols

The green symbols indicate that the protocol accounts for the factor in its calculation of the final score, whereas the red cross indicates that the protocol does not consider that factor. After analysing these two protocols, we can conclude that the ITACA protocol

provides a more comprehensive examination of the factors that contribute to the Heat Island (HUI) effect.

According to **Protocollo ITACA a Scala Urbana SINTETICO** [34], is reported how the factors shown in the table above were calculated for evaluating the Urban Heat Island (UHI) effect, according to the protocol's guidelines.

Solar Reflectance Index (SRI): is an indicator of a surface's ability to reflect solar radiation and dissipate the heat generated by absorbed radiation. The SRI is defined as a combination of solar reflectance (albedo) and surface thermal emittance and is expressed on a scale of 0 to 100, where higher values indicate greater heat dissipation ability.

H/D ratio: is a measure of the relationship between the height of buildings and the size of the open space they overlook. Specifically, it is calculated by dividing the height of the buildings (H) by the dimension of the open space (D). Lower values of the H/D ratio indicate that buildings are shorter in relation to the open space, resulting in a greater proportion of open sky. This increased exposure to the atmosphere facilitates greater heat loss, which can contribute to a reduction in temperature. From a technical perspective, this phenomenon can be explained by the fact that buildings act as barriers to the flow of air, reducing the rate of heat exchange between the surface of the earth and the atmosphere. By contrast, open spaces allow for greater convection and radiation, which facilitates heat loss to the atmosphere. This effect is particularly important in urban environments, where high concentrations of buildings can lead to a phenomenon known as the "urban heat island" effect. By optimizing the H/D ratio to maximize open space and minimize building height, it may be possible to mitigate the impacts of this effect and reduce overall energy consumption.

Sky view factor (SVF): is a measure of the visible portion of the sky from a given observation point. An SVF value of 0.0 indicates that the sky is completely obstructed and therefore not visible, while an SVF value of 1.0 indicates that the sky is completely visible from all angles. For this reason, the higher the calculated SVF, the greater the heat loss to the atmosphere. This effect is particularly important in urban environments, where the high concentration of buildings can obstruct the view of the sky and limit the potential

[34] Istituto Per L'innovazione E Trasparenza Degli Appalti E La Compatibilità Ambientale Itaca, Protocollo Itaca A Scala Urbana Sintetico, Metodologia E Strumento Di Verifica, Versione 2.02 (2020)

for heat exchange. By optimizing the SVF to maximize the visible sky, it may be possible to reduce the impacts of the "urban heat island" effect and improve overall energy efficiency.

Soil permeability: it refers to the ability of soil to allow water or other fluids to pass through it. Soils with high permeability allow water to flow through them easily, while soils with low permeability tend to hold water. In the context of urban design and management, permeable soils are important in reducing the urban heat island effect. The scenario will be satisfied if has been considered and reached the criteria in the evaluation criteria area n°5.01 of the document mentioned in citation [34].

Green roof and vegetation: The use of **green roofs** or materials that guarantee a Solar Reflectance Index (SRI) of at least 29 [35] is preferred if the slope is greater than 8.5°, and 76 in the case of roofs with a slope of less than or equal to 8.5°. **Green surfaces:** the presence of green surfaces provides solar protection, cooling of the ambient air through evapotranspiration, and improves air quality. The presence of green roofs and vertical gardens can also significantly contribute to the reduction of urban temperatures. The scenario will be satisfied if has been considered and reached the criteria in the evaluation criteria area n°6 of the document mentioned in citation [34].

Waste heat from urban anthropogenic activities: reduce the presence in open spaces of waste heat generated by anthropogenic activities (transportation, air conditioning, etc.). The scenario is satisfied by highlighting, during the pre-operation phase, which activities generate waste heat and the strategies adopted to reduce their effects, which will then be verified during the post-operation phase.

Presence of water: The presence of natural low-temperature reservoirs such as ponds, puddles, fountains, etc. contributes to lowering the temperature through evapotranspiration. The scenario is satisfied by highlighting in the pre-operative phase what natural heat dissipators are present and/or expected. Their presence can later be verified in the post-operative phase.

Materials: Materials: in open spaces, on the facades of buildings facing these spaces, and on their roofs, reflective or so-called cool materials should be favored, which have

[35] Istituto Per L'innovazione E Trasparenza Degli Appalti E La Compatibilità Ambientale Itaca, Protocollo Itaca A Scala Urbana Sintetico, Metodologia E Strumento Di Verifica, Versione 2.02 (2020)

high reflectivity to solar radiation and a high coefficient of emissivity. In the case of materials with low absorption of solar radiation and high infrared emission, the surface temperature of the materials is minimized, thus reducing the amount of heat released into the atmosphere. These materials include, for example: natural materials that have high reflectivity to solar radiation; artificial coatings in white or light colors; coatings colored with high reflectivity in the infrared spectrum; "smart" coatings mixed with thermochromic substances or paints and phase change materials (PCM) that improve the thermal and optical properties of the coating.

The evaluation or choice of materials can be carried out in the following way: external paved or pedestrian or bicycle use surfaces (e.g., pedestrian paths, sidewalks, squares, courtyards, cycle paths, etc.): the use of permeable materials (e.g., draining materials, green surfaces, pavements with open meshes or grated elements, etc.) with a Solar Reflectance Index (SRI) of at least 29 should be provided [36].

Overall, both protocols aim to mitigate the UHI effect, but they have some minimal differences in terms of the calculation and mitigation strategies, reflecting the different priorities and contexts of the regions in which they were developed.

LITERATURE TABLE

[36] Istituto Per L'innovazione E Trasparenza Degli Appalti E La Compatibilità Ambientale Itaca, Protocollo Itaca A Scala Urbana Sintetico, Metodologia E Strumento Di Verifica, Versione 2.02 (2020)

2. Studying urban heat islands: the case of Turin

In this research, we have selected the *city of Turin*, Italy, as a case study to investigate the phenomenon of urban heat islands. Turin is the capital city of the Piedmont region in northern Italy and has a population of approximately 900,000 people. It is a historically and culturally rich city with a significant industrial past, that is the reason of the need of significant urban transformation. Despite these efforts, **the city is still susceptible to the formation of urban heat islands** due to its dense urban fabric and high concentration of anthropogenic heat sources.

In this study, we aim to analyse the spatial and temporal patterns of surface temperature in Turin, and to evaluate the intensity of the urban heat island effect and its impact on the local climate.

Furthermore, the study will examine the *impact of large-scale urban interventions*, such as the redevelopment of former industrial areas and the expansion of green infrastructure, on the urban heat island phenomenon.



2.1 Main purposes of the UHI analysis in Turin

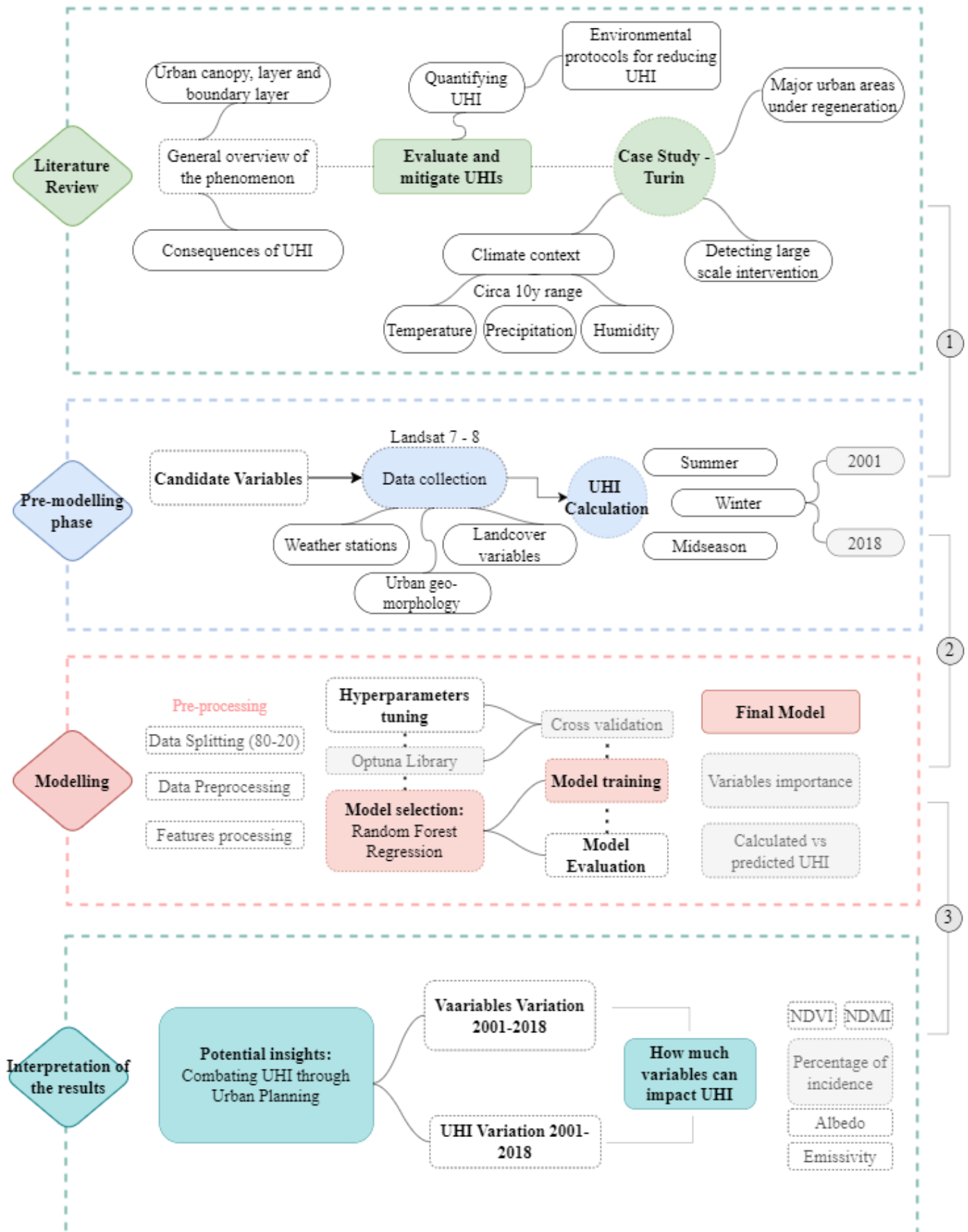
Within the framework of this thesis, dedicated to the exploration of the *urban heat island (UHI) phenomenon in Turin*, the analysis serves crucial objectives aligned with the specific urban and environmental dynamics of the city. In particular, the survey focuses on the former *TEKSID area*, a major industrial area undergoing a comprehensive redevelopment initiative. This context injects additional layers of significance into the analysis as it examines UHI dynamics in the wake of urban redevelopment. The analysis is driven by the following primary purposes:

- **Contextual Understanding:** a paramount objective is to dissect Turin's distinctive UHI dynamics within the context of the Ex-area TEKSID's transformation. By discerning how urbanization, combined with post-industrial regeneration actions, influences local microclimate patterns, the study yields insights tailored to the specific nuances of this area. Understanding how regeneration impacts UHI dynamics is essential for refining mitigation strategies for areas undergoing similar urban rejuvenation efforts.
- **Mitigation Strategy Formulation:** an integral aspect of this analysis is to pinpoint variables that significantly contribute to UHI intensity within the Ex-area TEKSID. The goal is to quantitatively assess whether the regeneration intervention has mitigated UHI effects. By evaluating the impacts of various factors on UHI, the analysis informs the development of targeted strategies that harness the benefits of urban regeneration to alleviate temperature differentials.

In essence, this analysis of UHI, focusing on the transformation of the Former TEKSID area, seeks to contribute scientifically rigorous insights into the effectiveness of regeneration actions in mitigating the effects of UHI.

By linking scientific inquiry with pragmatic concerns of urban redevelopment, the study seeks to provide a data-driven basis for resilient, comfortable, and sustainable urban environments in revitalized areas.

2.2.1 Visualizing Thesis Methodology: Flowchart Overview



2.2 Understanding Climate Context: A Brief Overview

A climate context overview provides a broad understanding of the climate conditions in a particular region. It encompasses various factors such as temperature, precipitation, seasonal variations, and notable climate patterns. This overview aims to provide a foundation for understanding the typical climate characteristics and trends in the area.

The overview usually includes information about the average annual temperature range, highlighting the general temperatures experienced throughout the year. It may also mention the seasonal temperature variations, describing the differences in temperature between the various seasons.

Precipitation is another essential aspect covered in the overview. It typically includes the average annual precipitation amount, indicating the level of rainfall or snowfall the region receives over a year. The overview might also highlight the seasonal distribution of precipitation, pointing out the wettest and driest months or any notable patterns observed.

In addition to temperature and precipitation, a climate context overview may touch upon other relevant aspects. These can include humidity levels, prevailing winds, the occurrence of extreme weather events, or specific climatic phenomena associated with the region.

Furthermore, if applicable, the overview may address the impact of climate change on the area. It might briefly mention any observed changes or trends resulting from climate change, such as temperature shifts, altered precipitation patterns, or increased frequency of extreme weather events.

A climate context overview aims to provide a concise yet informative summary of the climate conditions in each region. It serves as a starting point for understanding the typical weather patterns, seasonal variations, and potential climate-related challenges in that area.

The climatic data comes from the weather station chosen in: “**Torino, Via della Consolata, 10**”.

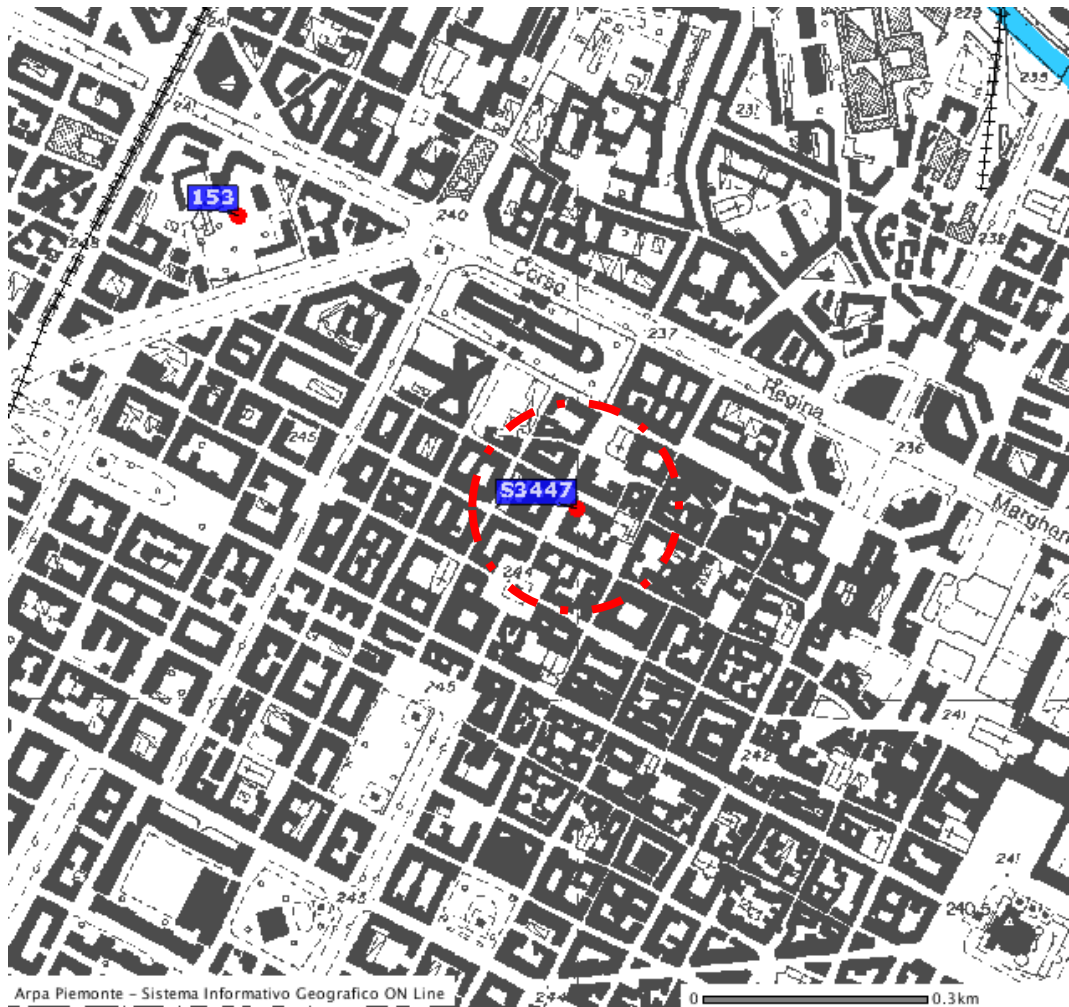


Figure 6 - Weather station localization

■ Temperature chart

Year	J	F	M	A	M	J	J	A	S	O	N	D
2010	1.7	4.4	8.9	14.6	17.6	22.3	26.7	23.4	19.2	12.9	8.5	2.5
2011	3.3	6.8	9.8	17.3	20.5	21.7	23	25.6	22.5	14.7	9	6.1
2012	4.4	2.7	13.6	12.6	18.5	23.7	25.3	25.9	19.9	15	10	3.9
2013	4.6	3.9	7.9	13.6	16	22.4	25.9	24.6	20.5	14.6	9.2	5.5
2014	5.7	7.1	12	15.6	17.9	22.7	22.6	22.2	20.1	16.2	10.6	6.3
2015	5.6	5.4	10.8	15.1	19.2	23.4	28.3	24.4	18.9	13.6	9.8	6.4
2016	5.1	7.2	10	14.9	17.1	22	25.5	24.7	22	13.5	9.6	5.5
2017	2.9	6.9	12.9	15.1	19.4	24.9	25.6	25.7	19	15.8	8.7	3.1
2018	6.5	4	7.9	16.2	18.5	23.7	26	25.9	21.7	15.7	9.7	4.9
2019	4	7.6	12.1	13.5	15.9	24.4	26	24.8	20.3	15.7	8.8	6.7
2020	4.3	9.4	9.8	15	19	21	24.9	25	20.8	13.3	9.9	5.1
2021	3.9	8	10.9	12.6	16.9	23.7	24.3	24.3	21.5	14.2	8.8	4.8
2022	4.8	8.2	9.6	14.1	20.5	25.4	28	25.9	20.4	17.8	10.1	4.6
Mean Temperature	4.37	6.28	10.48	14.63	18.23	23.18	25.55	24.80	20.52	14.85	9.44	5.03
Std.	1.27	2.01	1.79	1.36	1.51	1.30	1.65	1.09	1.15	1.38	0.65	1.28
(+) Std	5.64	8.28	12.27	15.99	19.74	24.47	27.20	25.89	21.67	16.23	10.09	6.31
(-) Std	3.10	4.27	8.68	13.28	16.72	21.88	23.90	23.71	19.37	13.46	8.79	3.76

Table 2 - Mean monthly temperature over years [°C]

■ Precipitation chart

Year	J	F	M	A	M	J	J	A	S	O	N	D
2010	38.6	58.6	49.4	43.4	172.4	261.4	24.2	230.8	84.2	113.8	173.8	76.4
2011	13.6	35	196.8	44.6	37.4	249.8	171.2	28.8	150.6	13.8	229	0.8
2012	35	8	38.4	168.8	133.2	88.8	101.2	38.8	52.4	48.8	167.6	2
2013	11.8	12	112	188.6	166.4	60.2	107	100.4	34.4	67	104	79.8
2014	71	104.4	71.2	65.4	115.6	74.2	202.8	75.6	48	31.2	238.2	67.2
2015	13.4	102.6	122.2	87	29.8	73.4	29.8	195.6	51.4	197.4	0	0
2016	4.4	107.6	57	63.8	166.4	108.4	91.4	77	25.4	61.4	171.8	50.2
2017	2.6	33.4	58.4	42	66.6	127	40.8	35.6	24.2	0	37.6	24
2018	88.8	62.8	76	86.6	195	96.2	131.4	44.6	57.6	153.4	100	5.2
2019	2	19.8	10	122.8	135	90	247.8	70.6	43.8	129.6	269.2	72.2
2020	0	0	21	129.6	89	224.4	43.6	184.4	32.6	90	1.6	31.8
2021	85	19.2	0.8	52.4	78.6	181.8	172	51.6	65.4	33	116.2	11.8
2022	0.4	4.4	8.2	12.2	57.2	13.8	39.2	42	25.8	40.4	26.6	40
Mean	28.20	43.68	63.18	85.17	110.97	126.88	107.88	90.45	53.52	75.37	125.82	35.49
Std	32.99	39.70	55.03	52.91	55.12	77.83	73.04	68.26	34.01	58.47	91.34	30.87
Std +	61.19	83.37	118.21	138.07	166.09	204.71	180.92	158.71	87.53	133.84	217.16	66.36
Std -	-4.79	3.98	8.16	32.26	55.84	49.05	34.83	22.19	19.52	16.90	34.47	4.62

Table 3 - Mean monthly precipitation over years [mm]

The table depicting temperature and precipitation data displays the monthly averages for precipitation in *millimetres* and temperature in *degrees Celsius* for each year. At the bottom of the table, are present the basic statistics for each month, which will be utilized in the construction of the climograph.

▪ **Climograph – mean temperature and precipitation tendency (from 2010 to 2022)**

Climograph offer a concise way to visually represent complex climate data. It combines **average temperature and precipitation data** into a single graph, using line graphs for temperatures and bars for precipitation. In this study, we focus on **Turin** and use data **from 2010 to 2022** to create a climograph. This graph provides a quick overview of Turin's typical temperature and rainfall patterns over this period, forming a basis for deeper analysis of the city's climate.

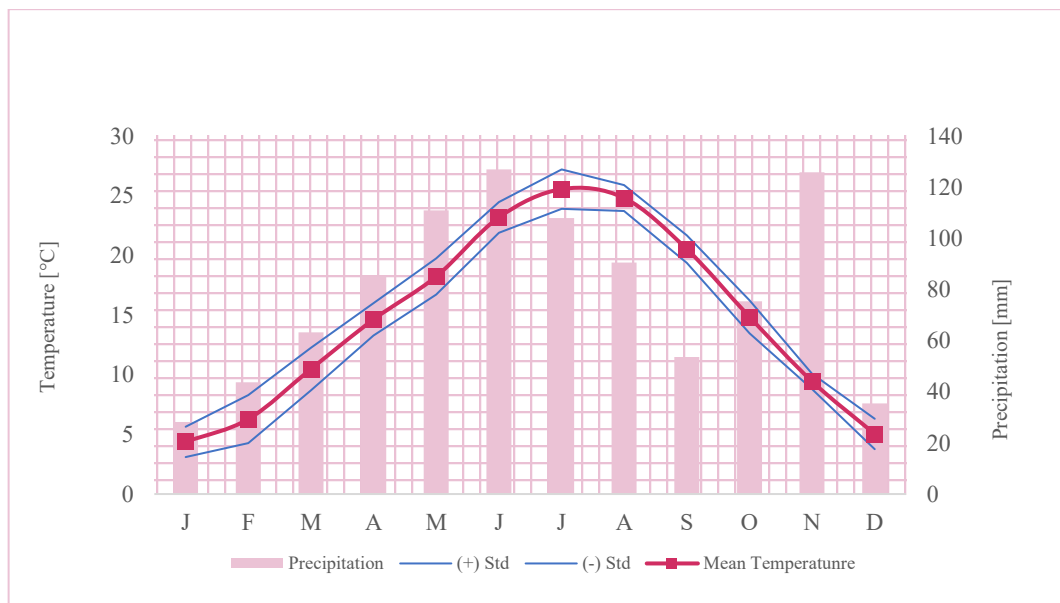


Table 4 – Climograph showing temperature (line) and precipitation (bar)

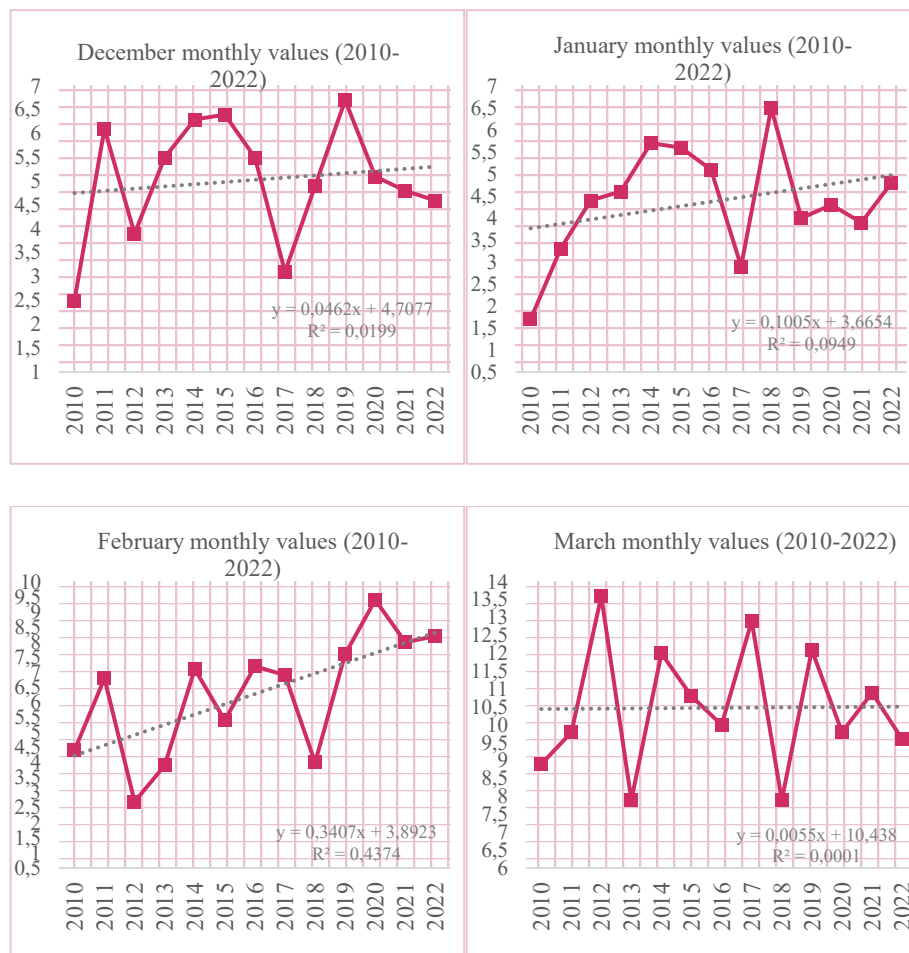
Following the presentation of the graph, we will delve into the process of calculating trends over the years.

▪ Trends calculation

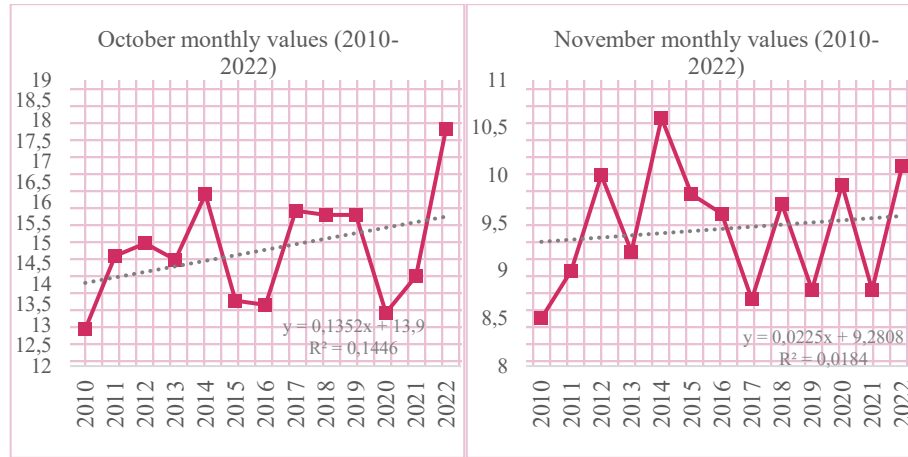
To calculate the trends in precipitation and temperature over the years has been followed these steps:

- For each year, it was created a graph reporting the average temperature/precipitation by each month.
- It was displayed the trend line and the equation of the graph.
- To calculate the trend, it was taken the average value coming from the averages of all the years for a given month, then multiplied for the angular coefficient of the equation displayed in the graph related to the month considered.

Here following, is shown the process using our data:







Graph 1 - Temperature [°C] trend analysis

Month	Temperature trend over years [°C]
J	0.437
F	2.134
M	0.052
A	-1.170
M	0.250
J	3.477
J	2.299
A	1.934
S	0.944
O	2.004
N	0.212
D	0.231
Average trend from 2010 to 2022	1.067

Table 5 – Temperature [°C] trend over years (2010 - 2022)

■ Temperature Trends in Turin (2010-2022): A Comprehensive Analysis of Annual and Monthly Variations

- Annual Temperature Trends:** The temperature data from 2010 to 2022 reveals a picture of Turin's climate dynamics. The overall trend indicates a gradual increase in average temperatures over this time span. The *calculated average trend of 1.067°C* confirms this **warming trend**, suggesting that the city has experienced a notable rise in its average annual temperature during this period.

- **Monthly Temperature Variations:** Analysing the monthly temperature trends provides insights into the specific months that have contributed to this warming trend:
 - **January:** The data indicates a positive temperature trend over the years, with an increase of approximately 0.44°C per year. This reflects the warming trend even in the typically colder months, likely influenced by broader climate shifts.
 - **February:** Like January, February demonstrates a significant warming trend, with an average annual temperature increase of around 2.13°C .
 - **March:** March shows a minor positive trend, with an average increase of approximately 0.05°C per year. Although the rise is slight, it contributes to the overall warming pattern.
 - **April:** In contrast to the previous months, April exhibits a negative temperature trend, with an average decrease of about -1.17°C per year. This may be influenced by various factors, including changes in cloud cover and atmospheric circulation.
 - **May:** May experiences a slight positive trend of approximately 0.25°C per year. This contributes to the gradual warming observed during the spring months.
 - **June:** June's temperature trend is notably positive, with an increase of around 3.48°C per year. This substantial warming could have significant implications for summer temperatures and associated climate phenomena.
 - **July:** July follows suit with June, displaying a substantial warming trend of about 2.30°C per year. The combined effects of urbanization and broader climatic factors could be contributing to this increase.
 - **August:** August mirrors the trends seen in June and July, with a positive temperature trend of approximately 1.93°C per year. This consistency across the summer months underlines the significance of the warming pattern.
 - **September:** September's trend is positive, albeit relatively moderate, with an average increase of around 0.94°C per year. This contributes to the transition from summer to autumn temperatures.
 - **October:** October's trend indicates a warming of about 2.00°C per year. This considerable rise in autumn temperatures can have implications for various ecological and climatic interactions.

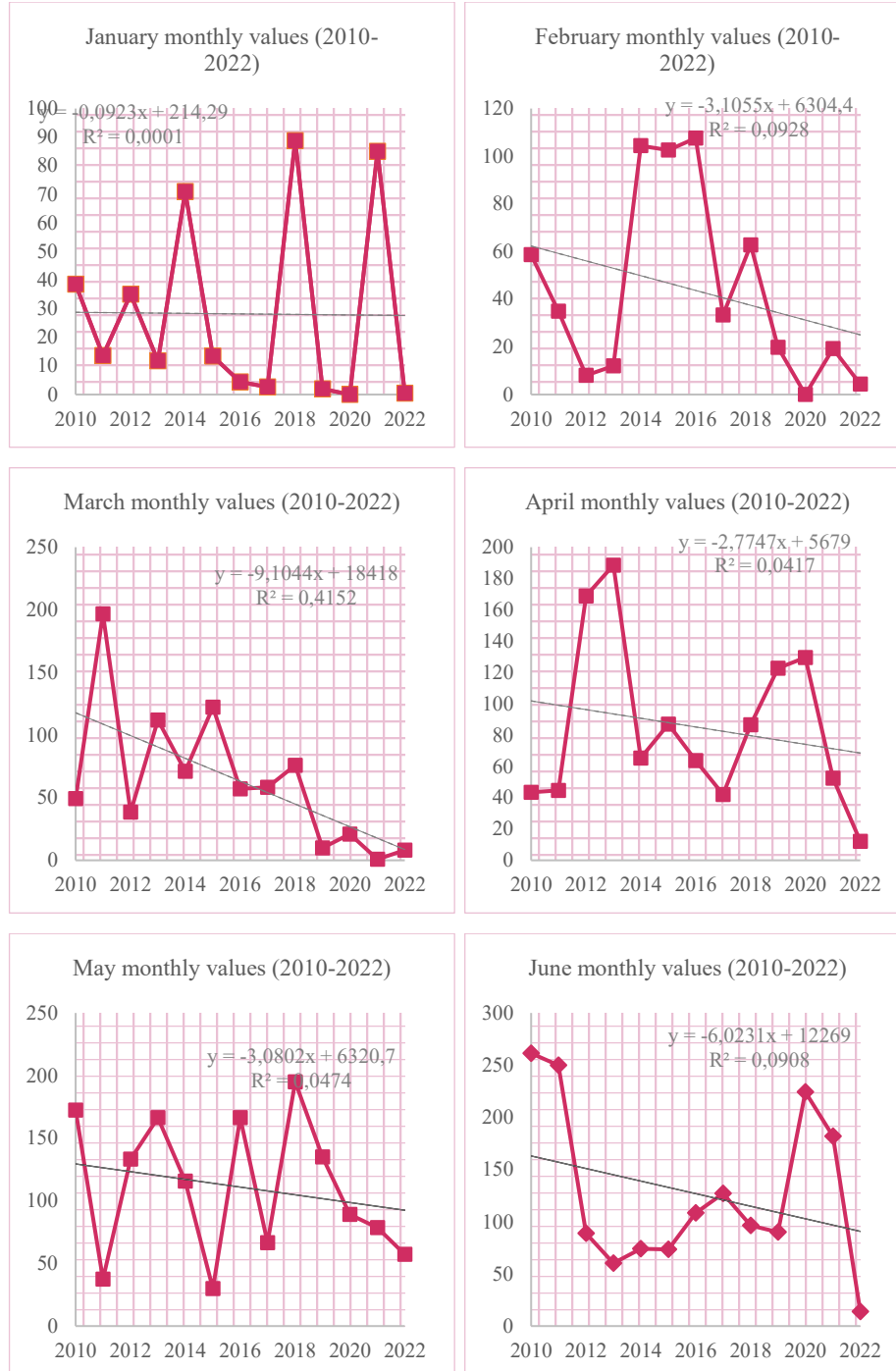
- **November:** The positive trend in November is relatively small, with an average increase of about 0.21°C per year. Nonetheless, it contributes to the overall upward temperature trajectory.
- **December:** December shows a minor positive trend, with an average increase of approximately 0.23°C per year. This indicates that even the colder months are experiencing a warming trend.

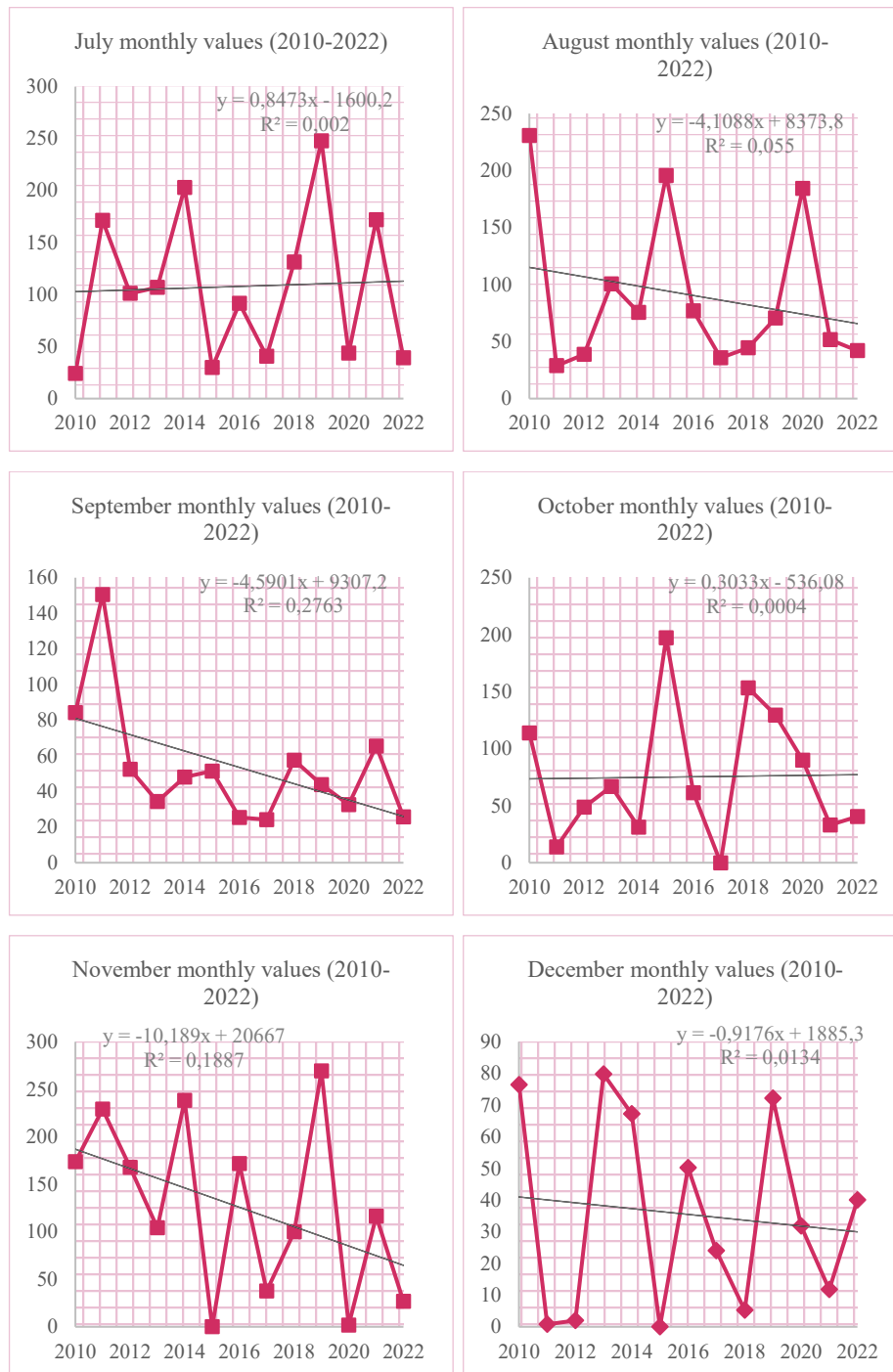
▪ **Climatic Implications:**

The observed temperature trends in Turin from 2010 to 2022 have broader climatic implications. The warming trend aligns with global patterns [37] of temperature increase attributed to **climate change**. It can impact various aspects of the city's ecosystem, including agriculture, urban planning, and water resources. Understanding these trends is vital for adapting to changing climate conditions and making informed decisions about sustainable development and resilience.

[37] IPCC Report 2023, “Summary for Policymakers”, page 16-17: IPCC_AR6_SYR_SPM.pdf

As for the temperature trends, the same process has been followed for the **precipitation**, here following are reported the related graphs:





Graph 2 - Precipitation [mm] trend analysis

▪ **Precipitation Trends in Turin (2010-2022): A Comprehensive Analysis of Annual and Monthly Variations**

Month	Precipitation trend over years [mm]
J	-19.455
F	-253.772
M	-668.936
A	-287.421
M	-352.660
J	-832.642
J	150.672
A	3.365
S	-234.501
O	2.487
N	-1222.762
D	-38.474
Average trend from 2010 to 2022	-312.842

Table 6 - Precipitation trend over years [mm] from 2010 to 2022

Annual Precipitation Trends:

The precipitation data from 2010 to 2022 reveals insights into Turin's precipitation patterns. The calculated delta of **-312.842 mm** indicates a notable decrease in average annual precipitation during this period.

Monthly Precipitation Variations:

Analysing the monthly precipitation trends provides insights into the specific months that have contributed to this overall decrease in precipitation:

- **January:** The negative precipitation trend in January (-19.45 mm) suggests a decrease in average precipitation over the years. This could impact the winter hydrological balance and water resources.
- **February:** February exhibits a significant negative precipitation trend of approximately -253.77 mm per year. This could influence early spring water availability and related ecological dynamics.

- **March:** The negative trend in March (-668.93 mm) points to a substantial reduction in average precipitation. This could impact soil moisture and vegetation during the transition to spring.
- **April:** April follows the trend of reduced precipitation, with a negative trend of -287.41 mm. This could influence agricultural activities and groundwater recharge.
- **May:** The negative precipitation trend in May continues (-352.60 mm), indicating a decrease in water availability during this critical growing season.
- **June:** June's trend is highly negative (-832.64 mm), highlighting a significant reduction in precipitation during the early summer months.
- **July:** July's positive trend of 150.67 mm suggests a slight increase in average precipitation. This could have implications for summer water availability and related sectors.
- **August:** August's positive trend of 3.36 mm indicates a modest increase in precipitation. However, this increase might not be substantial enough to counterbalance the negative trends of the preceding months.
- **September:** September's negative trend of -234.50 mm implies a reduction in fall precipitation, which can affect soil moisture and late-season vegetation.
- **October:** October's positive trend of 2.48 mm is relatively small. While any increase is valuable, the cumulative negative trends in previous months might still impact overall water resources.
- **November:** The pronounced negative trend in November (-1222.76 mm) suggests a substantial decrease in average precipitation during this month. This could influence autumn water availability and ecological processes.
- **December:** The negative trend in December (-38.47 mm) points to reduced precipitation during the late autumn and early winter months.

Climatic Implications:

The observed precipitation trends in Turin from 2010 to 2022 have significant climatic implications. The *overall decrease in precipitation aligns with broader patterns of changing precipitation regimes attributed to climate change*. Decreased water

availability can impact multiple sectors, including agriculture, water supply, and ecosystem health [38].

Interplay with Temperature:

It's important to note that *the negative precipitation trends coincide with the generally warming temperatures observed during this period*. The combination of warming temperatures and decreased precipitation raises concerns about *potential shifts in local climate dynamics, including altered hydrological cycles and increased risk of drought* [39].

In conclusion, the precipitation trends in Turin over this period **underscore the city's vulnerability to changing climate conditions**. The data reflects a notable decrease in average precipitation, which can have wide-ranging impacts on various aspects of the city's environment and economy. The need for informed decision-making, water management strategies, and climate resilience measures becomes crucial considering these trends [40].

Month	Temperature trend over years (°C)	Month	Precipitation trend over years (mm)
J	0.437	J	-19.455
F	2.134	F	-253.772
M	0.052	M	-668.936
A	-1.170	A	-287.421
M	0.250	M	-352.660
J	3.477	J	-832.642
J	2.299	J	150.672
A	1.934	A	3.365
S	0.944	S	-234.501
O	2.004	O	2.487
N	0.212	N	-1222.762
D	0.231	D	-38.474
Av temp 2010 to 2022	1.067	Av Precipitation 2010 to 2022	-312.842

Table 7 - Comparative example about temperature increase and related precipitation

[38] IPCC Report 2023, "Longer Report", page 49: IPCC_AR6_SYR_LongerReport.pdf

[39] IPCC Report 2023, "Longer Report", page 50: IPCC_AR6_SYR_LongerReport.pdf

[40] IPCC Report 2023, "Longer Report", page 105: IPCC_AR6_SYR_LongerReport.pdf

To support our analyses, the IPCC 2023 reports provide a global overview of the phenomenon of climate change. Below, we present two figures from the previously mentioned IPCC Report 2023. These figures will enhance our analysis in explaining the phenomena of 'Hot extremes,' 'Heavy precipitation,' 'Agricultural and ecological droughts,' and 'Observed impacts and related losses and damages of climate change.'

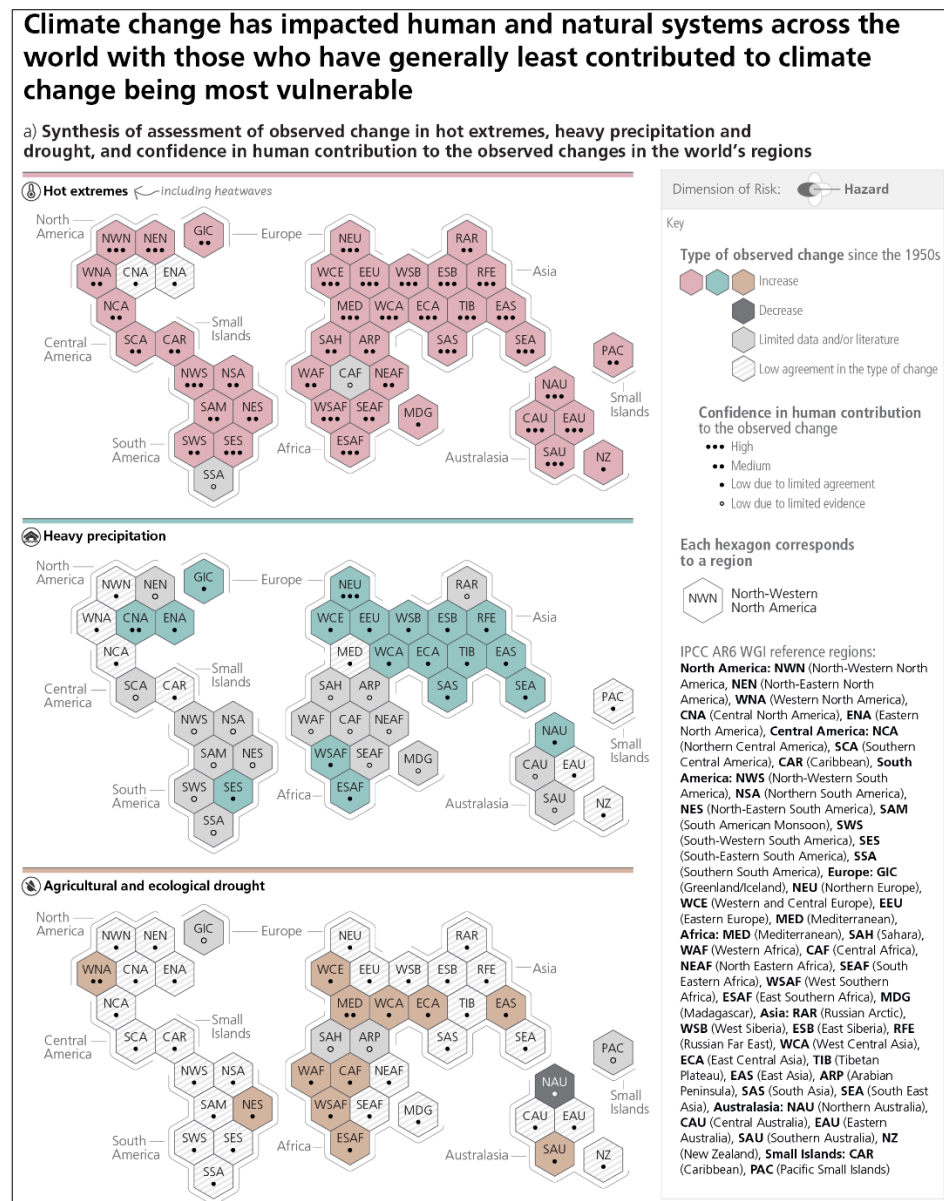
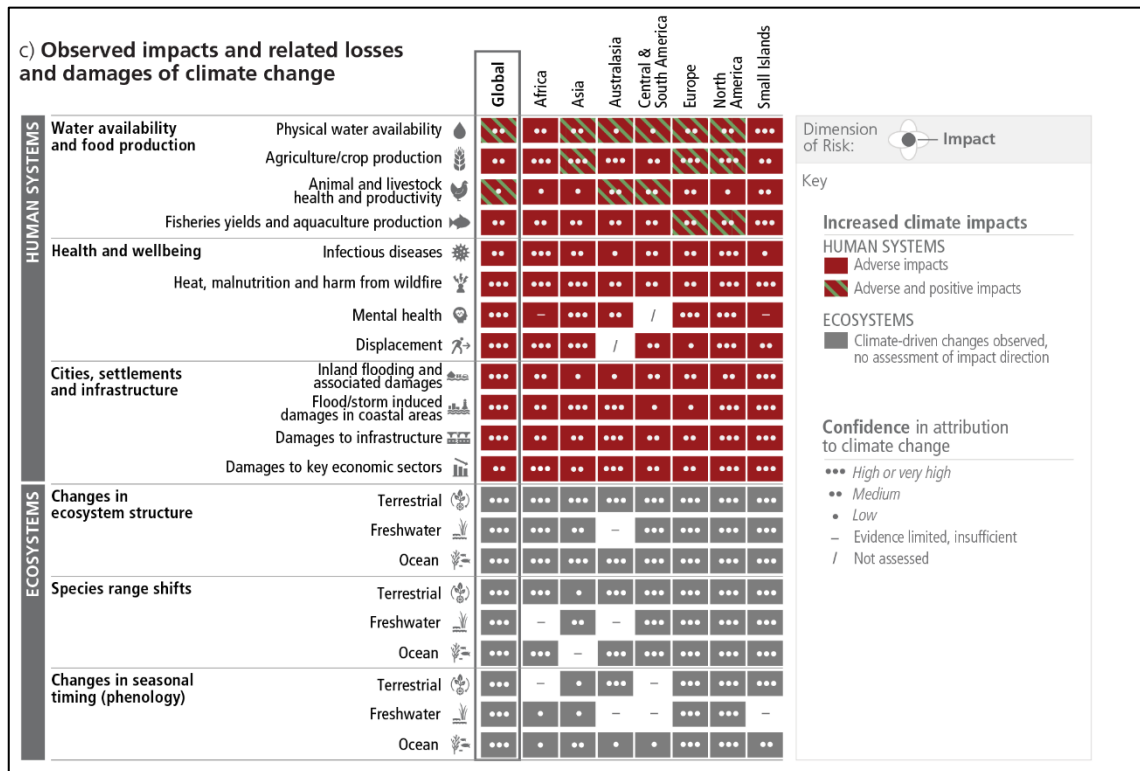


Figure 7 - IPCC Longer Report, 2023 (MED area means Mediterranean area, it also includes Italy) ⁴¹

[41] IPCC Report 2023, "Longer Report", page 48: IPCC_AR6_SYR_LongerReport.pdf

Figure 8 - IPCC Longer Report, 2023 ⁴²

[42] IPCC Report 2023, "Longer Report", page 49: IPCC_AR6_SYR_LongerReport.pdf

2.3 Transforming the Urban Landscape: Large-Scale Interventions in Turin



Figure 9 – Area EX TEKSID captured in the orthophoto of CGR IT2000 Provincia di Torino (year 2000)

▪ **How the TEKSID area was born?**

The EX-AREA TEKSID's origin is intertwined with the history of Ferriere Piemontesi, a significant industrial entity in Turin. It traces its roots back to the French company Vandel & C., which relocated its nail and wire manufacturing facilities from Ferrière-sous-Jougne to Avigliana in 1891. The enterprise initially focused on shoe nails, "punte di Parigi" (Paris points), and wire production. However, by 1896, with the installation of a Martin furnace and a rolling mill, Vandel expanded into steel manufacturing and rod rolling. This evolution continued, and in 1899, after further expansion, the company transformed into Ferriere di Buttigliera Alta e di Avigliana, becoming a joint-stock company. The growing demand for processing from the Piedmontese industry prompted the company to establish a larger facility in Turin in 1906. This led to the birth of Ferriere Piemontesi, characterized by an operational steelwork with multiple furnaces and rolling equipment. The Turin complex commenced operations in 1907 near the Dora Station, boasting multiple Martin Siemens furnaces, electric furnaces, and rolling mills.

During World War I, the Ferriere Piemontesi complex, employing approximately 550 workers, played a role not only in steel production but also in manufacturing steel parts for weaponry. This strategic significance attracted Fiat's interest, and in 1917, Fiat acquired the Piemontese Group, incorporating significant metallurgical companies like Ferriere Piemontesi. This move was regarded as a pivotal business decision during and after the war.

The complex underwent further expansion during and after the war, with Fiat's involvement. Its capabilities expanded beyond metallurgy to encompass machinery, automotive parts, and more. Throughout World War II, Ferriere Piemontesi endured bombings, yet remained a hub of resistance and refuge for partisans. The factory became

a focal point of anti-fascist activities, a hub for distributing clandestine materials, and a site of collaboration for the resistance.

After the war, the factory experienced a resurgence in production, culminating in the establishment of the Teksid company in 1978. Teksid consolidated Fiat's metallurgical and siderurgical operations, becoming a subsidiary of Finsider in 1982 [43].

44	AREA EX FERRIERE FIAT VITALI	AREA EX FERRIERE FIAT VALDOCCO	AREA EX FERRIERE FIAT INGEST
Construction	1920	1917	1939
Expansion	1973	1925	-
Bombardment	-	1942	-
Abandonment	1992	1992	1992
Demolition/Transformation	2001	2005	1995

Table 8 – Ex Area TEKSID divided in FIAT's ex-industrial plant, source: Museo Torino



Figure 10 - Ex ferriere Fiat Valdocco (left), Ex ferriere Fiat Vitali (right)

[43] Torino 1938|45 Luoghi memoria: Torino 1938|45 - La città delle fabbriche. Fiat Ferriere (istoreto.it)

[44] Ex Stabilimento TEKSID, ex Ferriere Fiat Vitali: Ex stabilimento Teksid, ex Ferriere Fiat Vitali - MuseoTorino

[44] Ex Stabilimento TEKSID, ex Ferriere Fiat Valdocco: Ex stabilimento Teksid, ex Ferriere Fiat Valdocco - MuseoTorino

[44] Ex Stabilimento TEKSID, ex Ferriere Fiat Ingest: Ex Stabilimento Teksid, ex Ferriere Fiat Ingest - MuseoTorino

With the passage of time and changing economic dynamics, many of these industrial areas underwent ***transformation and decline***. This shift led to a unique challenge: how to preserve the industrial heritage while adapting to new socio-economic and environmental realities. Here below are shown two images about one of the renewals, the source images are “**Immagini del cambiamento: [PD04 \(polito.it\)](https://polito.it)**”

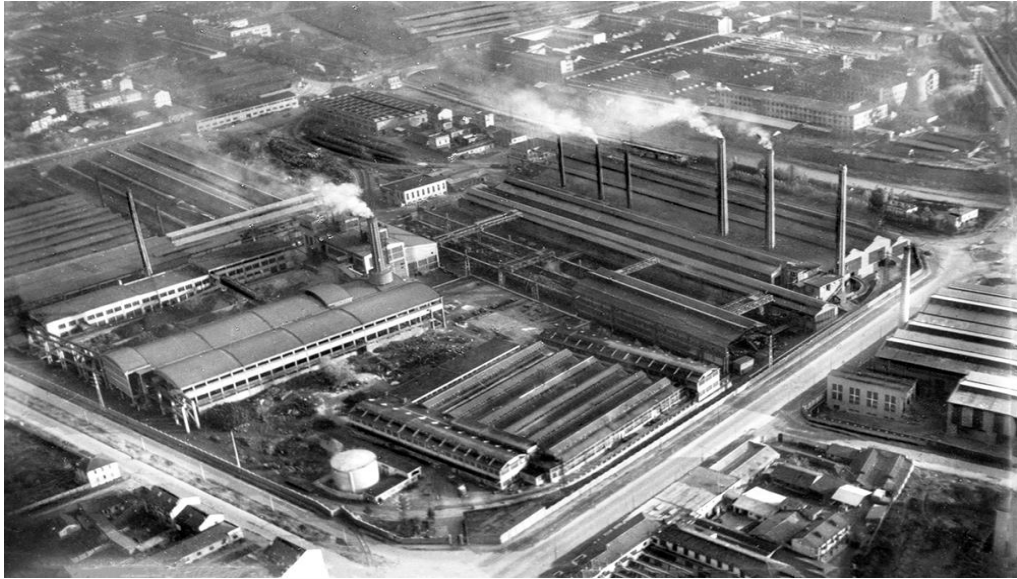


Figure 11 - Teksid area and other factories, from Verolengo Street (view looking south), 1970

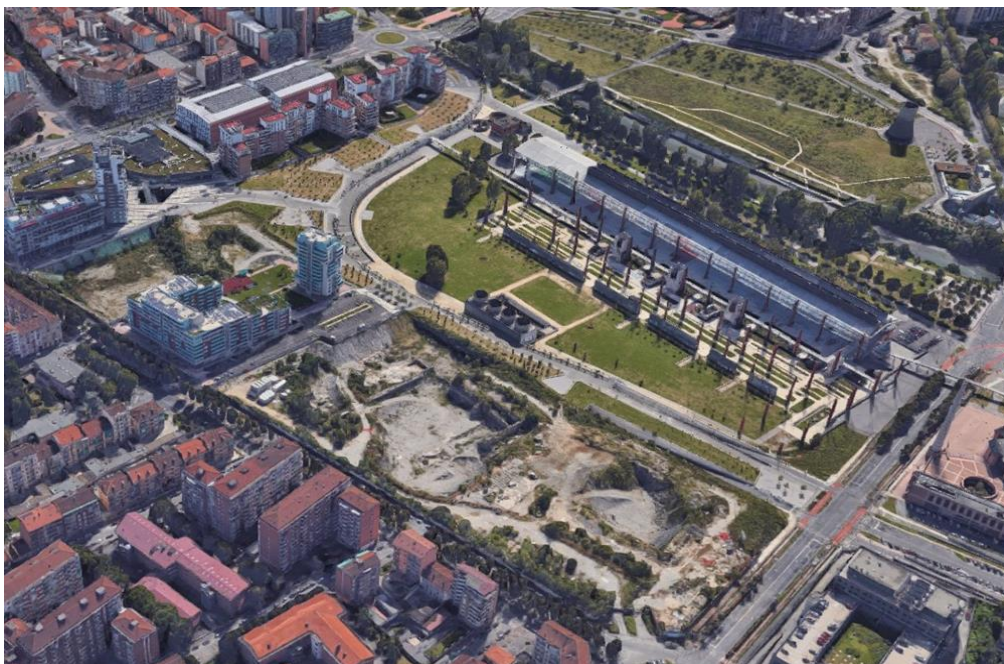


Figure 12 - Dora Park, condominiums, urban empty residue, 2018

3. Pre-modelling: Assessment of Environmental Variables

3.1 Data collection and processing

In the subsequent step, we conducted the calculation of indexes utilizing Landsat 8 images. The acquisition of the data was facilitated by a QGIS tool called "STAC API Browser Plugin," which allows direct downloading of the required bands into the software. This streamlined process enabled seamless access to the necessary materials for index calculation within the QGIS environment.

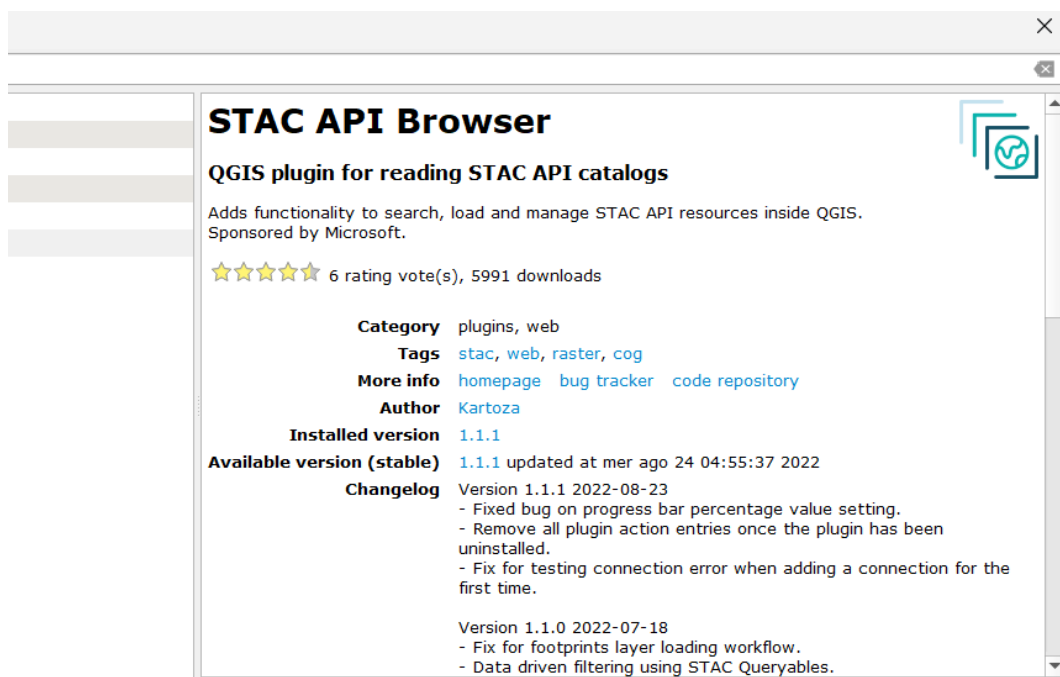


Figure 13 - STAC API Browser (Qgis Plugin)

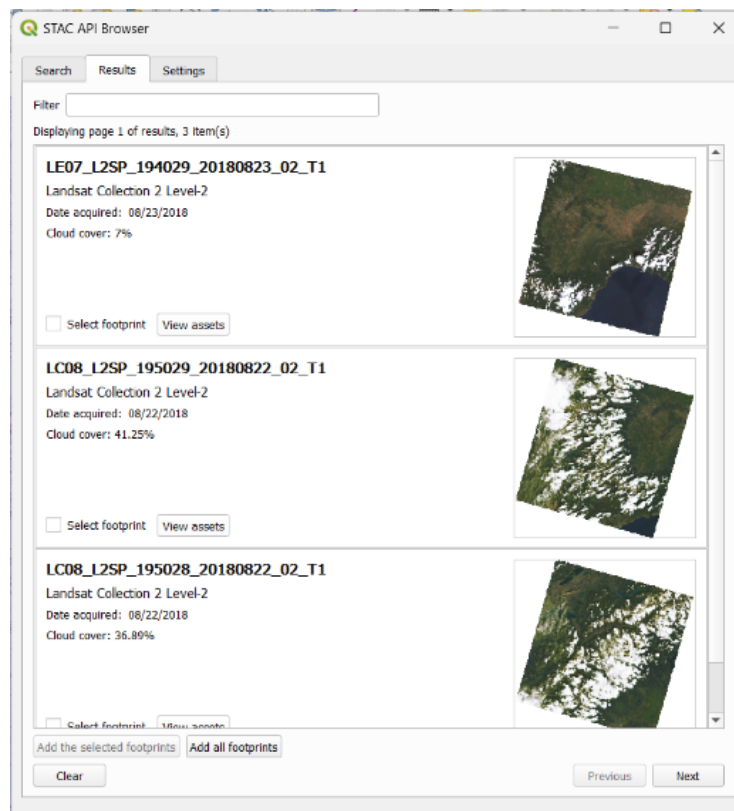
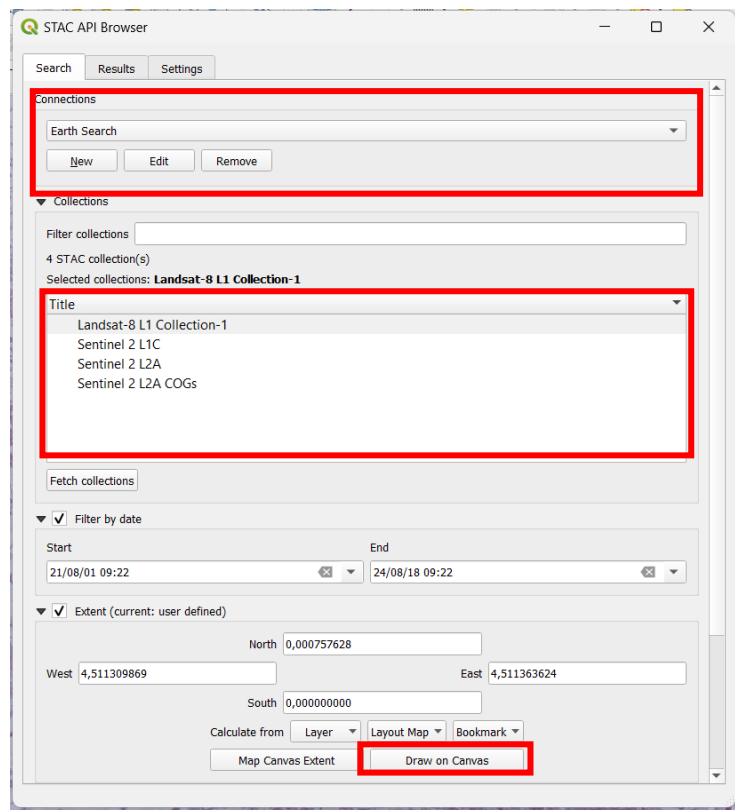


Figure 14 - Selection of the images

For the **preprocessing of Landsat images**, the “**Semi-Automatic Classification Plugin (SCP) tool**” within the QGIS software has been utilized. SCP offers a comprehensive set of functions for processing remote sensing data.

This tool assists in tasks such as **radiometric calibration, atmospheric correction, geometric correction, and cloud masking**. By integrating SCP into QGIS, we can streamline the preprocessing workflow, ensuring accurate and well-prepared data for subsequent analyses.

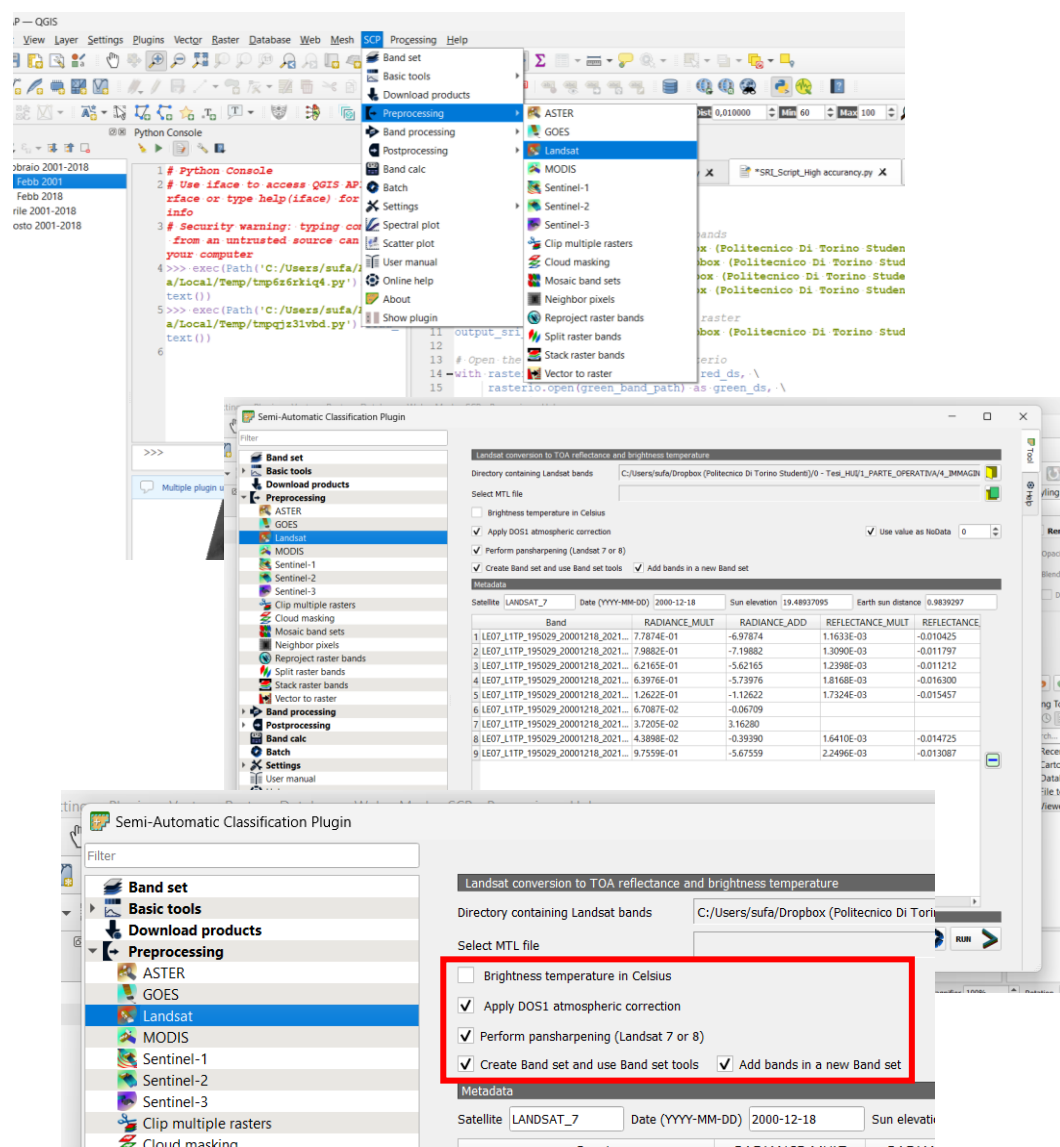


Figure 15 - SCP Preprocessing for satellite images (Qgis)

The subsequent steps involve calculating various indexes using the raster bands extracted from these images. To streamline and expedite this process, we harnessed the power of the **Python Console in QGIS**. By utilizing the Python programming language, we were able to **write custom codes that automated the index calculation, replacing the need for the "raster calculator tool"** that is typically used in a manual workflow.

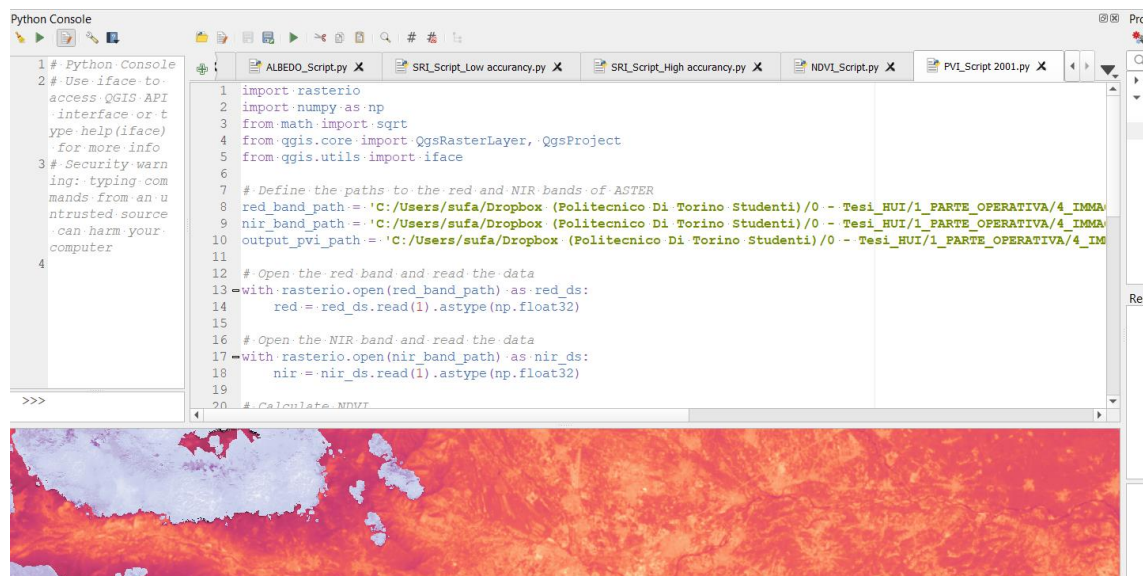


Figure 16 - Python Console (QGis)

The advantage of using Python scripts in the Python Console lies in the **automation and efficiency** it brings to the process. Rather than manually executing the calculations for each image and index using the graphical user interface, the Python scripts allowed us to execute the necessary commands automatically.

To achieve this automation, we adapted the codes in a way that made it easy to **modify just the band paths for each image**. Instead of hardcoding specific bands for individual images, we made the scripts flexible by allowing the input of different band paths or filenames as variables. This way, the same script could be reused for multiple images, reducing the need for repetitive manual tasks.

3.2 Candidate variables

Candidate data for modelling refers to the subset of data that is selected or considered as **potential input for the modelling phase** of a data analysis project. It is the data that is believed to have a high likelihood of being useful and informative for building a model to address a particular research question.

Category	Variables
LAND COVER VARIABLES	Use of soil
	Normalized difference vegetation index (NDVI)
	Proportion Vegetation index (PVI)
	Normal difference water index (NDWI)
	Normal difference moisture index (NDMI)
	Albedo
	Emissivity
	SRI (Solar reflectance index)
URBAN GEOMORPHOLOGY	Buildings - geometry and human activities - (2018)
	Sky view factor
	Normalized Difference Built-Up Index (2001)
	S/V ratio
Weather stations	Relative humidity
	Wind speed
	Wind direction
	Incoming solar irradiation
	Air Temperature

Table 9 - Candidate variables for modelling

▪ **Satellite images for the calculation of the variables – LANDSAT 8 [2018]:**

- **SUMMER:**

LANDSAT_PRODUCT_ID	LC08_L2SP_195029_20180822_2020083 1 02 T1
DATE_ACQUIRED	2018-02-11
SCENE_CENTER_TIME	10:16:47.63
CLOUD_COVER	41.25
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

- **WINTER:**

LANDSAT_PRODUCT_ID	LC08_L2SP_195029_20180211_2020090 2 02 T1
DATE_ACQUIRED	2018-08-22
SCENE_CENTER_TIME	10:17:08.07
CLOUD_COVER	22.91
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

- **MID SEASON:**

LANDSAT_PRODUCT_ID	LC08_L2SP_195029_20180416_2020090 1 02 T1
DATE_ACQUIRED	2018-04-16
SCENE_CENTER_TIME	10:16:36.47
CLOUD_COVER	28.46
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

▪ **Satellite images for the calculation of the variables – LANDSAT 7 [2001]:**

- **SUMMER:**

LANDSAT_PRODUCT_ID	LE07_L1TP_194029_20010824_2020091 7 02 T1
DATE_ACQUIRED	2001-08-24
SCENE_CENTER_TIME	09:59:54.42
CLOUD_COVER	4.00
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

- **WINTER:**

LANDSAT_PRODUCT_ID	LE07_L1TP_195029_20001218_2021112 2_02_T1
DATE_ACQUIRED	2000-12-18
SCENE_CENTER_TIME	10:07:43.39
CLOUD_COVER	5.00
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

- **MID SEASON:**

LANDSAT_PRODUCT_ID	LE07_L1TP_195029_20010527_2020091 7_02_T1
DATE_ACQUIRED	2001-05-27
SCENE_CENTER_TIME	10:07:06.43
CLOUD_COVER	10.00
GRID_CELL_SIZE_REFLECTIVE / THERMAL	30

■ **Day temperature [°C]**

Seasonality	2001		2018	
	Landsat 7		Landsat 8	
	gg	°C	gg	°C
Summer	24-ago	26,1	22-ago	26,2
Mid-Season	27-mag	23,2	16-apr	17,9
Winter	18-dic	2,1	11-feb	3,8

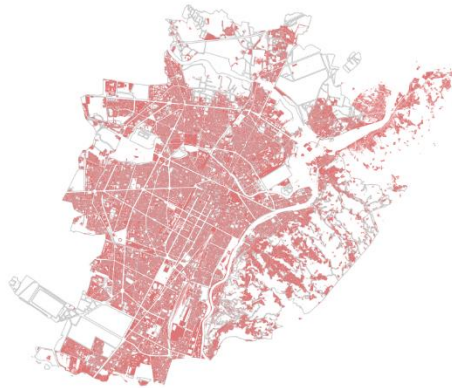
Satellite imagery from 2001 and 2018 was selected for analysis based on comparable temperature parameters during each respective season. The process involved careful curation of **satellite images that exhibited similar temperature patterns across corresponding seasons in both years**. This methodology aimed to provide a refined basis for evaluating climate-related changes and their consequences on the environment. The resulting alignment of images with similar temperatures facilitated the examination of the evolving relationship between temperature variations and corresponding shifts in land use and climatic conditions.

3.2.1 Land cover variables

The data production method used for the classification of the land uses is an experimental application of a new European model proposed by EAGLE. It involves using the regional geo-topographic database (BDTRE) as a foundation and enriching it with thematic information such as the Forest Map and Regional Landscape Plan. Additionally, new information derived from satellite imagery (Copernicus and Telerilevamento Piemonte) is integrated and correlated with cadastral mapping parcels.

This approach aims to enhance land monitoring by utilizing a combination of existing databases and satellite imagery to gather comprehensive and up-to-date information about land cover and related characteristics.

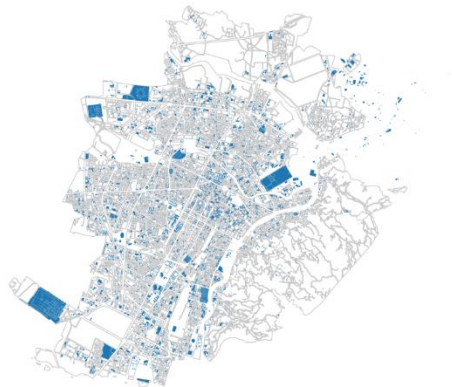
- **Use of soil:** The composition of land cover, such as concrete, asphalt, vegetation, or open spaces, significantly influences the urban heat island effect. Evaluating land uses allows for a better understanding of how different areas contribute to heat retention or mitigation.



f) Residential areas



d) Industrial areas



c) Commercial and services areas



e) Logistical spaces



a) Resource extraction and waste disposal areas



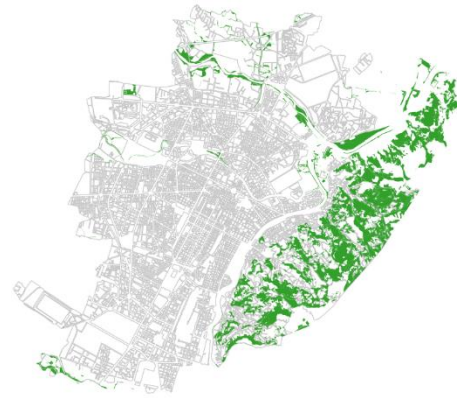
b) Urban green areas

Soil Coverage

■ Residential areas	■ Urban green areas
■ Industrial areas	■ Agricultural areas
■ Commercial and services areas	■ Forest areas
■ Logistical spaces	■ Grazing areas
■ Resource extraction and waste disposal areas	■ Aquatic areas



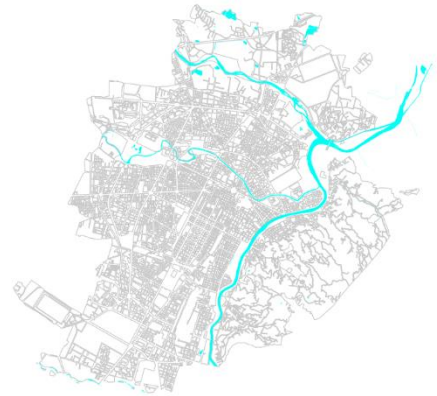
i) Agricultural areas



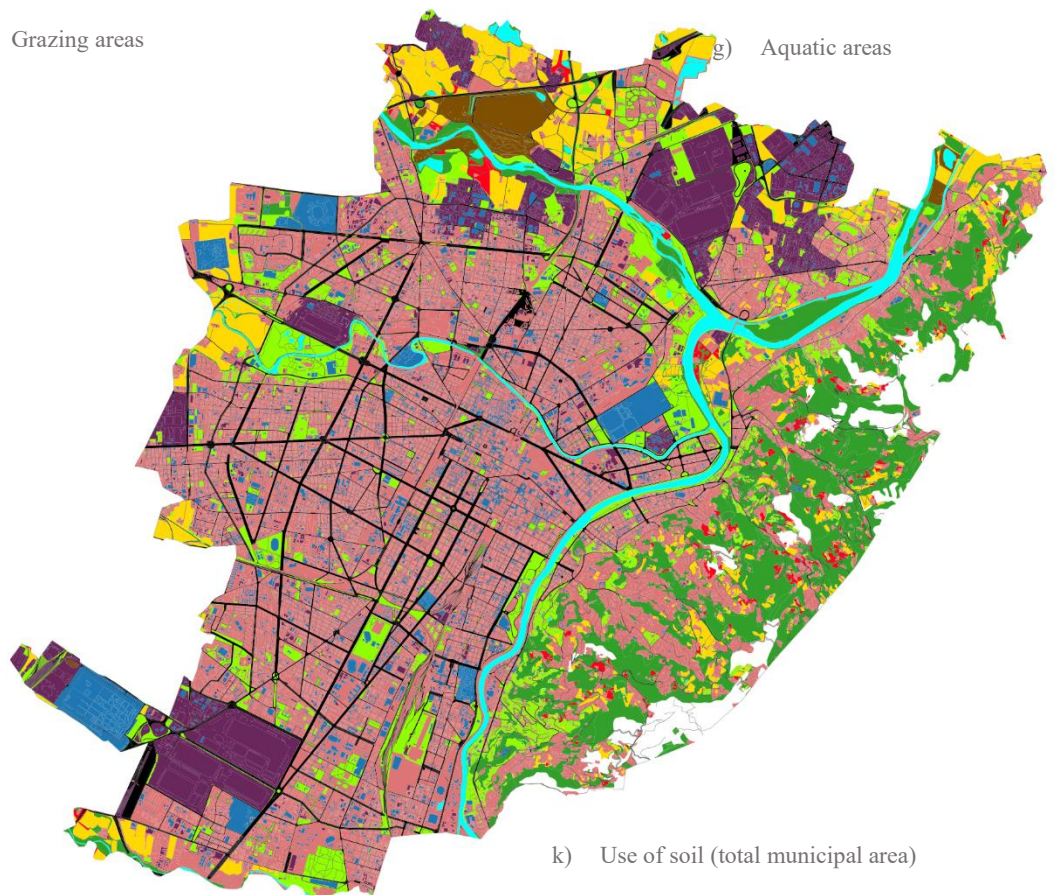
j) Forest areas



h) Grazing areas



g) Aquatic areas



k) Use of soil (total municipal area)

▪ **Fractional roof cover:**

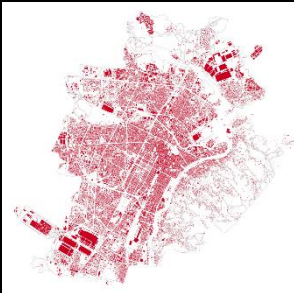
Asphalt Area	Total Metropolitan Area	Asphalt Area	Total Metropolitan Area	FRACTIONAL ROOF COVER 
[m ²]	[m ²]	[ha]	[ha]	
25534029,9	142653650,2	2553,40	14265,37	18%

Table 10 - Fractional roof cover [%] – Elaboration on data by Geoportale Comune di Torino

▪ **Fractional soil cover:**

Category	Total area [m2]	Total area [ha]	Fractional soil cover %
Residential areas	53741782,27	5374,178227	37,67%
Industrial areas	13069847,23	1306,984723	9,16%
Commercial & services areas	9450942,313	945,0942313	6,63%
Logistical spaces	16341443,45	1634,144345	11,46%
Resource extraction and waste disposal areas	1679569,661	167,9569661	1,18%
Urban green areas	13390711,83	1339,071183	9,39%
Agricultural areas	8936226,606	893,6226606	6,26%
Forest areas	20922900,9	2092,29009	14,67%
Grazing areas	1214062,255	121,4062255	0,85%
Aquatic areas	3906163,707	390,6163707	2,74%
TOT	142653650,2	14265,36502	100,00%

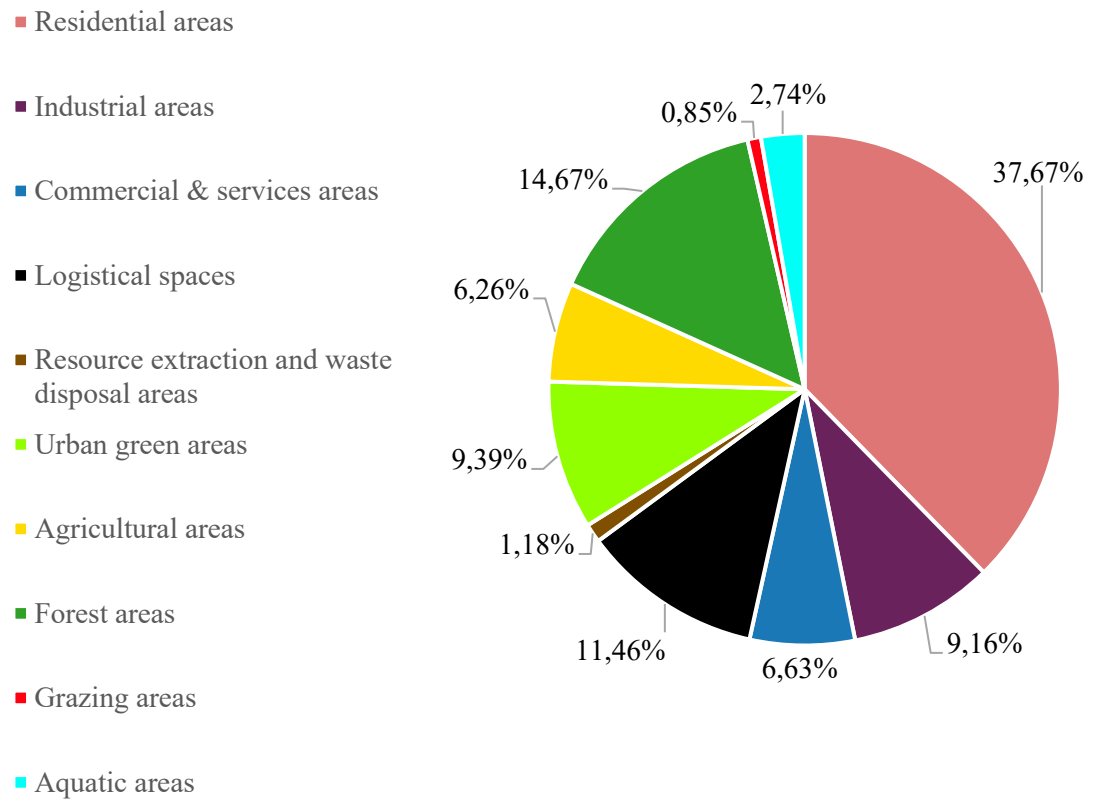


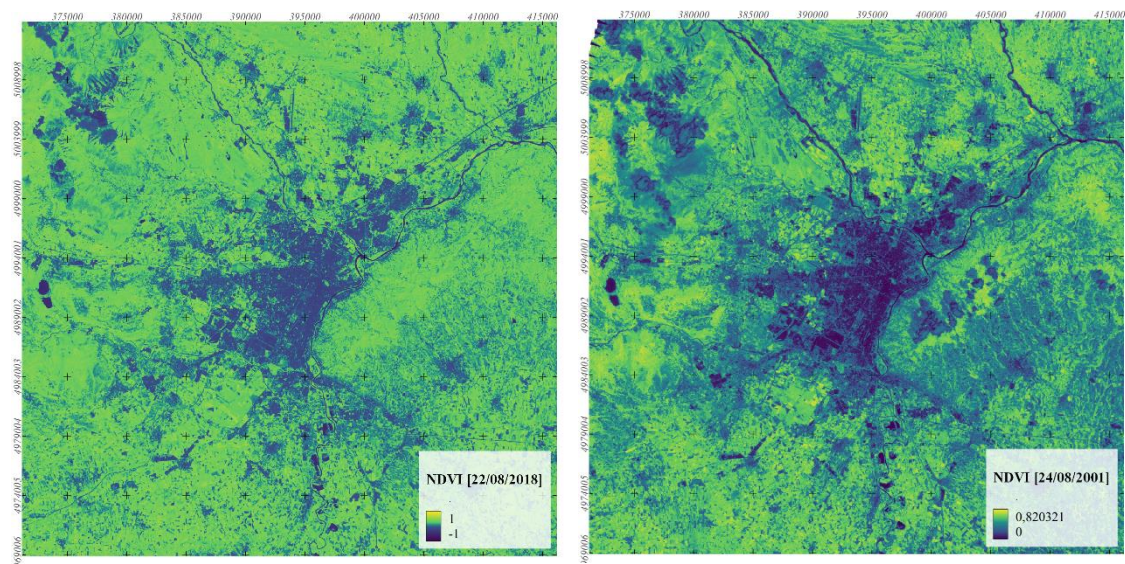
Figure 17 - Fractional soil cover [%] - Elaboration on data by Copernicus

▪ Normalized difference vegetation index (NDVI):

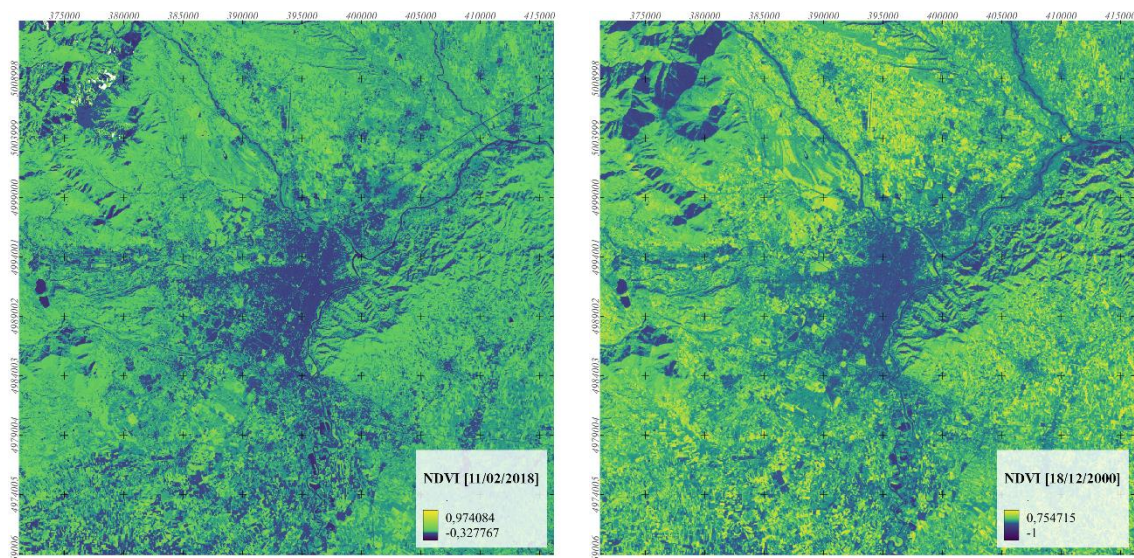
The NDVI is a vegetation index derived from satellite imagery that **measures the density of green vegetation**. It compares the reflectance values of the near-infrared (NIR) and red bands of the electromagnetic spectrum, using the calculator in QGIS.

$$\text{NDVI} = \frac{\text{NIR} - \text{VIS}}{\text{NIR} + \text{VIS}} = \frac{B5 - B4}{B5 + B4}$$

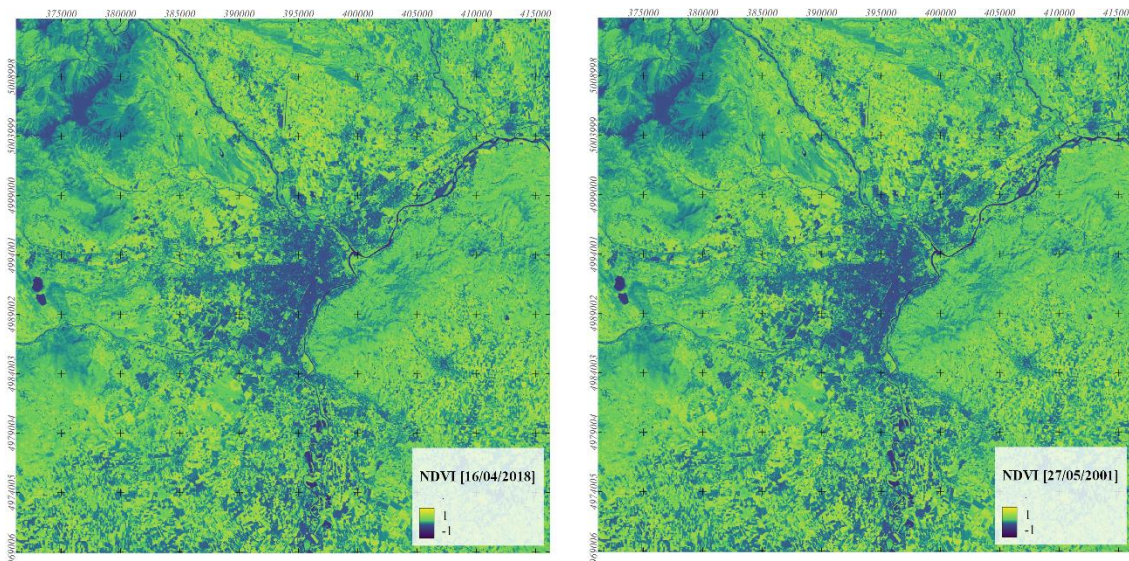
PYTHON SCRIPT OF THE AUTOMATIZED PROCESS



NDVI calculation [QGIS] – SUMMER 2018 & 2001



NDVI calculation [QGIS] – WINTER 2018 & 2000



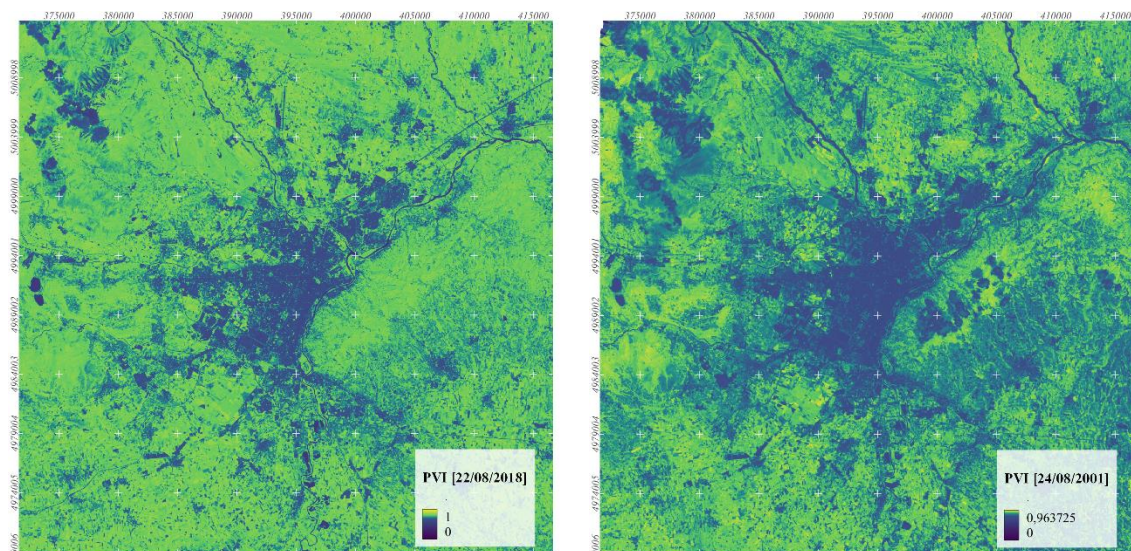
NDVI calculation [QGis] – MID SEASON 2018 & 2001

▪ **Proportion vegetation index (PVI):**

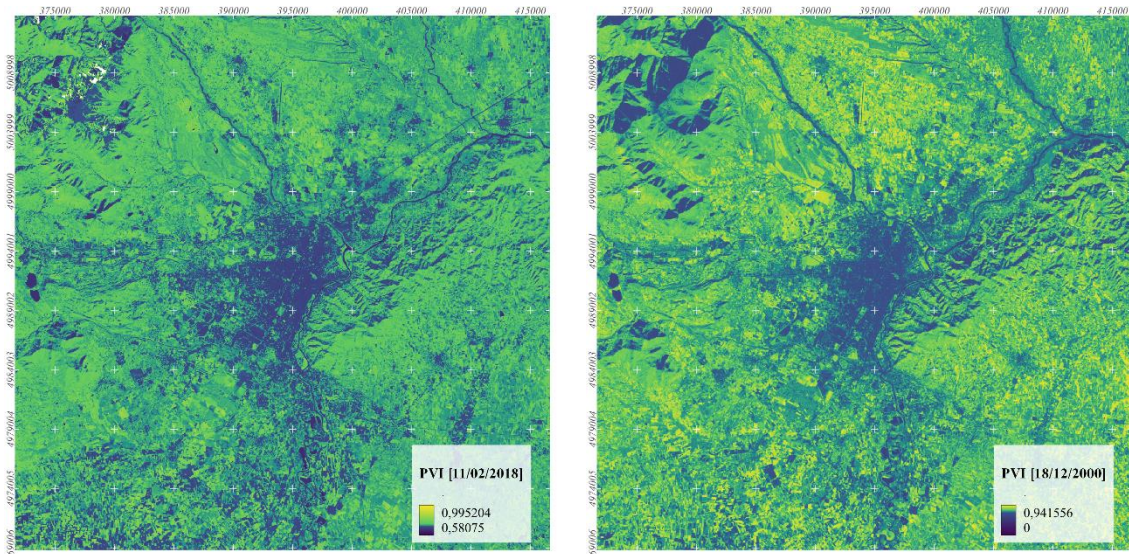
A vegetation index that **quantifies the proportion of vegetation cover within a specific area** (0 to 1 index). It is calculated by comparing the reflectance values of vegetation and non-vegetation elements in remote sensing imagery, using QGis.

$$PVI = \left[\frac{NDVI - NDV_{min}}{NDVI + NDV_{max}} \right]^2$$

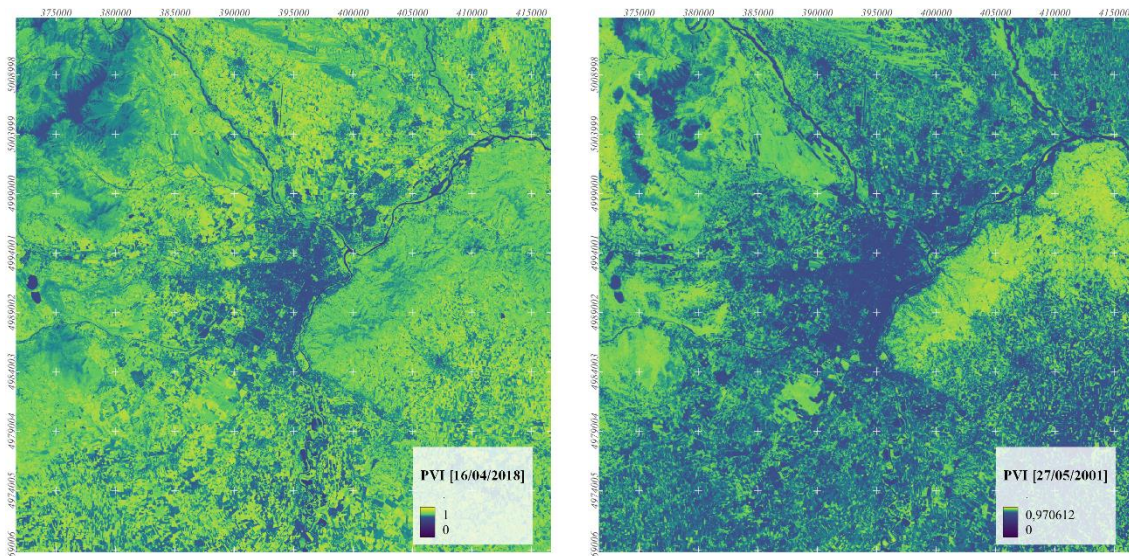
PYTHON SCRIPT OF THE AUTOMATIZED PROCESS



PVI Calculation [QGis] – SUMMER 2018 & 2001



PVI Calculation [QGis] – WINTER 2018 & 2000



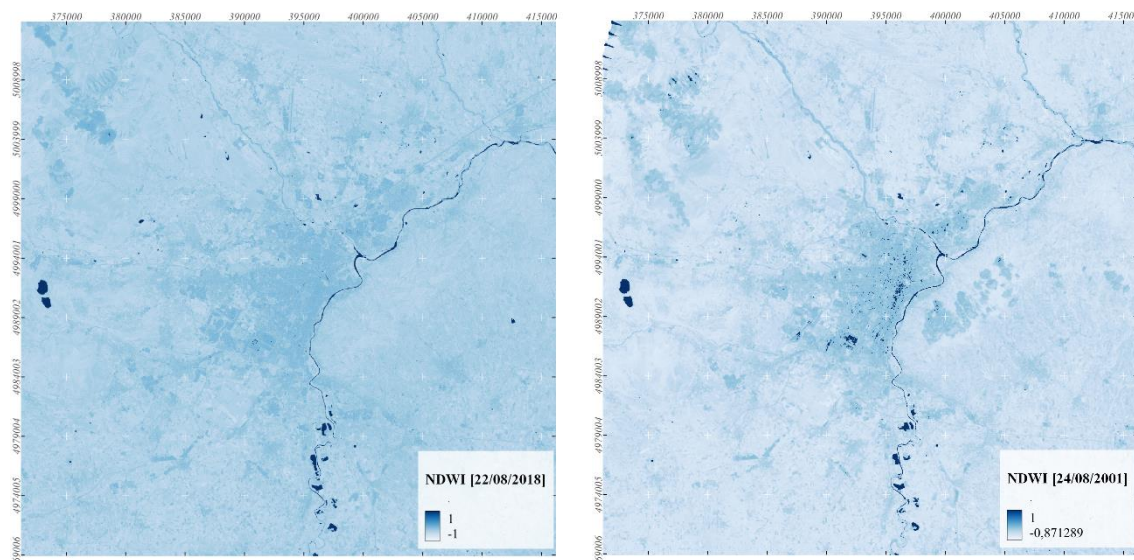
PVI Calculation [QGis] – MID SEASON 2018 & 2001

▪ **Normal difference water index (NDWI)**

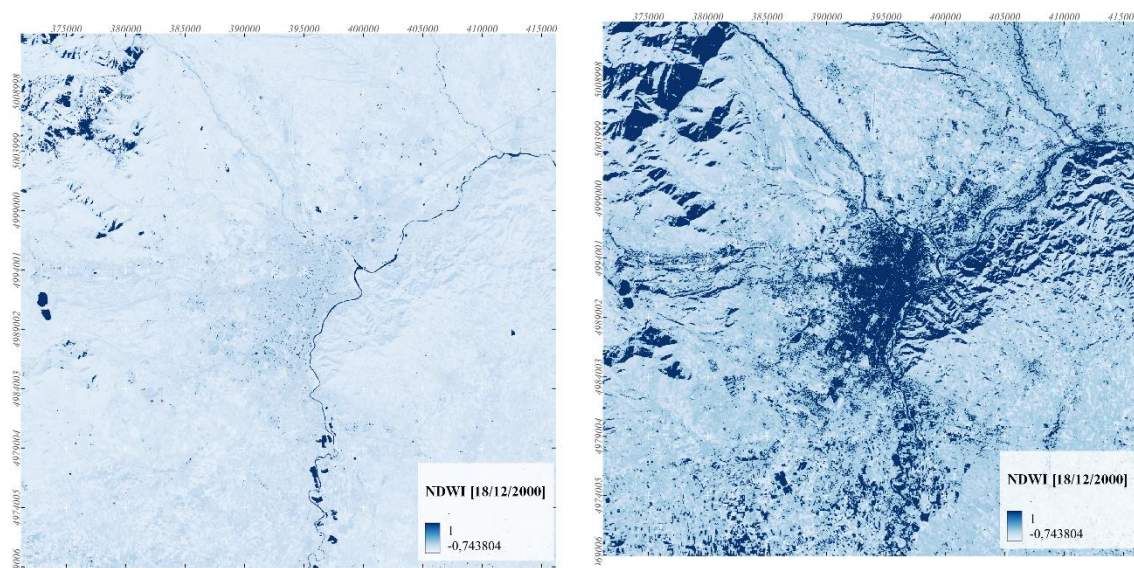
NDWI highlights the presence and extent of water within a given area. We need to highlight that, the presence of noise in the image, does not allow the perfect distinction of the waterbodies.

$$\text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}} = \frac{B3 - B5}{B3 + B5}$$

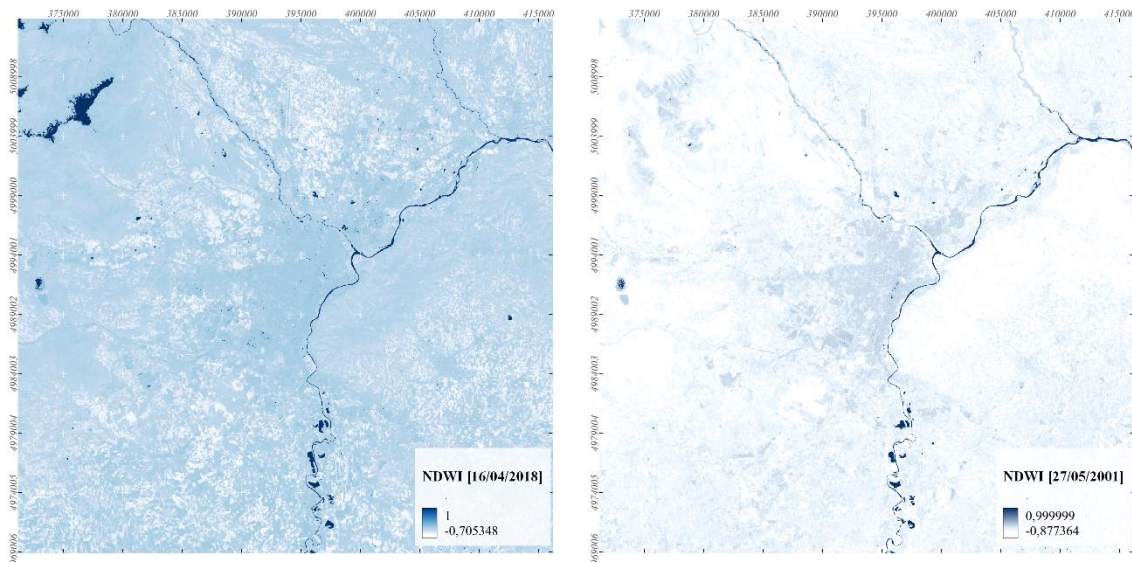
PYTHON SCRIPT OF THE AUTOMATIZED PROCESS



NDWI Calculation [QGIS] – SUMMER 2018 & 2001



NDWI Calculation [QGIS] – WINTER 2018 & 2000



NDWI Calculation [QGIS] – MID SEASON 2018 & 2001

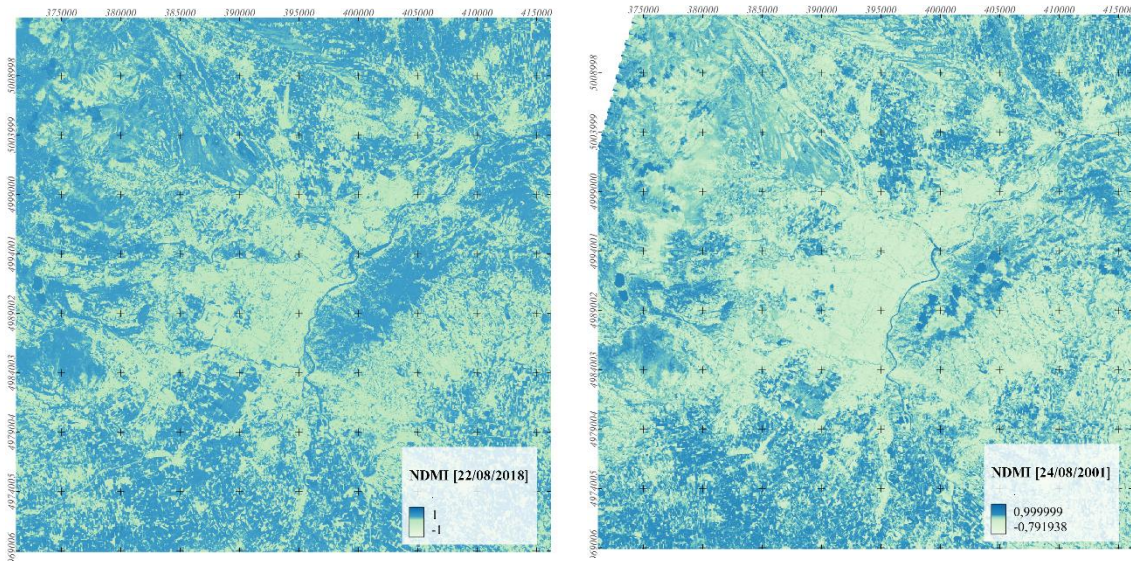
- **Normal difference moisture index (NDMI)**

It measures the moisture content and water stress of vegetation. By comparing the near-infrared (NIR) and shortwave infrared (SWIR) bands of the satellite data, NDMI indicates the relative moisture content of vegetation. Higher NDMI values indicate healthier vegetation with higher moisture content, while lower values suggest stressed vegetation with lower moisture content.

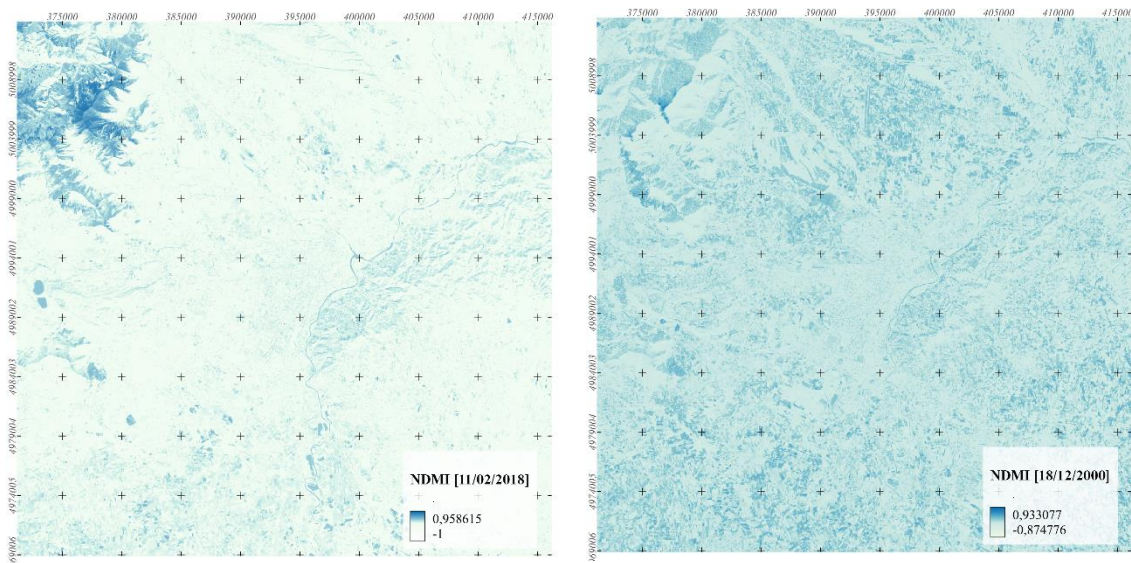
$$\text{NDMI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} = \frac{B5 - B6}{B5 + B6}$$

[PYTHON SCRIPT OF THE AUTOMATIZED PROCESS](#)

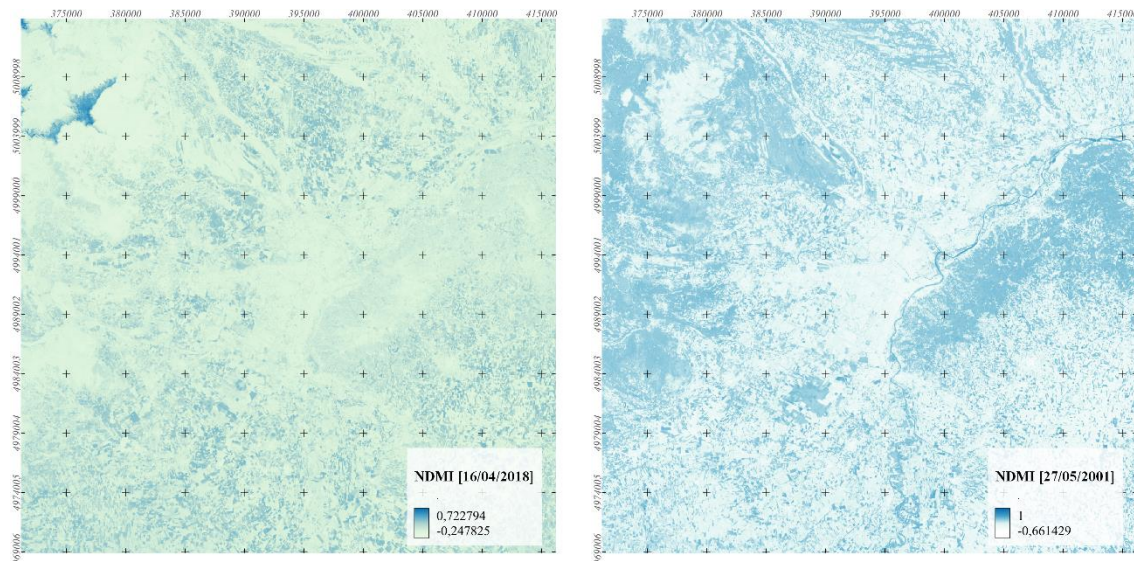
*Analysing Urban Heat Island Dynamics in Turin: A Machine Learning-based Investigation of Climate
Changes and Intervention Effects | A. Scalise, X. Sufa
Supervisor Professor G. Mutani
Politecnico di Torino, 2023*



NDMI Calculation [QGIS] - SUMMER 2018 & 2001



NDMI Calculation [QGIS] - WINTER 2018 & 2000



NDMI Calculation [QGis] –MID SEASON 2018 & 2001

NDMI	CODING
-1 - -0,8	Bare soil
-0,8 - -0,6	Almost absent canopy cover
-0,6 - -0,4	Very low canopy cover
-0,4 - -0,2	Low canopy cover
-0,2 - 0	Mid low canopy cover
1 - 0,2	Average canopy cover
0,2 - 0,4	Mid high canopy cover
0,4 - 0,6	High canopy cover
0,6 - 0,8	Very high canopy cover
0,8 - 1	Total canopy cover

Table 11 - NDMI Coding

▪ Albedo

Albedo is expressed as a percentage or a value between 0 and 1, with 0 representing a perfectly absorbing surface (no reflection) and 1 representing a perfectly reflecting surface (complete reflection).

PYTHON SCRIPT OF THE AUTOMATIZED PROCESS

$$\text{Albedo} = \frac{\rho_{NIR} + \rho_{SWIR}}{2 * (1 - \rho_{Red})}$$

ρ NIR represents the reflectance in the near-infrared band (band 5), ρ SWIR represents the reflectance in the shortwave infrared band (band 7), and ρ Red represents the reflectance in the red band (band 4).

$$\text{Reflectance } \rho = \text{DN} * \text{scale factor} + \text{offset}$$

N.B. The scale factor and offset values can be found in the metadata associated with the Landsat 8 image.

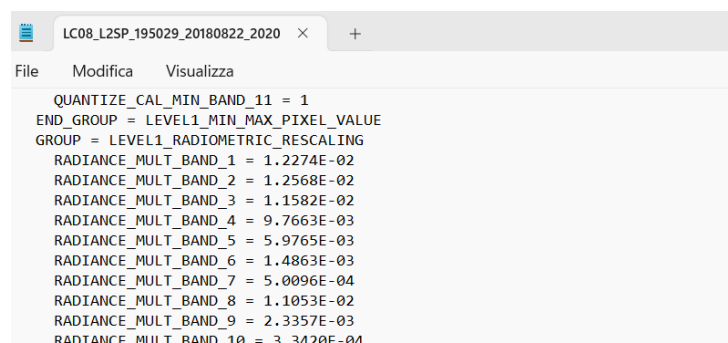
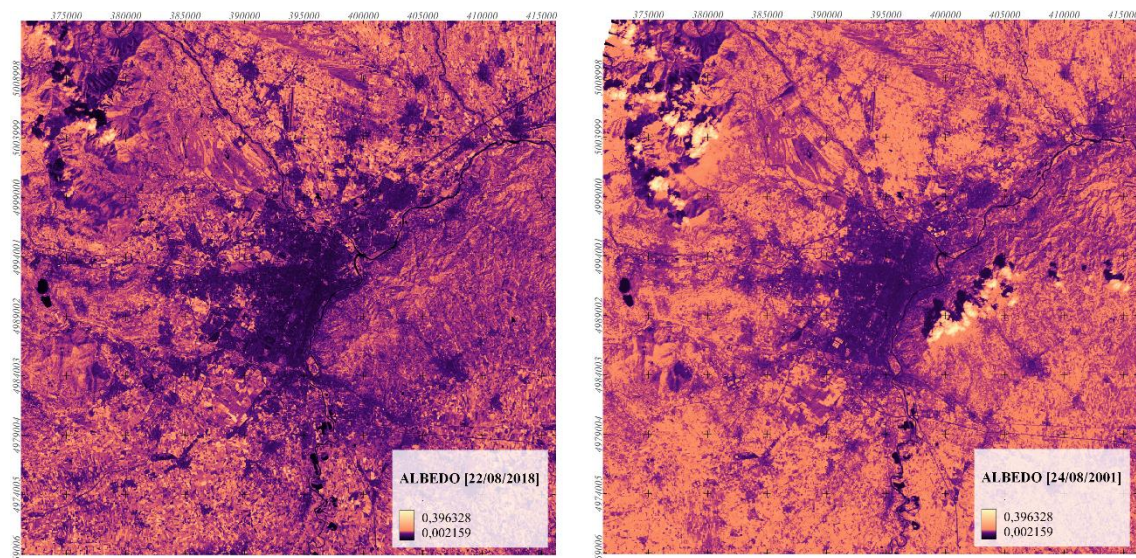
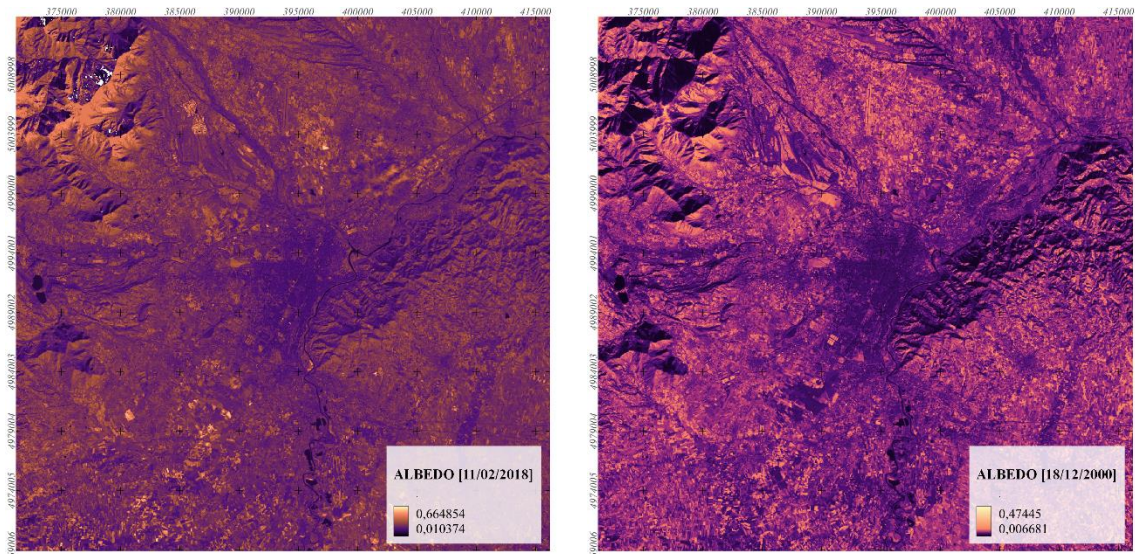


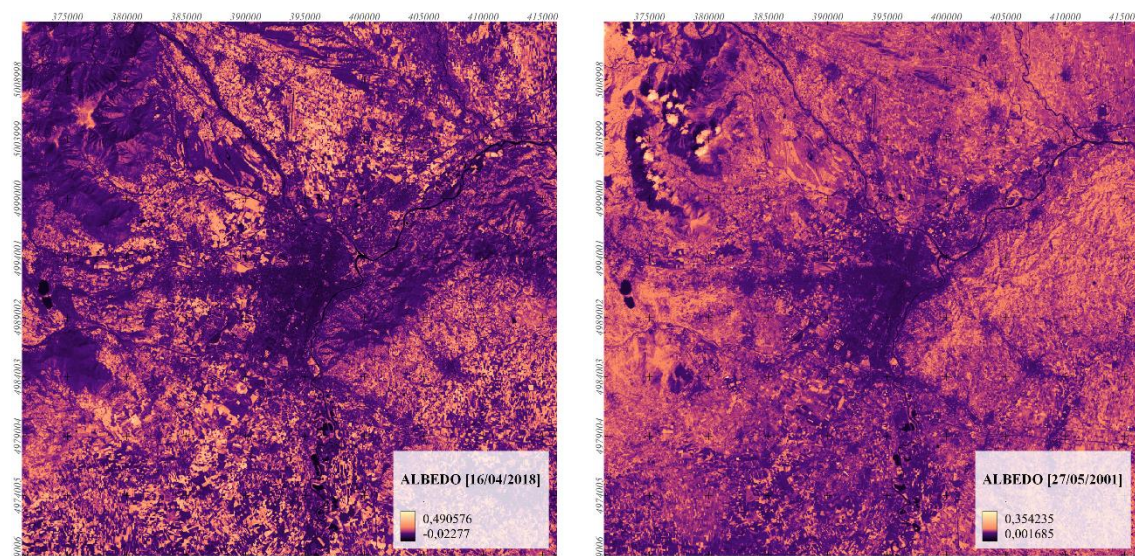
Figure 18 – Scale factor by Landsat 8's metadata



Albedo calculation [QGIS] - SUMMER 2018 & 2001



Albedo calculation [QGIS] - WINTER 2018 & 2000



Albedo calculation [QGIS] - MID SEASON 2018 & 2001

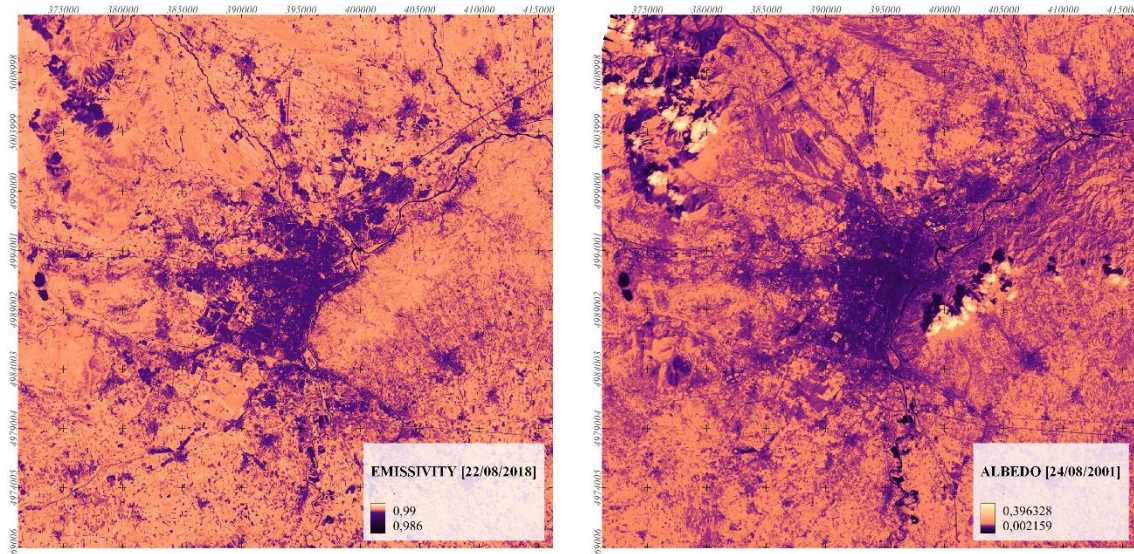
- **Emissivity**

Emissivity refers to the **ability of a surface to emit thermal radiation**. It is a measure of how efficiently a surface radiates heat compared to an ideal blackbody radiator.

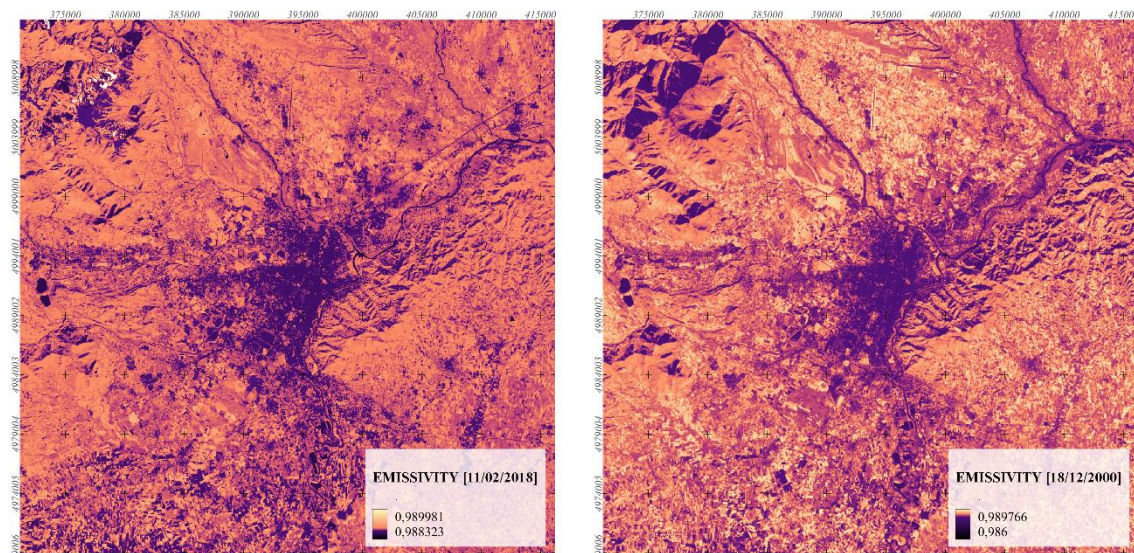
$$E = (0,004 \cdot PVI) + 0,986$$

[PYTHON SCRIPT OF THE AUTOMATIZED PROCESS](#)

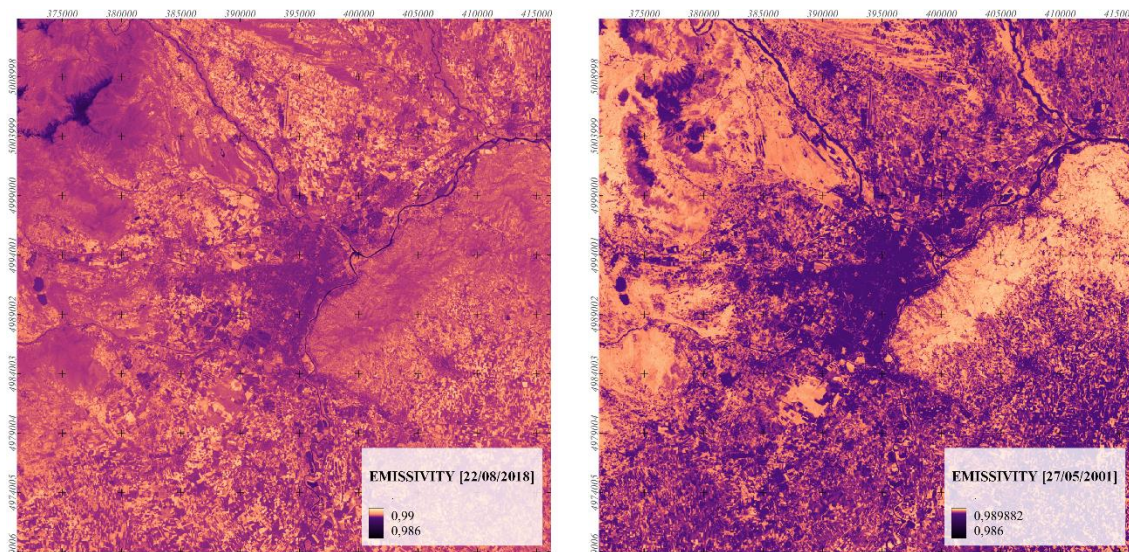
*Analysing Urban Heat Island Dynamics in Turin: A Machine Learning-based Investigation of Climate
Changes and Intervention Effects | A. Scalise, X. Sufa
Supervisor Professor G. Mutani
Politecnico di Torino, 2023*



Emissivity calculation [QGis] – SUMMER 2018 & 2001



Emissivity calculation [QGis] – WINTER 2018 & 2000



Emissivity calculation [QGis] – MID SEASON 2018 & 2001

▪ Solar reflectance index (SRI)

The Solar Reflectance Index (SRI) is a measure of a **material's ability to reflect solar heat** (solar reflectance) **and emit absorbed heat** (thermal emittance). It is used to evaluate the potential of a surface to stay cooler in the sun by reducing the amount of solar radiation it absorbs. The SRI is expressed on a scale from 0 to 100, where 0 represents a surface that absorbs all solar radiation and heats up significantly, and 100 represents a surface that reflects all solar radiation and remains cool. Higher SRI values indicate a material's better ability to reflect sunlight and lower its surface temperature, reducing the heat island effect in urban areas and lowering cooling demands in buildings.

[PYTHON SCRIPT OF THE AUTOMATIZED PROCESS](#)

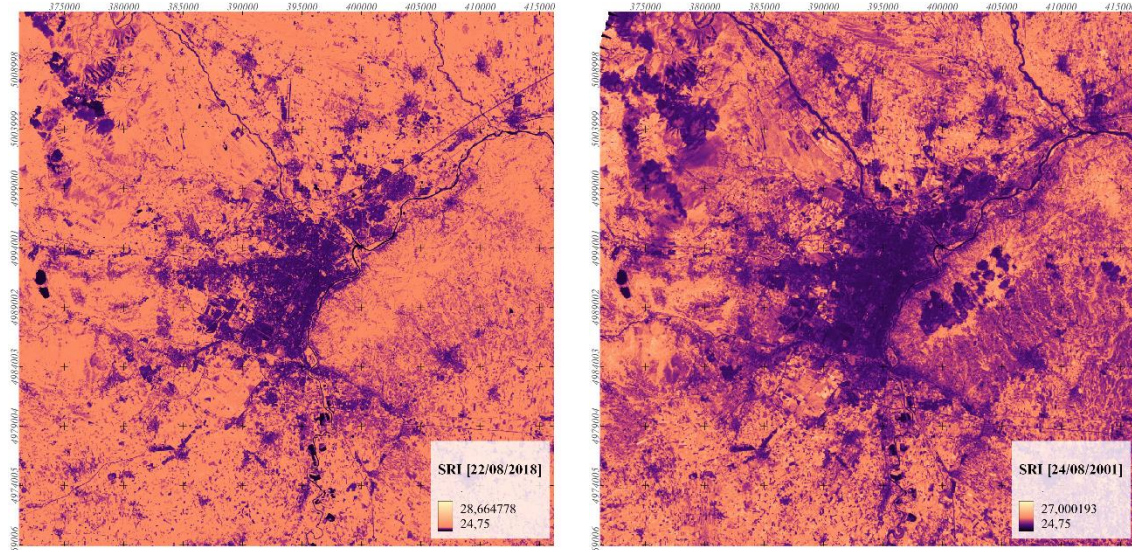
$$\text{SRI} = \left[\frac{\text{SR} - \text{SR}_{\min}}{\text{SR}_{\max} - \text{SR}_{\min}} \right] * 100$$

In this formula:

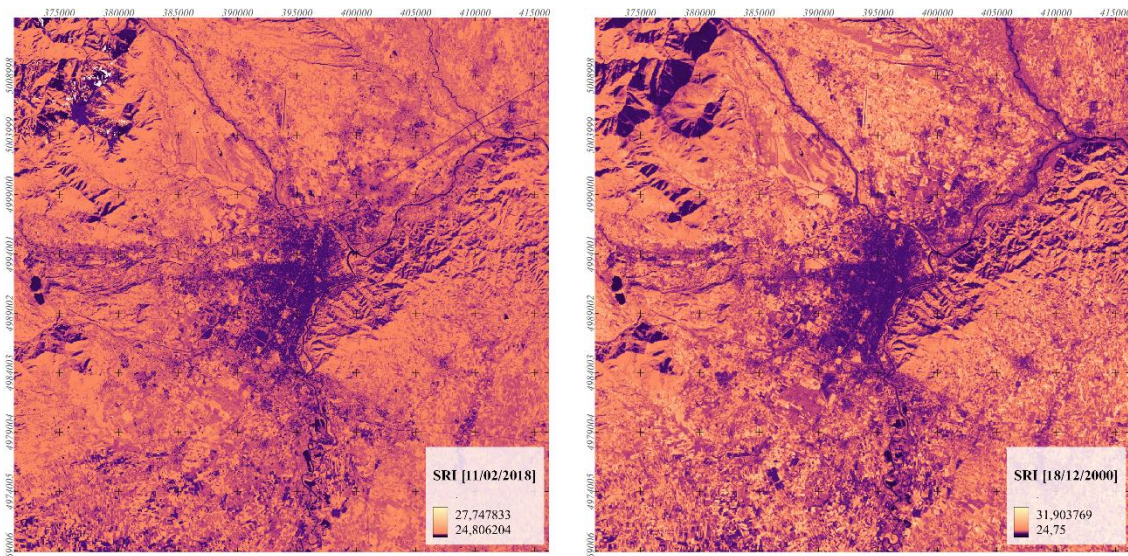
SR represents the solar reflectance of the surface being analyzed.

SR_min is the minimum solar reflectance value for a standard black surface (typically 0.05).

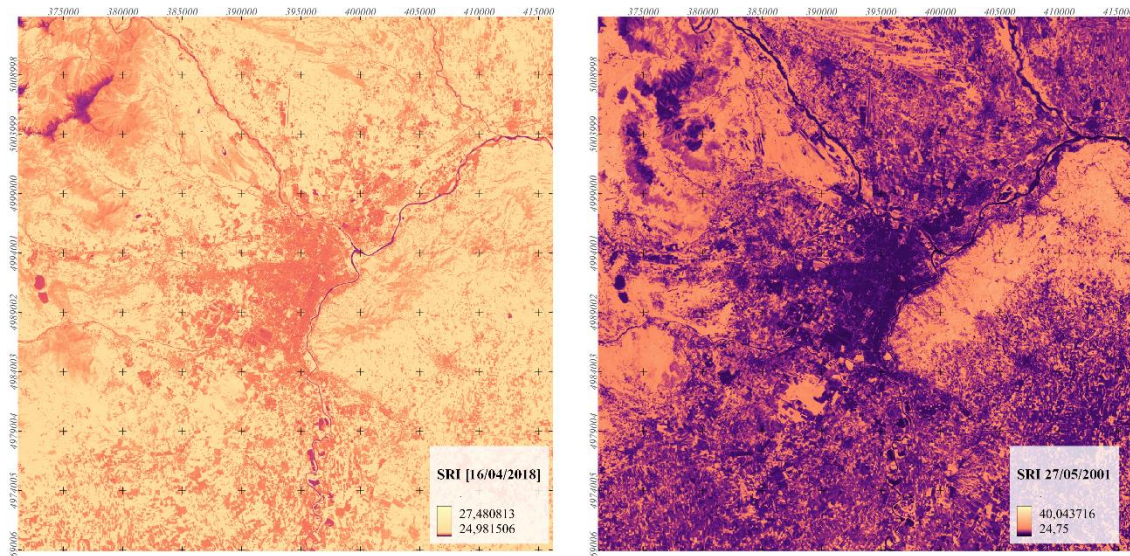
SR_max is the maximum solar reflectance value for a standard white surface (typically 0.80).



SRI calculation [QGIS] – SUMMER 2018 & 2001



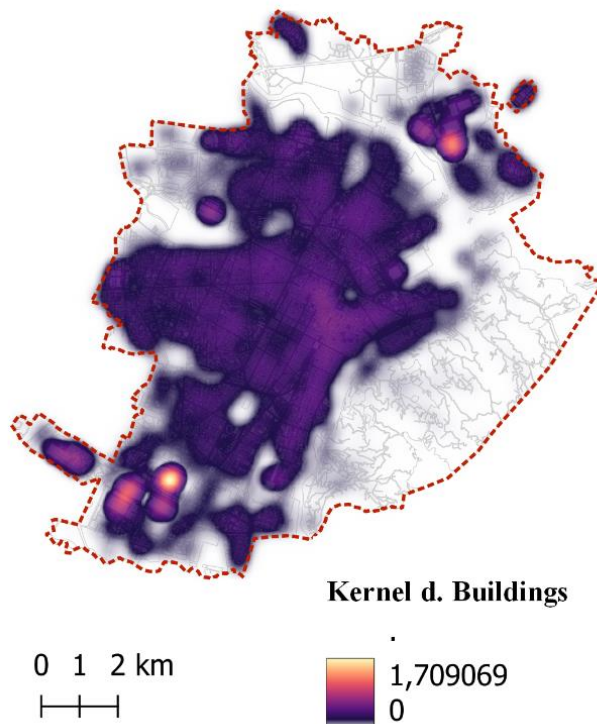
SRI calculation [QGIS] – WINTER 2018 & 2001



SRI calculation [QGis] – MID SEASON 2018 & 2001

3.2.2 Urban geo-morphology variables

- Buildings Density



Statistics:

NND (Nearest Neighbour Distance) analysis - Buildings
Observed mean distance: 18.14
Expected mean distance: 28.29
Nearest neighbour index: 0.64

Nearest Neighbour Analysis has been conducted to determine if there is any clustering or dispersion in the point distribution. The output of this analysis will give us the average nearest neighbour distance, which can help in **guiding the selection of the search radius** for the kernel density analysis.

Kernel Density Analysis was conducted on the buildings using the shapefile provided by Turin Municipality. The analysis involved converting the buildings' spatial distribution into centroids and then into a density map through the application of Kernel Density Analysis.

▪ **Sky view factor**

The Sky View Factor (SVF) is a metric used to measure the **proportion of visible sky from a location**. It ranges from 0 to 1, where 0 means no visible sky (fully obstructed), and 1 means a completely unobstructed view of the sky. SVF is valuable to assess the urban heat island effect. The **saga tool “sky view factor”** has been used for the final evaluation. The sky view factor has been performed using a DSM 1 meter of precision, later exported in 30mx30m size in order to be correctly scaled to be used also with the other satellite images.

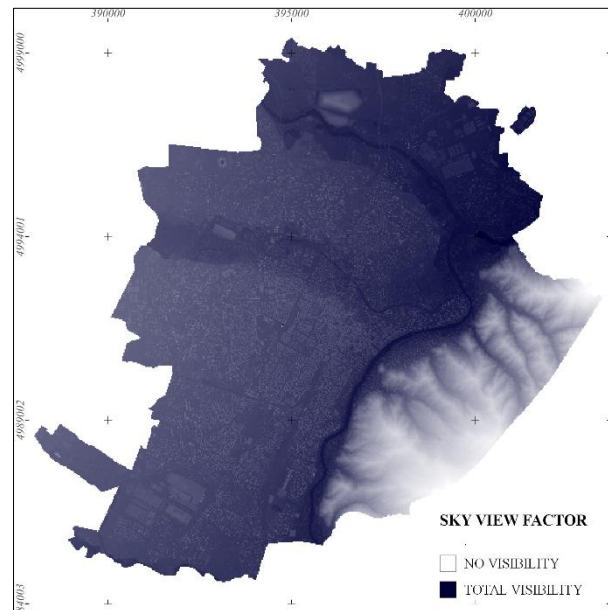


Figure 21 – SKY VIEW FACTOR

▪ S/V Ratio

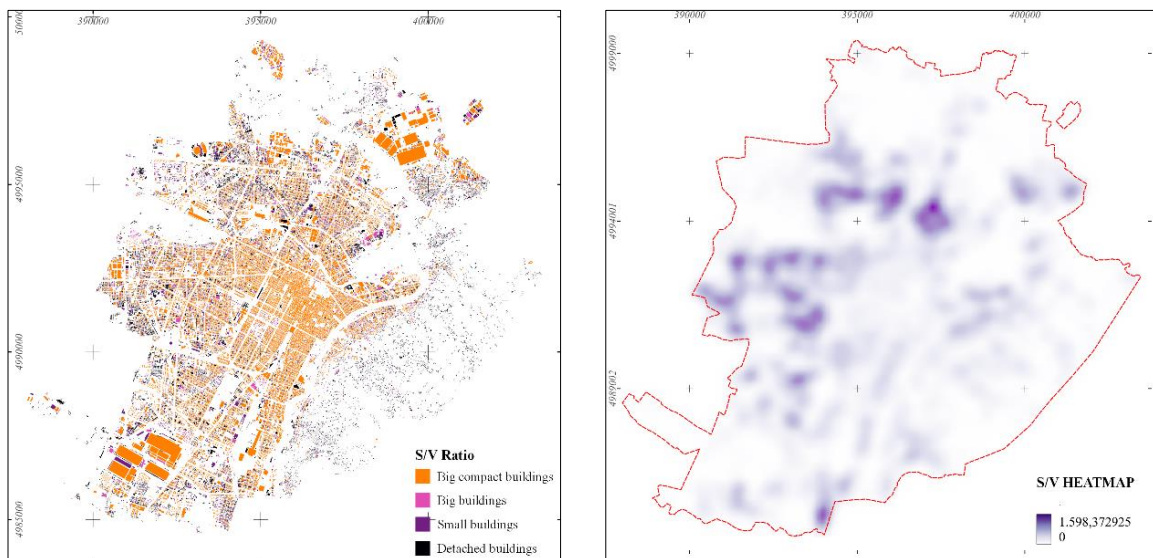


Figure 20 - S_V [Shape and heatmap QGis]

The surface-to-volume ratio (s/v ratio) plays a role in the urban heat island effect, affecting **heat absorption and retention in buildings**. We evaluated the s/v ratio using building characteristics in a shapefile and converted the data into points. This allowed us to create a kernel density map, **highlighting areas with specific s/v ratio concentrations**.

▪ **Land Surface Temperature: Landsat 7 (2001) and LANDSAT 8 (2018)**

Land Surface Temperature (LST) provides valuable information on surface heat patterns and energy exchange. To ensure accuracy and reliability of the LST, radiometric calibration is essential. This calibration process corrects for various atmospheric and sensor-related influences on the thermal data, resulting in precise LST measurements.

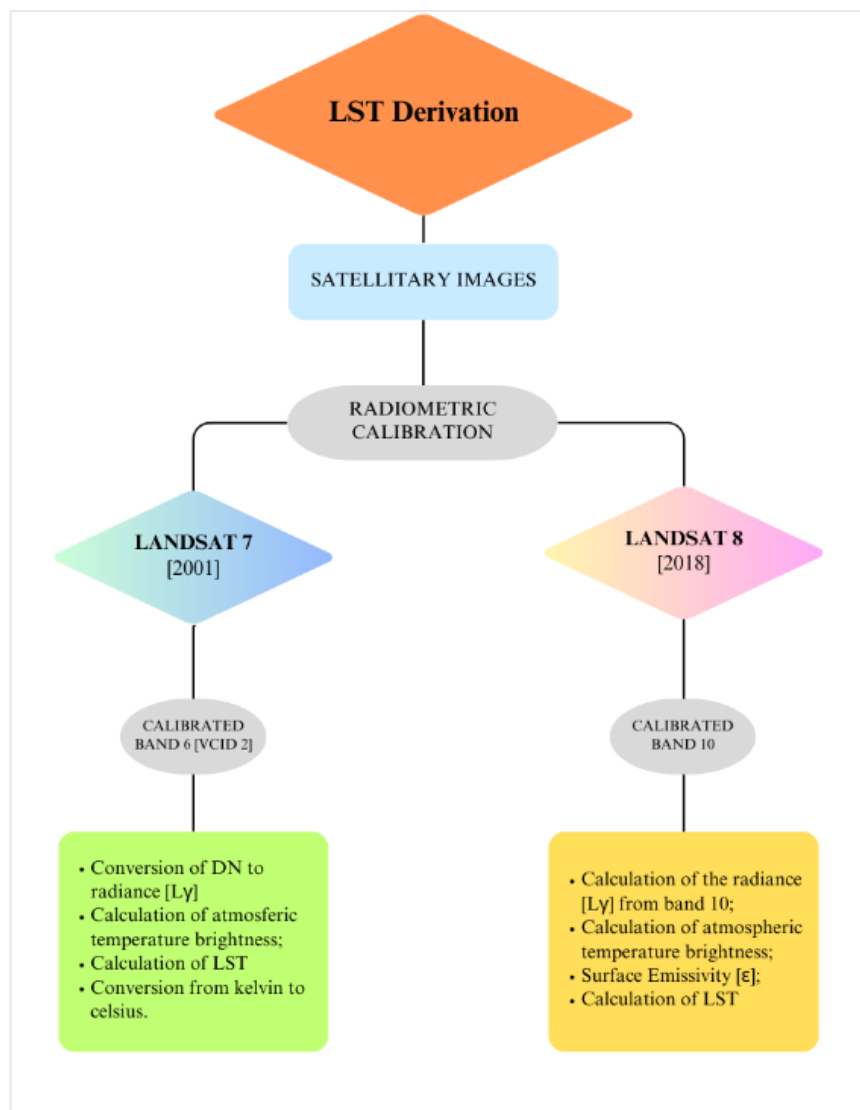


Figure 22 - LST retrieving process (Authors elaboration)

Landsat 7 (2001) – LST Calculation

For the calculation of the LST with Landsat 7 images, band 6 has been used, in fact, several steps characterised the evaluation [45].

1. Conversion from Digital Number [DN] to Radiance [L_γ]:

$$L_\gamma = \left(\frac{LMAX_\gamma - LMIN_\gamma}{QCALmax - QCALmin} \right) * (QCAL - QCALmin) + Lmin_\gamma$$

$$L_\gamma = \left(\frac{17.040 - 0.000}{255 - 1} \right) * (\text{Band 6} - 1) + 0$$

Equation 1 - Metadata information

2. Calculation of temperature brightness in Celsius:

$$T_c = \left(\frac{K2}{\ln \left(\frac{K1}{L_\gamma} + 1 \right)} \right) - 273.15$$

$$T_c = \left(\frac{1282.71}{\ln \left(\frac{666.09}{L_\gamma} + 1 \right)} \right) - 273.15$$

Equation 2 - Metadata information

[45] Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>

Landsat 8 (2018) – LST Calculation

For the calculation of the LST with Landsat 8 images, band 10 has been used, in fact, several steps characterised the evaluation [46].

1. Conversion from Digital Number [DN] to Radiance [L_γ]:

$$L_\gamma = Ml \times QCal + Al$$

$$L_\gamma = 0.1 \times \text{Band 10} + 3.3420E - 04$$

Equation 3 - Metadata information

2. Calculation of temperature brightness in Celsius:

$$T_c = \left(\frac{K2}{\ln \left(\frac{K1}{L_\gamma} + 1 \right)} \right) - 273.15$$

3.

$$TBc = \left(\frac{1321.0789}{\ln \left(\frac{774.8853}{L_\gamma} + 1 \right)} \right) - 273.15$$

Equation 4 - Metadata information

[46] Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>

Calculation of emissivity: See paragraph 3.2.1 (Land cover variables - Emissivity)

4. Calculation of Land Surface temperature in Celsius

$$T_c = T_{Bc} / (1 + (\gamma * \frac{T_B}{C_2}) \ln(E))$$

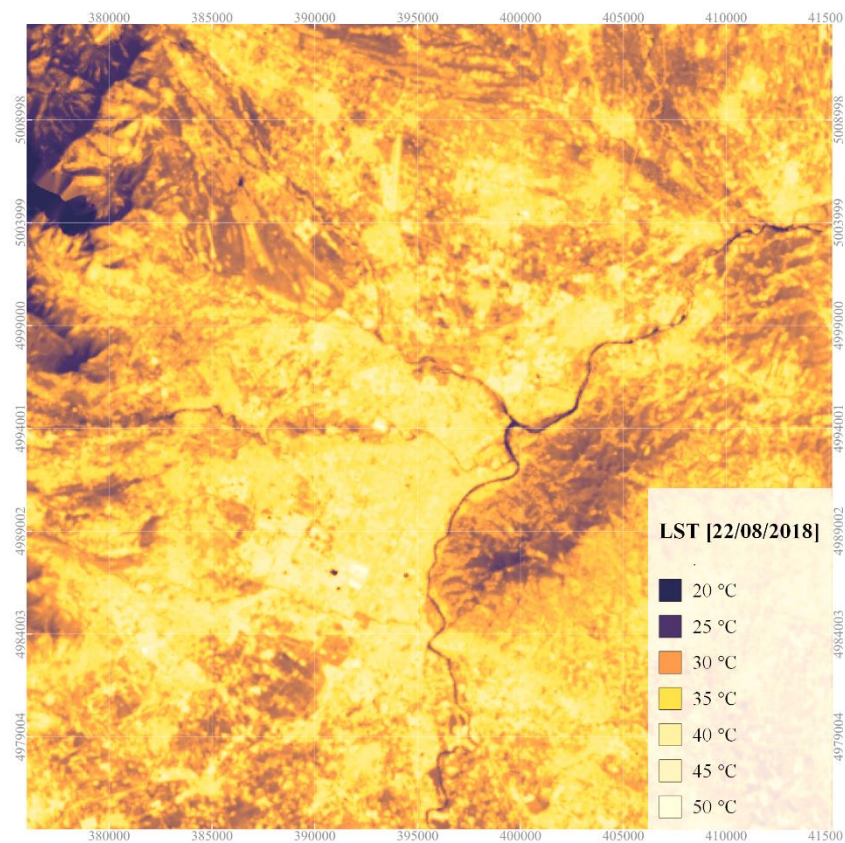
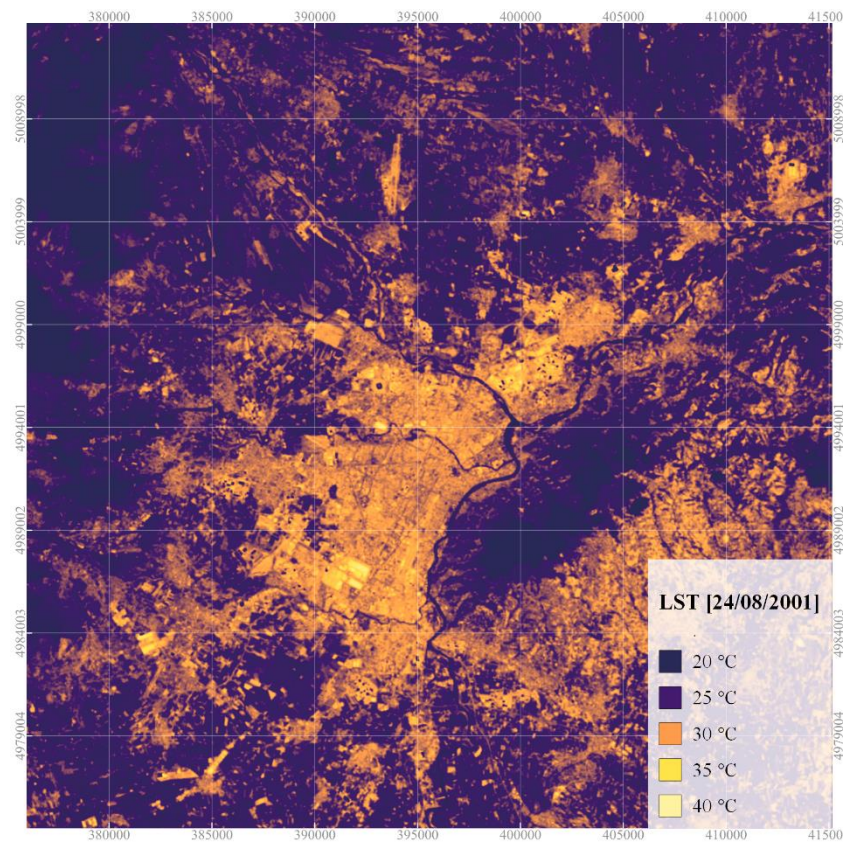
$$\gamma = 10.6$$

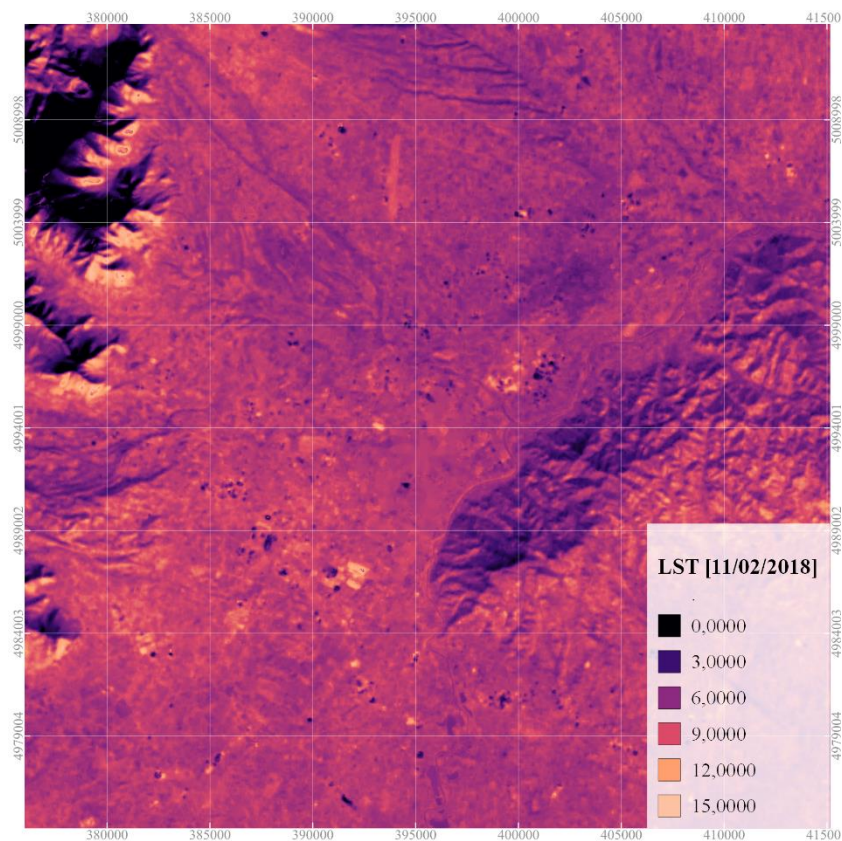
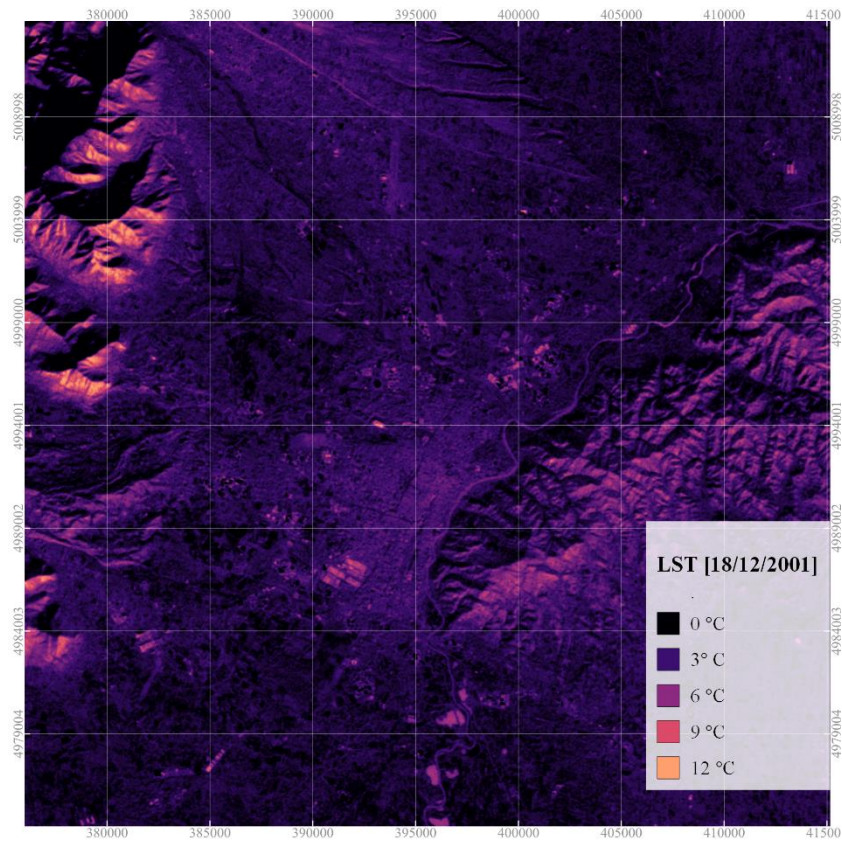
$$C_2 = 14,388 \text{ umK}$$

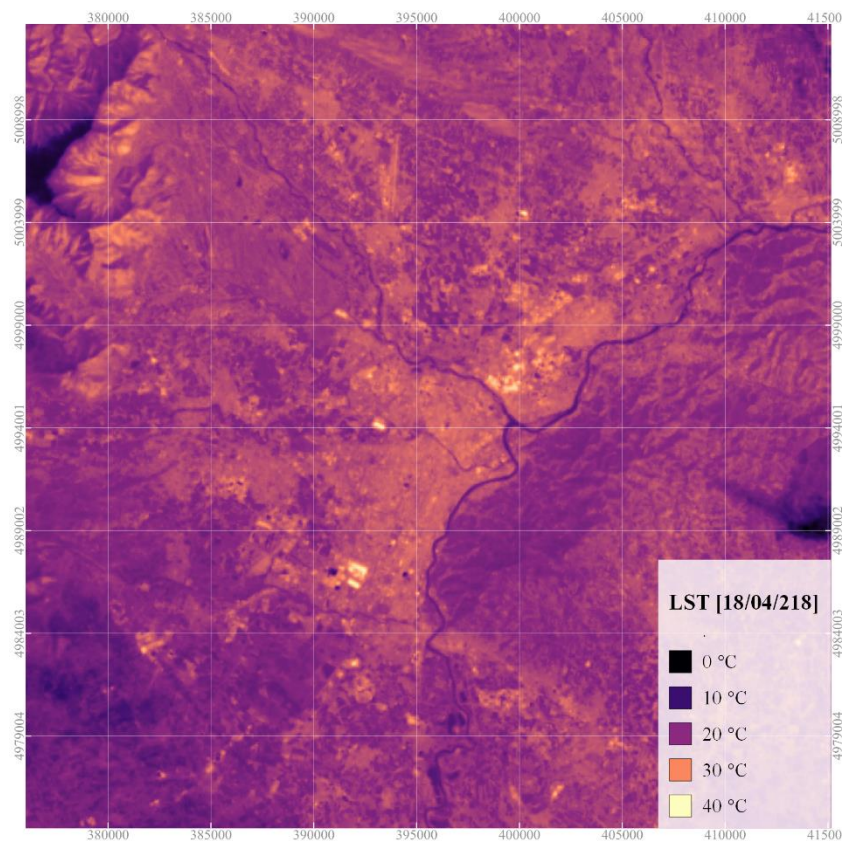
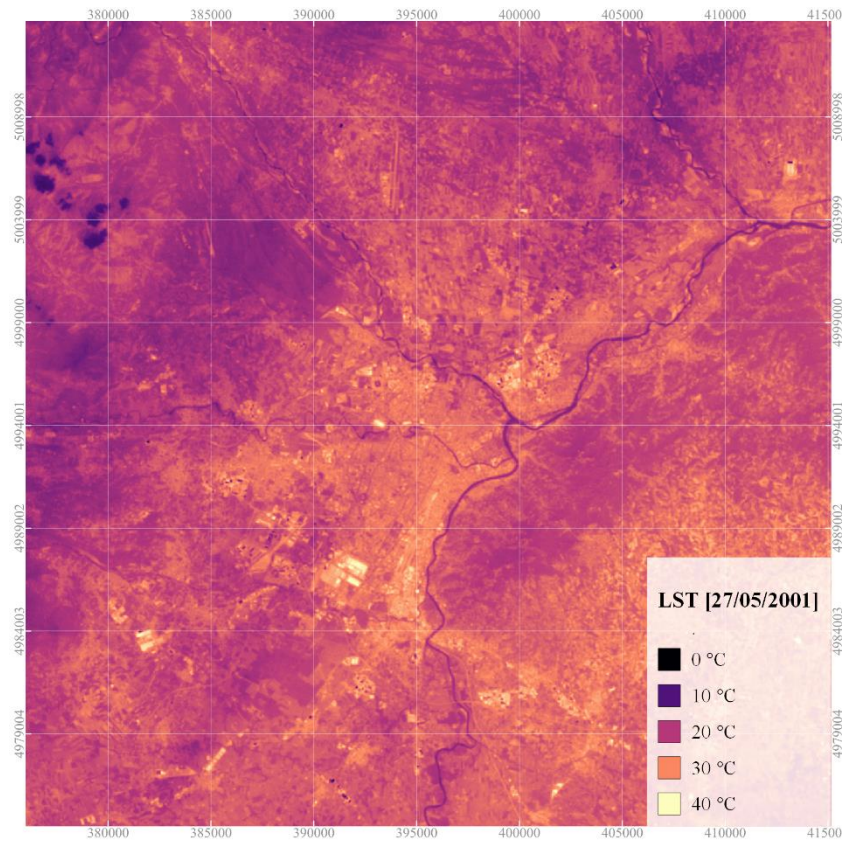
$$E = \text{Emissivity}$$

This study delves into the dynamic shifts in Land Surface Temperature (LST) by examining LST maps across distinct periods in two pivotal years, 2001 and 2018. To facilitate direct visual assessment, **LST values were categorized into standardized ranges**. This classification process establishes a harmonized framework that **enables seamless visual comparisons across different periods and years**.

Recognizing the **multifaceted influences on LST**, the **analysis extends beyond air temperature considerations**. Factors such as land use transformations, albedo modifications, cloud cover fluctuations, and broader climate shifts were acknowledged as supplementary contributors to LST variations.







3.2.3 Weather stations variables

AIR TEMPERATURE

In the context of our research, a crucial step involved the interpolation of air temperature data. To achieve this, we employed the Smart Map interpolation technique within the QGIS software. This approach was essential due to the spatially limited nature of temperature measurements. The **Smart Map interpolation method** utilizes advanced algorithms that consider the spatial relationships between data points. This enables the estimation of temperature values at locations where direct measurements were not available.

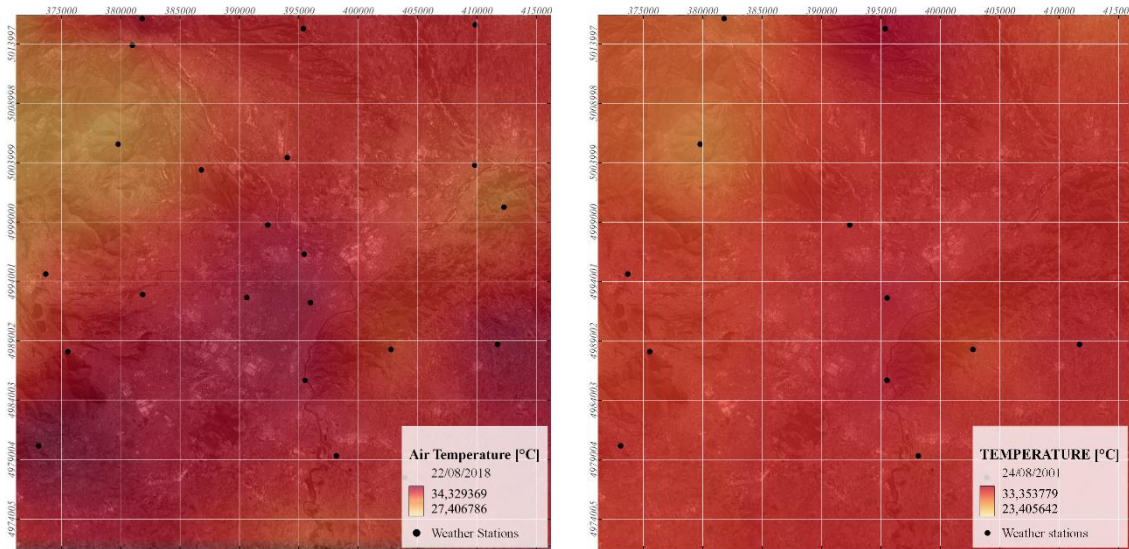
- SUMMER [2018/08/22]:

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	24.1	32.2	16.7
BAUDUCCHI	398146	4979330	24.8	33.3	16.1
BRANDIZZO MALONE	409800	5003797	24.8	32.2	17.2
CALUSO	409839	5015610	23.9	33.1	16
CANDIA LAGO	413672	5019220	23.8	32.3	17.9
CARMAGNOLA	396234	4971145	23.7	31.5	15.9
CASELLE	394011	5004435	25.2	32.3	19
CASTAGNETO PO	412254	5000266	25.5	30.6	20.8
COAZZE	361678	4989124	21.3	26.5	17.6
CUMIANA	373092	4980177	25.9	34.8	17.2
FRONT MALONE	395359	5015296	23.4	32.8	14.9
LANZO	381803	5016138	23.1	33.8	15.2
LANZO STURA DI LANZO	380979	5013879	24.5	31.3	19.1
MARENTINO	411709	4988709	26.1	34.4	19.1
PINO TORINESE	402746	4988284	25.9	31.5	21.6
RIVOLI LA PEROSA	381856	4992919	25.9	32.9	20.5

SANTENA BANNA	403909	4977517	24.5	32.8	16.8
TORINO ALENIA	390620	4992658	26.4	33.7	19.3
TORINO REISS ROMOLI	395446	4996307	26.2	33.5	19
TORINO VALLERE	395514	4985692	25.2	33.9	17.1
TORINO VIA DELLA CONSOLATA	395970	4992224	26.9	34.2	20.8
TRANA SANGONE	375548	4988109	24.4	32.7	16.4
VARISELLA	379790	5005562	23.3	28.9	19
VENARIA CERONDA	392378	4998772	24.5	33.2	15.9
VENARIA LA MANDRIA	386790	5003396	23.6	32.1	15.7
VEROLENGO	422299	5003982	24	33.7	15

- **SUMMER [2001/08/24]:**

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	0,6	8	-3,5
BAUDUCCHI	398146	4979330	1,9	8,6	-2
CARMAGNOLA	396234	4971145	1	7,9	-2,7
COAZZE	361678	4989124	1,5	5	-0,2
CUMIANA	373092	4980177	2,2	8,3	-1,6
FRONT MALONE	395359	5015296	3,7	9	0,6
LANZO	381803	5016138	2	7,3	-1,3
MARENTINO	411709	4988709	2,8	8,7	-1,1
PINO TORINESE	402746	4988284	4,8	7,7	3,3
SANTENA BANNA	403909	4977517	0,5	8,9	-3,4
TORINO BUON PASTORE	395526	4992621	3,2	9,4	0
VARISELLA	379790	5005562	1,9	7	0,1
VEROLENGO	422299	5003982	0,2	7,6	-3,5



Air temperature - Weather Station interpolation [QGIS] – SUMMER 2018 & 2001

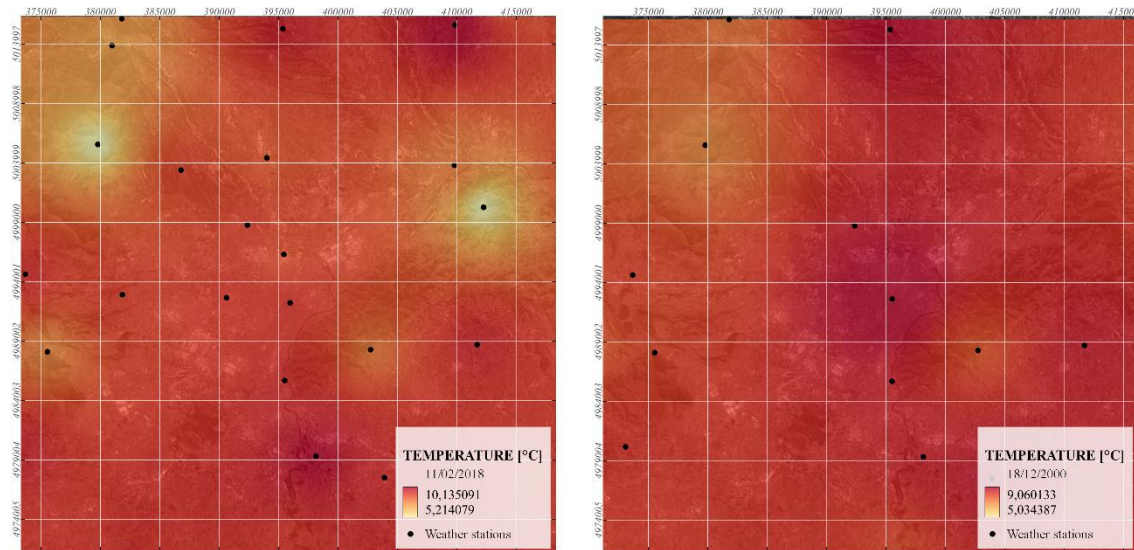
- **WINTER [2018/02/11]:**

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	2.60	9.40	-3.30
BAUDUCCHI	398146	4979330	2.70	10.20	-2.50
BRANDIZZO MALONE	409800	5003797	2.30	7.90	-2.20
CALUSO	409839	5015610	2.00	10.40	-2.70
CANDIA LAGO	413672	5019220	1.90	7.90	-2.30
CARMAGNOLA	396234	4971145	2.20	9.10	-2.70
CASELLE	394011	5004435	3.10	8.50	-1.60
CASTAGNETO PO	412254	5000266	2.40	5.10	-0.40
COAZZE	361678	4989124	0.40	4.10	-2.10
CUMIANA	373092	4980177	3.50	9.60	-1.20
FRONT MALONE	395359	5015296	1.10	9.60	-4.60
LANZO	381803	5016138	2.50	7.40	-2.00
LANZO STURA DI LANZO	380979	5013879	3.20	7.10	0.10
MARENTINO	411709	4988709	4.00	9.50	1.20
PINO TORINESE	402746	4988284	3.10	7.90	0.60
RIVOLI LA PEROSA	381856	4992919	3.20	9.00	-2.10
SANTENA BANNA	403909	4977517	2.30	9.40	-2.80
TORINO ALENIA	390620	4992658	4.00	9.10	0.40
TORINO REISS ROMOLI	395446	4996307	3.80	8.60	0.30

TORINO VALLERE	395514	4985692	3.30	9.40	-1.70
TORINO VIA DELLA CONSOLATA	395970	4992224	4.60	9.20	1.60
TRANA SANGONE	375548	4988109	1.40	7.60	-3.00
VARISELLA	379790	5005562	1.90	5.00	-1.00
VENARIA CERONDA	392378	4998772	1.70	8.90	-2.90
VENARIA LA MANDRIA	386790	5003396	1.70	8.80	-3.50
VEROLENGO	422299	5003982	1.40	9.10	-4.30

- **WINTER [2000/12/18]:**

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	0,6	8	-3,5
BAUDUCCHI	398146	4979330	1,9	8,6	-2
CARMAGNOLA	396234	4971145	1	7,9	-2,7
COAZZE	361678	4989124	1,5	5	-0,2
CUMIANA	373092	4980177	2,2	8,3	-1,6
FRONT MALONE	395359	5015296	3,7	9	0,6
LANZO	381803	5016138	2	7,3	-1,3
MARENTINO	411709	4988709	2,8	8,7	-1,1
PINO TORINESE	402746	4988284	4,8	7,7	3,3
SANTENA BANNA	403909	4977517	0,5	8,9	-3,4
TORINO BUON PASTORE	395526	4992621	3,2	9,4	0
VARISELLA	379790	5005562	1,9	7	0,1
VEROLENGO	422299	5003982	0,2	7,6	-3,5



Air temperature - Weather Station interpolation [QGis] – WINTER 2018 & 2000

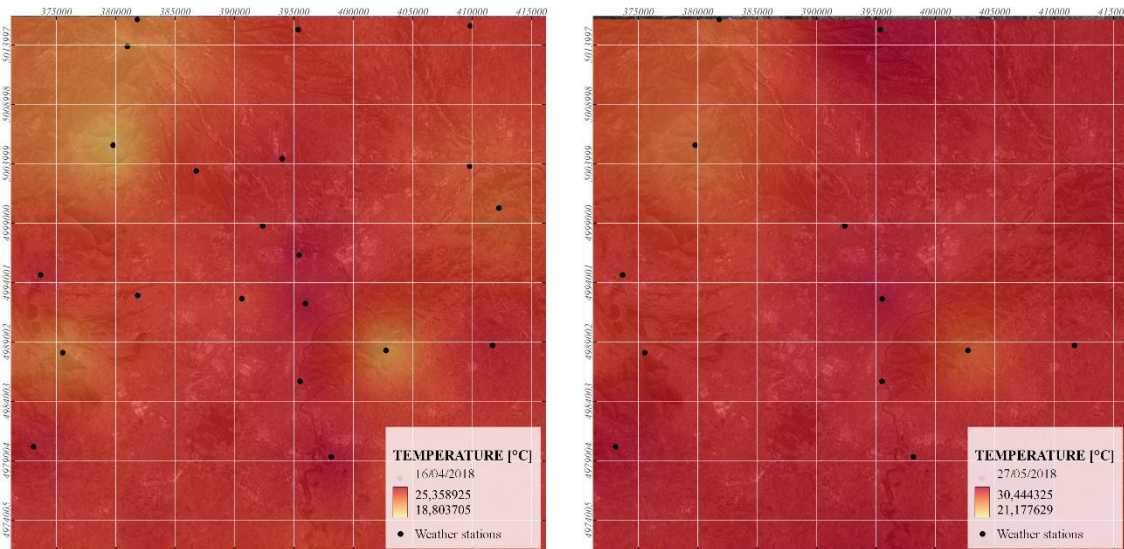
- **MID SEASON [2018/04/16]:**

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	15.90	24.60	5.80
BAUDUCCHI	398146	4979330	16.00	25.30	6.90
BRANDIZZO MALONE	409800	5003797	16.70	23.20	8.90
CALUSO	409839	5015610	15.90	24.00	7.50
CANDIA LAGO	413672	5019220	16.30	24.30	8.90
CARMAGNOLA	396234	4971145	15.30	23.70	6.90
CASELLE	394011	5004435	16.50	24.70	7.50
CASTAGNETO PO	412254	5000266	16.70	22.80	10.50
COAZZE	361678	4989124	12.60	18.80	6.90
CUMIANA	373092	4980177	16.10	24.90	6.50
FRONT MALONE	395359	5015296	14.70	24.00	4.50
LANZO	381803	5016138	14.60	22.30	6.40
LANZO STURA DI LANZO	380979	5013879	15.60	21.80	8.70
MARENTINO	411709	4988709	17.10	24.80	9.00
PINO TORINESE	402746	4988284	16.50	21.60	11.30
RIVOLI LA PEROSA	381856	4992919	16.60	23.90	7.10

SANTENA BANNA	403909	4977517	15.50	23.80	7.40
TORINO ALENIA	390620	4992658	17.40	23.90	8.90
TORINO REISS ROMOLI	395446	4996307	17.90	25.50	9.20
TORINO VALLERE	395514	4985692	16.10	24.80	7.00
TORINO VIA DELLA CONSOLATA	395970	4992224	18.30	25.40	10.40
TRANA SANGONE	375548	4988109	14.50	22.10	6.00
VARISELLA	379790	5005562	14.80	20.70	8.90
VENARIA CERONDA	392378	4998772	15.40	24.40	6.10
VENARIA LA MANDRIA	386790	5003396	14.70	23.90	4.70
VEROLENGO	422299	5003982	16.20	24.70	8.30

- **MID SEASON [2001/05/27]:**

NAME	X	Y	T_AVERAGE	T_MAX	T_MIN
AVIGLIANA	373692	4994636	20.70	28.90	13.40
BAUDUCCHI	398146	4979330	22.60	29.80	14.70
CARMAGNOLA	396234	4971145	22.00	30.20	13.80
COAZZE	361678	4989124	17.80	21.10	14.20
CUMIANA	373092	4980177	22.30	30.30	14.40
FRONT MALONE	395359	5015296	21.80	30.90	12.60
LANZO	381803	5016138	20.10	27.00	13.10
MARENTINO	411709	4988709	23.10	29.40	15.60
PINO TORINESE	402746	4988284	22.90	27.00	18.80
SANTENA BANNA	403909	4977517	21.80	29.50	13.60
TORINO BUON PASTORE	395526	4992621	23.20	30.50	15.40
TORINO VALLERE	395514	4985692	22.10	29.50	14.10
VARISELLA	379790	5005562	20.90	26.50	15.50
VEROLENGO	422299	5003982	21.00	29.90	12.20



Air temperature - Weather Station interpolation [QGIS] – MID SEASON 2018 & 2001

WIND SPEED & DIRECTION

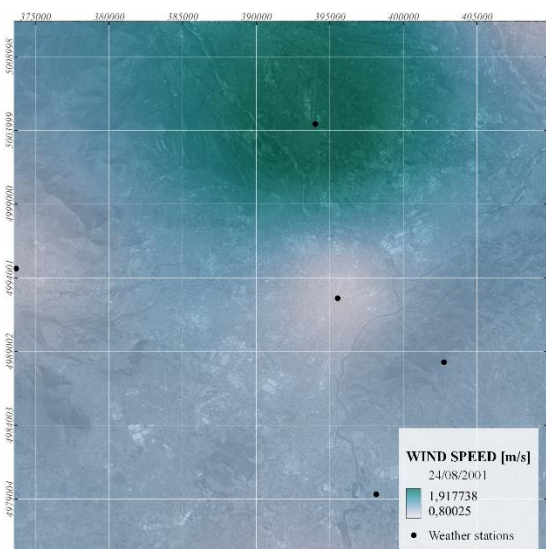
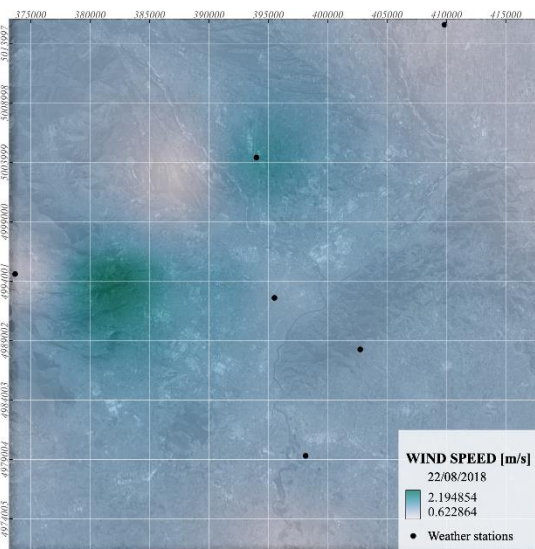
- SUMMER [2018/08/22]:

NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	0.8	WSW	247.5
BAUDUCCHI	398146	4979330	1.2	ESE	112.5
CALUSO	409839	5015610	0.8	NNW	337.5
CANDIA LAGO	413672	5019220	0.9	WSW	247.5
CARMAGNOLA	396234	4971145	0.8	S	180
CASELLE	394011	5004435	1.7	WNW	292.5
CUMIANA	373092	4980177	1.1	NW	315
PINO TORINESE	402746	4988284	1.2	SSE	157.5
RIVOLI LA PEROSA	381856	4992919	2.2	SW	225
TORINO ALENIA	390620	4992658	1.5	NW	315
TORINO REISS ROMOLI	395446	4996307	1.3	NE	45

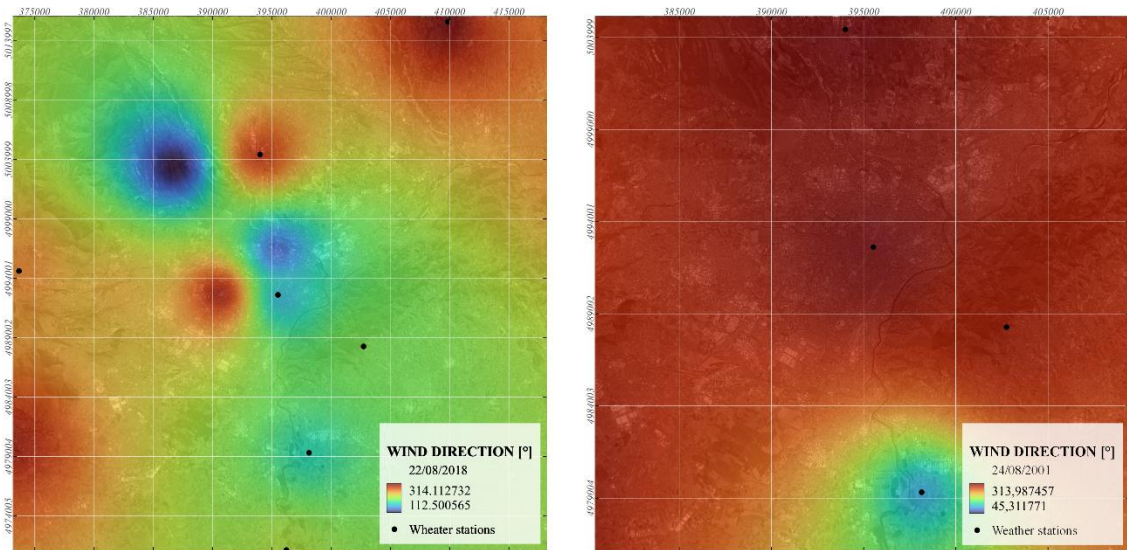
TORINO VIA DELLA CONSOLATA	395970	4992224	1.2	E	90
VENARIA LA MANDRIA	386790	5003396	0.6	N	0
VEROLENGO	422299	5003982	0.8	WSW	247.5

- **SUMMER [2001/08/24]:**

NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	1.00	W	270.00
BAUDUCCHI	398146	4979330	1.20	ESE	112.50
CARMAGNOLA	396234	4971145	0.80	NE	45.00
CUMIANA	373092	4980177	1.20	NW	315.00
PINO TORINESE	402746	4988284	1.20	WNW	292.50
TORINO BUON PASTORE	395526	4992621	0.90	NW	315.00
CASELLE	394011	5004435	2.00	NW	315.00
CALUSO	409839	5015610	0.80	W	270.00



Wind Speed - Weather Station interpolation [QGis] – SUMMER 2018 & 2001



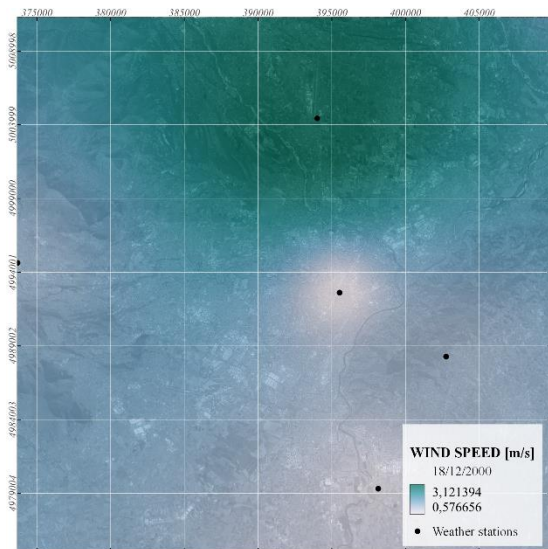
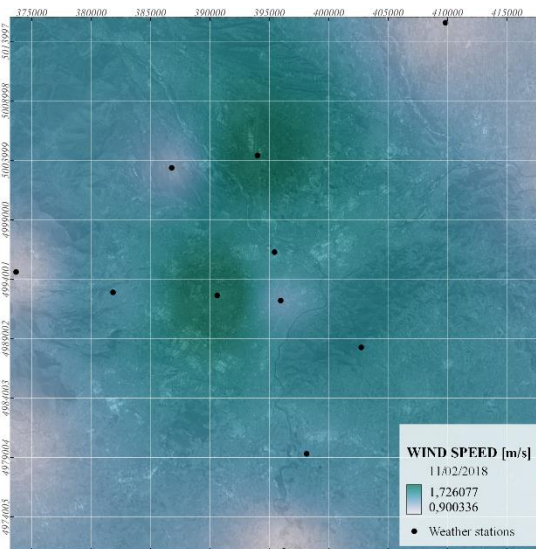
Wind Direction - Weather Station interpolation [QGis] – SUMMER 2018 & 2001

- WINTER [2018/02/11]:

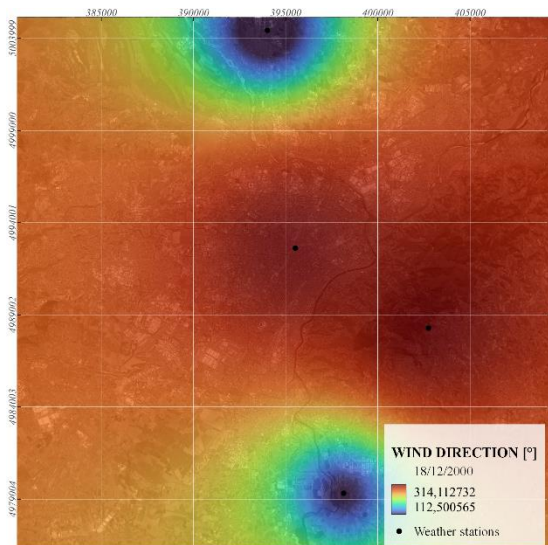
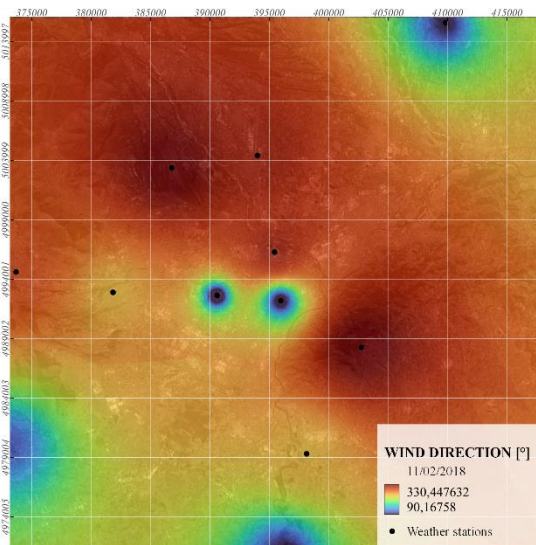
NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	1.00	WNW	292.50
BAUDUCCHI	398146	4979330	1.40	WSW	247.50
CALUSO	409839	5015610	1.00	E	90.00
CANDIA LAGO	413672	5019220	1.10	SW	225.00
CARMAGNOLA	396234	4971145	0.90	E	90.00
CASELLE	394011	5004435	1.80	NW	315.00
CUMIANA	373092	4980177	1.00	ESE	112.50
PINO TORINESE	402746	4988284	1.50	NNW	337.50
RIVOLI LA PEROSA	381856	4992919	1.40	WSW	247.50
TORINO ALENIA	390620	4992658	1.90	NNE	22.50
TORINO REISS ROMOLI	395446	4996307	1.50	NW	315.00
TORINO VIA DELLA CONSOLATA	395970	4992224	1.30	ENE	67.50
VENARIA LA MANDRIA	386790	5003396	1.20	NNW	337.50
VEROLENGO	422299	5003982	1.00	WSW	247.50

- **WINTER [2000/12/18]:**

NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	1.40	W	270.00
BAUDUCCHI	398146	4979330	0.70	ESE	112.50
CARMAGNOLA	396234	4971145	0.80	WSW	247.50
PINO TORINESE	402746	4988284	1.30	NW	315.00
TORINO BUON PASTORE	395526	4992621	0.40	NW	315.00
CASELLE	394011	5004435	3.20	E	90.00
CALUSO	409839	5015610	0.10	NW	315.00



Wind Speed - Weather Station interpolation [Qgis] – WINTER 2018 & 2000



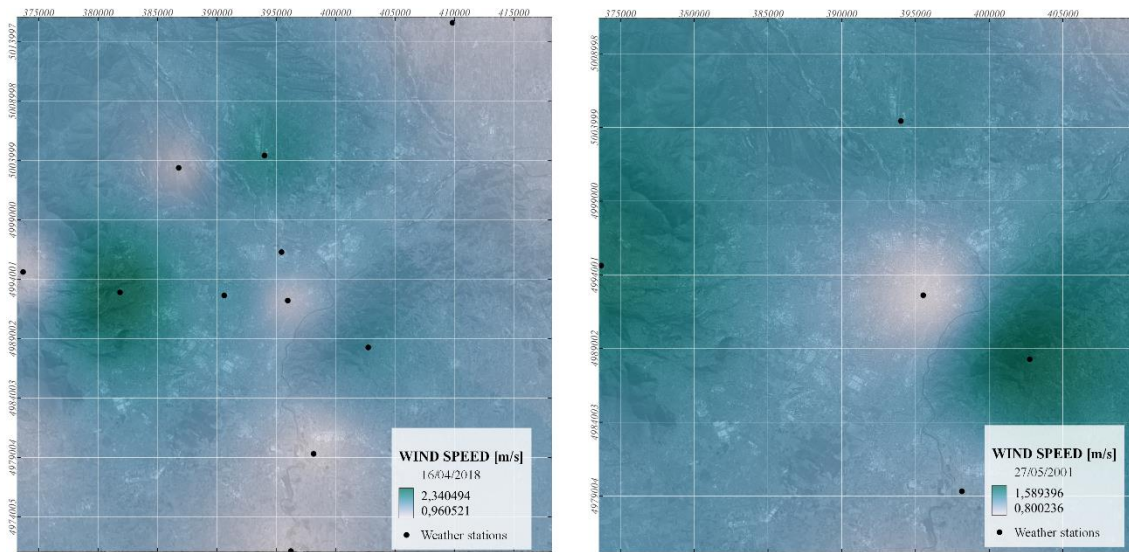
Wind Direction - Weather Station interpolation [Qgis] – WINTER 2018 & 2000

- **MID SEASON [2018/04/16]:**

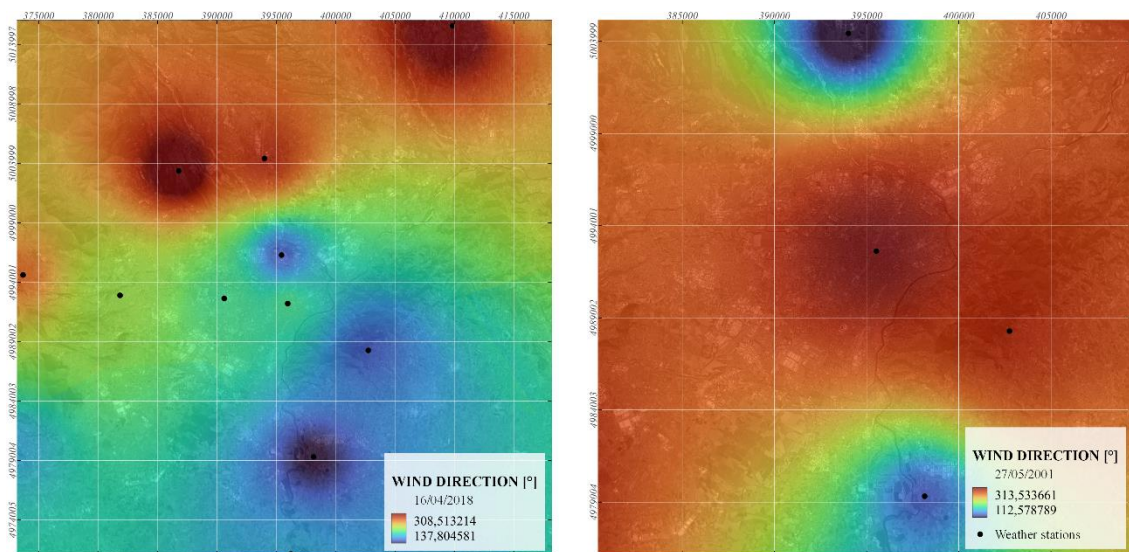
NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	1.10	W	270.00
BAUDUCCHI	398146	4979330	1.00	SE	135.00
CALUSO	409839	5015610	1.20	NNW	337.50
CANDIA LAGO	413672	5019220	0.90	SW	225.00
CARMAGNOLA	396234	4971145	1.00	S	180.00
CASELLE	394011	5004435	2.00	WNW	292.50
CUMIANA	373092	4980177	1.30	S	180.00
PINO TORINESE	402746	4988284	1.80	SSE	157.50
RIVOLI LA PEROSA	381856	4992919	2.40	SW	225.00
TORINO ALENIA	390620	4992658	1.80	SSW	202.50
TORINO REISS ROMOLI	395446	4996307	1.60	SSE	157.50
TORINO VIA DELLA CONSOLATA	395970	4992224	1.20	SSW	202.50
VENARIA LA MANDRIA	386790	5003396	1.20	NNW	337.50
VEROLENGO	422299	5003982	1.20	WSW	247.50

- **MID SEASON [2001/05/27]:**

NAME	X	Y	WIND_SPEED	SECTOR	SECTOR_DEGREE
AVIGLIANA	373692	4994636	1.50	W	270.00
BAUDUCCHI	398146	4979330	1.10	SE	135.00
CARMAGNOLA	396234	4971145	0.80	ESE	112.50
CUMIANA	373092	4980177	1.20	NW	315.00
PINO TORINESE	402746	4988284	1.60	WNW	292.50
TORINO BUON PASTORE	395526	4992621	0.80	NW	315.00
CASELLE	394011	5004435	1.30	E	90.00
CALUSO	409839	5015610	0.90	NW	315.00



Wind Speed - Weather Station interpolation [QGIS] – MID SEASON 2018 & 2001



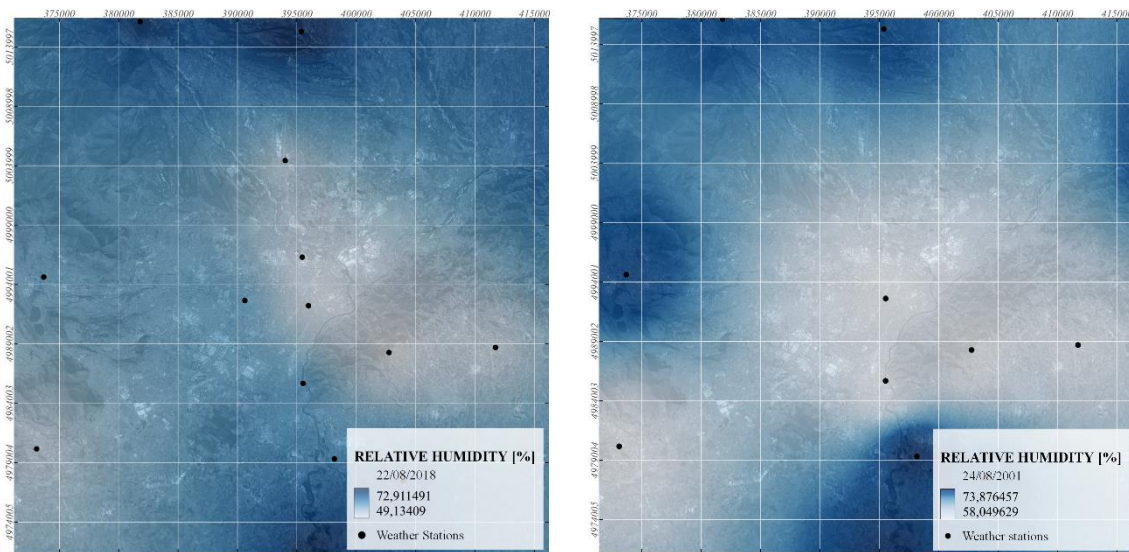
Wind Direction - Weather Station interpolation [QGIS] – MID SEASON 2018 & 2001

HUMIDITY**- SUMMER [2018/08/22]:**

NAME	X	Y	REL_HUMIDITY
AVIGLIANA	373692	4994636	60
BAUDUCCHI	398146	4979330	63
CANDIA LAGO	413672	5019220	65
CARMAGNOLA	396234	4971145	72
CASELLE	394011	5004435	55
COAZZE	361678	4989124	55
CUMIANA	373092	4980177	53
FRONT MALONE	395359	5015296	73
LANZO	381803	5016138	70
MARENTINO	411709	4988709	52
PINO TORINESE	402746	4988284	49
SANTENA BANNA	403909	4977517	68
TORINO ALENIA	390620	4992658	61
TORINO REISS ROMOLI	395446	4996307	53
TORINO VALLERE	395514	4985692	62
TORINO VIA DELLA CONSOLATA	395970	4992224	52
VEROLENGO	422299	5003982	72

- SUMMER [2001/08/24]:

NAME	X	Y	HUMIDITY
AVIGLIANA	373692	4994636	73.00
BAUDUCCHI	398146	4979330	76.00
CARMAGNOLA	396234	4971145	76.00
COAZZE	361678	4989124	73.00
CUMIANA	373092	4980177	56.00
FRONT MALONE	395359	5015296	72.00
LANZO	381803	5016138	72.00
MARENTINO	411709	4988709	59.00
PINO TORINESE	402746	4988284	54.00
SANTENA BANNA	403909	4977517	72.00
TORINO BUON PASTORE	395526	4992621	56.00
TORINO VALLERE	395514	4985692	57.00
VEROLENGO	422299	5003982	75.00



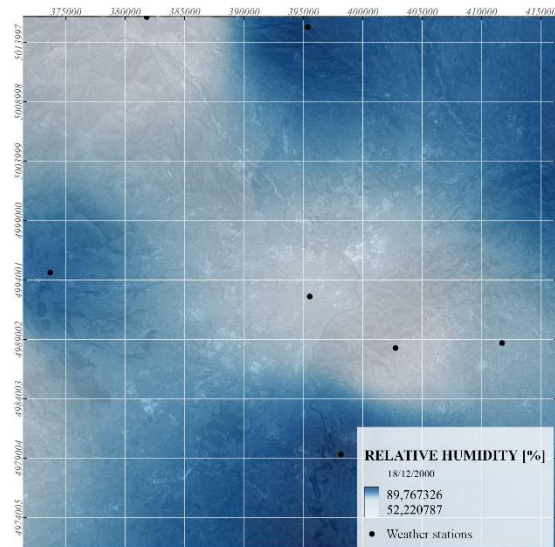
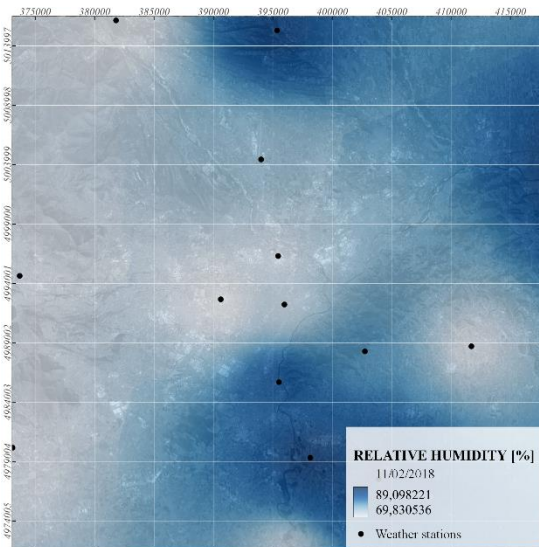
Humidity - Weather Station interpolation [QGis] – SUMMER 2018 & 2001

- **WINTER [2018/02/11]:**

NAME	X	Y	REL_HUMIDITY
AVIGLIANA	373692	4994636	71.00
BAUDUCCHI	398146	4979330	89.00
CANDIA LAGO	413672	5019220	79.00
CARMAGNOLA	396234	4971145	73.00
CASELLE	394011	5004435	76.00
COAZZE	361678	4989124	63.00
CUMIANA	373092	4980177	70.00
FRONT MALONE	395359	5015296	87.00
LANZO	381803	5016138	69.00
MARENTINO	411709	4988709	71.00
PINO TORINESE	402746	4988284	79.00
SANTENA BANNA	403909	4977517	89.00
TORINO ALENIA	390620	4992658	67.00
TORINO REISS ROMOLI	395446	4996307	72.00
TORINO VALLERE	395514	4985692	85.00
TORINO VIA DELLA CONSOLATA	395970	4992224	73.00
VEROLENGO	422299	5003982	91.00

- **WINTER [2001/12/18]:**

NAME	X	Y	T_MIN
AVIGLIANA	373692	4994636	84
BAUDUCCHI	398146	4979330	90
CARMAGNOLA	396234	4971145	91
COAZZE	361678	4989124	52
FRONT MALONE	395359	5015296	88
LANZO	381803	5016138	65
MARENTINO	411709	4988709	78
PINO TORINESE	402746	4988284	64
SANTENA BANNA	403909	4977517	84
TORINO BUON PASTORE	395526	4992621	74
VEROLENGO	422299	5003982	91



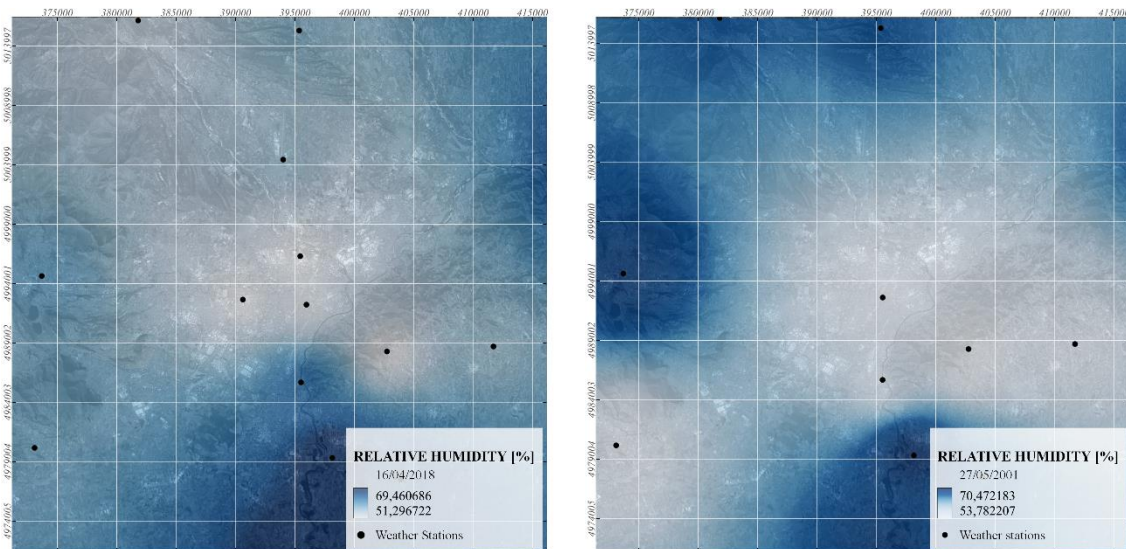
Humidity - Weather Station interpolation [QGIS] – WINTER 2018 & 2001

- **MID SEASON [2018/04/16]:**

NAME	X	Y	REL_HUMIDITY
AVIGLIANA	373692	4994636	60.00
BAUDUCCHI	398146	4979330	70.00
CANDIA LAGO	413672	5019220	62.00
CARMAGNOLA	396234	4971145	71.00
CASELLE	394011	5004435	57.00
COAZZE	361678	4989124	54.00
CUMIANA	373092	4980177	59.00
FRONT MALONE	395359	5015296	58.00
LANZO	381803	5016138	54.00
MARENTINO	411709	4988709	57.00
PINO TORINESE	402746	4988284	50.00
SANTENA BANNA	403909	4977517	66.00
TORINO ALENIA	390620	4992658	51.00
TORINO REISS ROMOLI	395446	4996307	51.00
TORINO VALLERE	395514	4985692	64.00
TORINO VIA DELLA CONSOLATA	395970	4992224	52.00
VEROLENGO	422299	5003982	71.00

- **MID SEASON [2001/05/27]:**

NAME	X	Y	HUMIDITY
AVIGLIANA	373692	4994636	71.00
BAUDUCCHI	398146	4979330	73.00
CARMAGNOLA	396234	4971145	71.00
COAZZE	361678	4989124	64.00
CUMIANA	373092	4980177	54.00
FRONT MALONE	395359	5015296	67.00
LANZO	381803	5016138	67.00
MARENTINO	411709	4988709	56.00
PINO TORINESE	402746	4988284	48.00
SANTENA BANNA	403909	4977517	66.00
TORINO BUON PASTORE	395526	4992621	54.00
TORINO VALLERE	395514	4985692	58.00
VEROLENGO	422299	5003982	69.00



Humidity - Weather Station interpolation [QGis] – MID SEASON 2018 & 2001

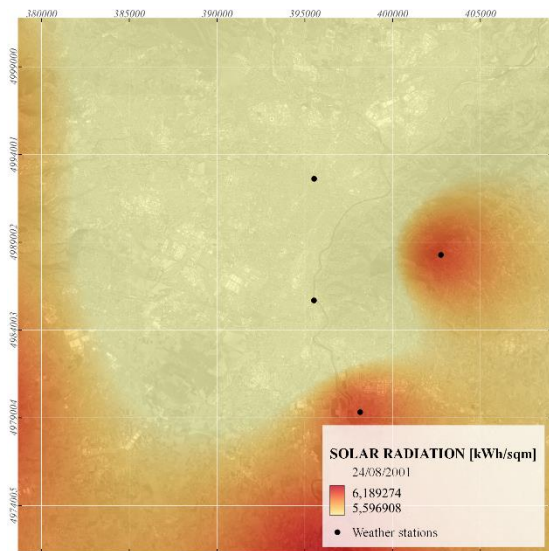
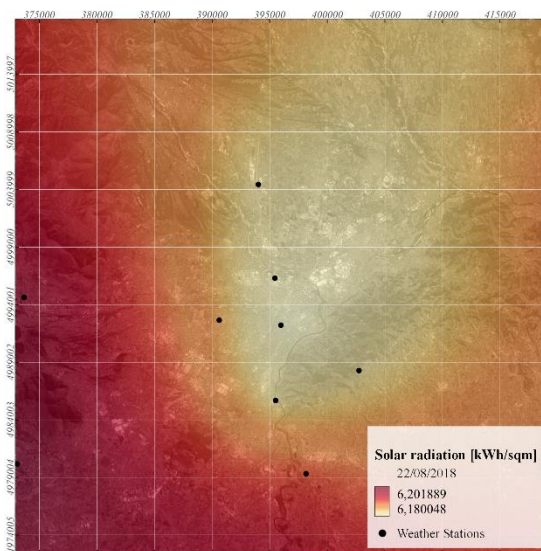
SOLAR RADIATION

- SUMMER [2018/08/22]:

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	23.6	6.56
BAUDUCCHI	398146	4979330	23.5	6.53
CANDIA LAGO	413672	5019220	21.6	6.00
CARMAGNOLA	396234	4971145	23.2	6.44
CASELLE	394011	5004435	21.6	6.00
CUMIANA	373092	4980177	23.5	6.53
PINO TORINESE	402746	4988284	22.1	6.14
TORINO ALENIA	390620	4992658	23.4	6.50
TORINO REISS ROMOLI	395446	4996307	22.2	6.17
TORINO VALLERE	395514	4985692	20.1	5.58
TORINO VIA DELLA CONSOLATA	395970	4992224	19.4	5.39
VEROLENGO	422299	5003982	23.1	6.42

- **SUMMER [2001/08/24]:**

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	21.30	5.92
BAUDUCCHI	398146	4979330	21.60	6.00
CARMAGNOLA	396234	4971145	22.00	6.11
CUMIANA	373092	4980177	22.30	6.19
PINO TORINESE	402746	4988284	21.60	6.00
TORINO BUON PASTORE	395526	4992621	14.70	4.08
TORINO VALLERE	395514	4985692	15.10	4.19
VEROLENGO	422299	5003982	20.50	5.69



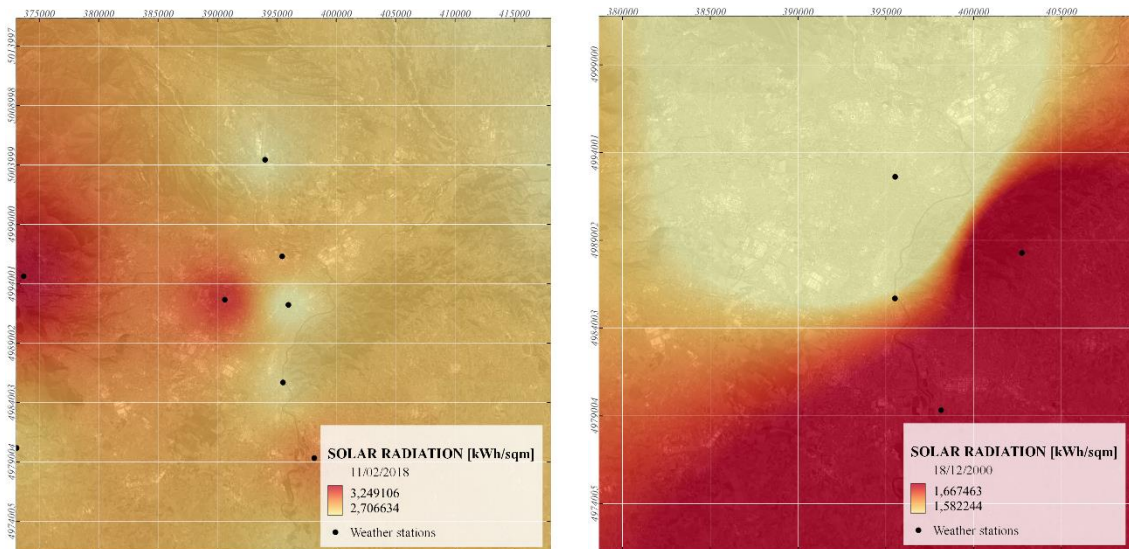
*Solar radiation - Weather Station interpolation [QGis] – **SUMMER 2018 & 2001***

- **WINTER [2018/02/11]:**

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	11.70	3.25
BAUDUCCHI	398146	4979330	10.70	2.97
CANDIA LAGO	413672	5019220	9.80	2.72
CARMAGNOLA	396234	4971145	10.00	2.78
CASELLE	394011	5004435	9.80	2.72
CUMIANA	373092	4980177	10.10	2.81
TORINO ALENIA	390620	4992658	11.40	3.17
TORINO REISS ROMOLI	395446	4996307	10.40	2.89
TORINO VALLERE	395514	4985692	9.90	2.75
TORINO VIA DELLA CONSOLATA	395970	4992224	9.60	2.67
VEROLENGO	422299	5003982	9.70	2.69

- **WINTER [2000/12/18]:**

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	5,8	1,6111124
BAUDUCCHI	398146	4979330	6,1	1,6944458
CARMAGNOLA	396234	4971145	6,3	1,7500014
PINO TORINESE	402746	4988284	6,3	1,7500014
TORINO BUON PASTORE	395526	4992621	4,5	1,250001
VEROLENGO	422299	5003982	6	1,666668



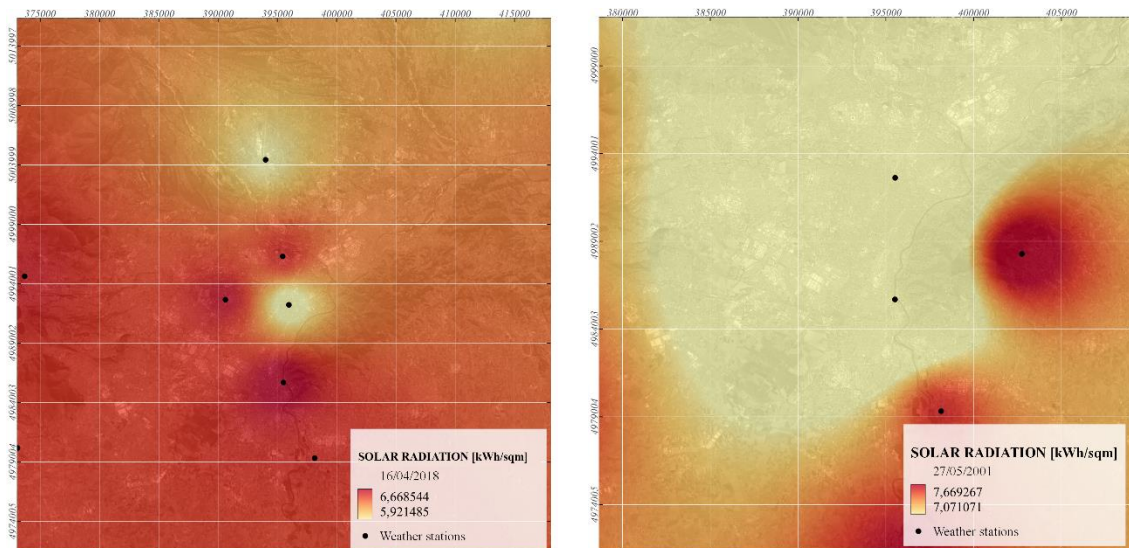
*Solar radiation - Weather Station interpolation [QGis] – **SUMMER 2018 & 2001***

- **MID SEASON [2018/04/16]:**

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	23.7	6.58
BAUDUCCHI	398146	4979330	23.5	6.53
CANDIA LAGO	413672	5019220	21.8	6.06
CARMAGNOLA	396234	4971145	23.3	6.47
CASELLE	394011	5004435	21.3	5.92
CUMIANA	373092	4980177	23.4	6.50
TORINO ALENIA	390620	4992658	23.9	6.64
TORINO REISS ROMOLI	395446	4996307	23.6	6.56
TORINO VALLERE	395514	4985692	24.3	6.75
TORINO VIA DELLA CONSOLATA	395970	4992224	20.7	5.75
VEROLENGO	422299	5003982	22.7	6.31

- **MID SEASON [2001/05/27]:**

NAME	X	Y	SOLAR RADIATION [MJ/mq]	SOLAR RADIATION [kWh/mq]
AVIGLIANA	373692	4994636	26.80	7.44
BAUDUCCHI	398146	4979330	27.30	7.58
CARMAGNOLA	396234	4971145	27.60	7.67
CUMIANA	373092	4980177	27.60	7.67
PINO TORINESE	402746	4988284	28.10	7.81
TORINO BUON PASTORE	395526	4992621	19.00	5.28
TORINO VALLERE	395514	4985692	19.20	5.33
VEROLENGO	422299	5003982	26.30	7.31



*Solar radiation - Weather Station interpolation [QGis] – **MID SEASON 2018 & 2001***

3.2.4 UHI evaluation

The Urban Heat Island (UHI) phenomenon, characterized by elevated temperatures in urban areas compared to their rural surroundings, can be quantified using a standardized approach. The UHI intensity is computed using the formula [47]:

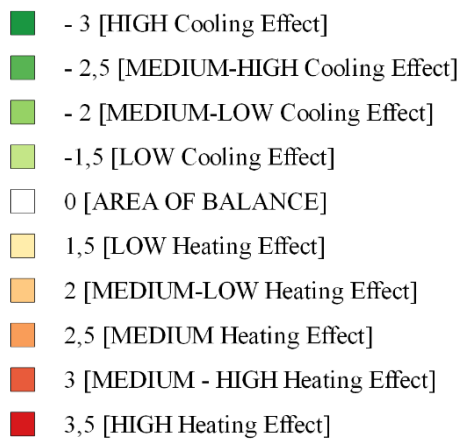
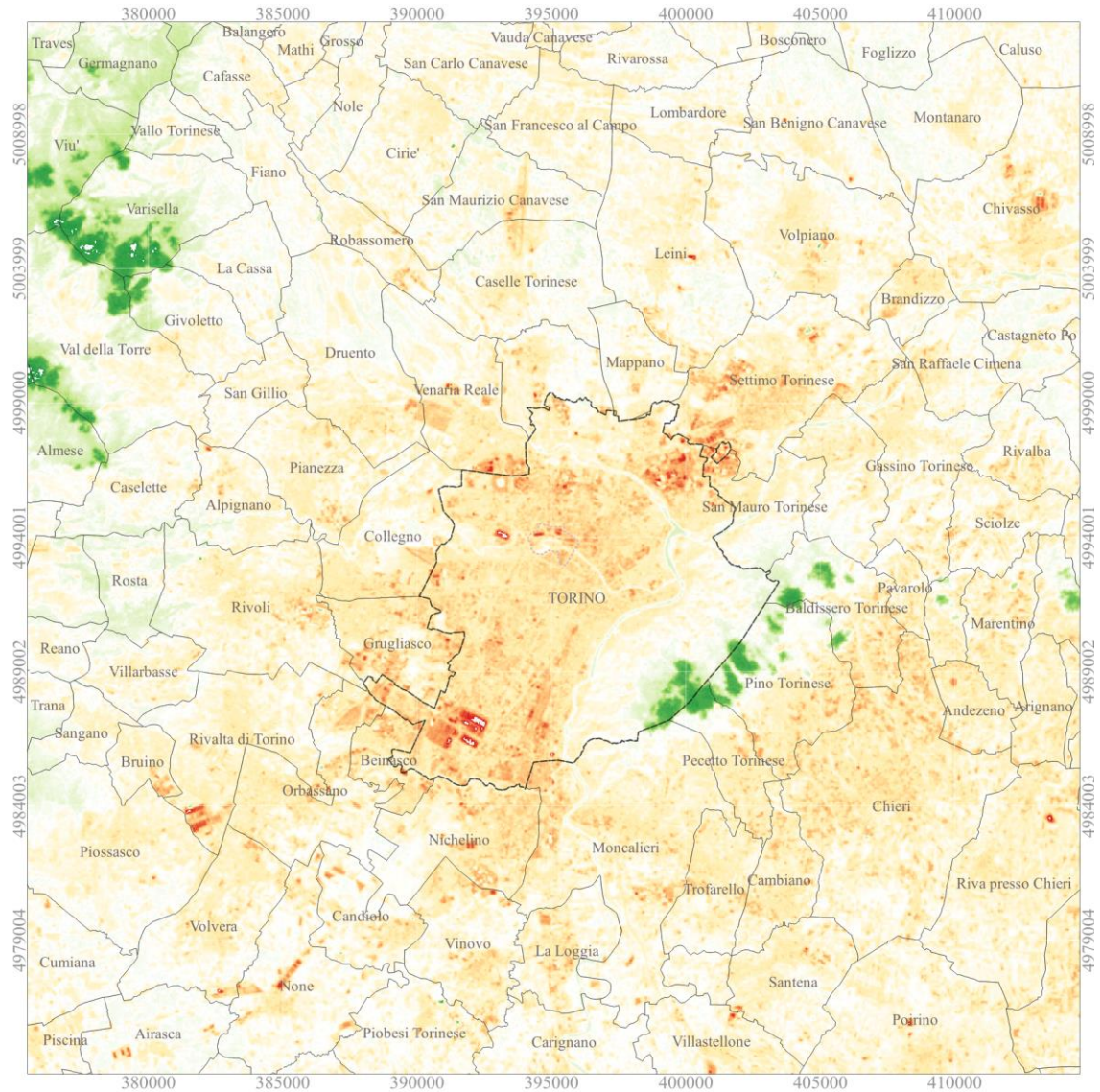
$$\text{UHI} = \frac{\text{LST} - \text{LST}_{\text{mean}}}{\text{LST}_{\text{stddev}}}$$

This formula serves **to normalize the LST value by considering its deviation from the mean LST and scaling it by the variability** (standard deviation) of LST within the region. A positive UHI value signifies that the LST at the given location is higher than the regional average, indicating an urban area with intensified heat. Conversely, a negative UHI value suggests that the LST is lower than the regional average, signifying a cooling effect in comparison to the surroundings.

By employing this standardized calculation, UHI assessments gain a context-independent measure of temperature deviations, enabling meaningful comparisons between diverse urban and rural regions. This approach contributes to a more comprehensive understanding of UHI dynamics and supports informed urban planning and environmental management strategies.

[47] Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>

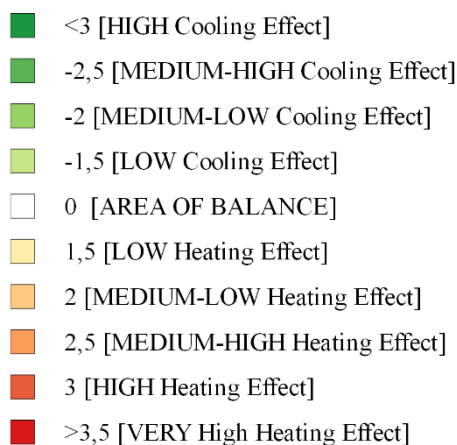
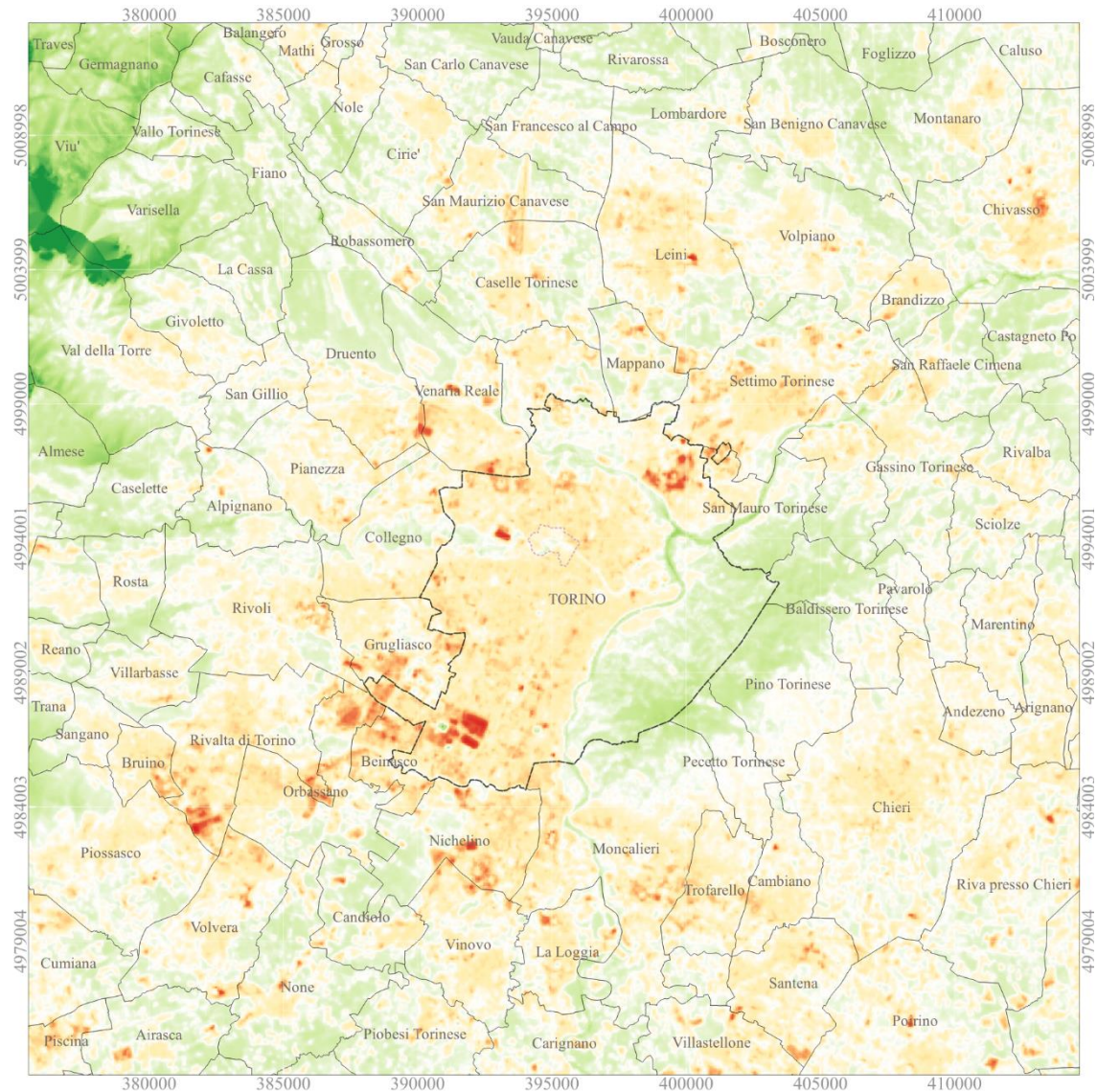
UHI Intensity 24/08/2001 - SUMMER



Notes:

As evident from the map, the Urban Heat Island (UHI) effect has been notably pronounced within regions marked by industrial land utilization, as well as areas characterized by high building density. Specifically, this phenomenon is discernible in the **southern and northern sectors of the city of Turin**. Meanwhile, in the provincial zone the effect registers higher values in correspondence of built-up areas where human activities take place.

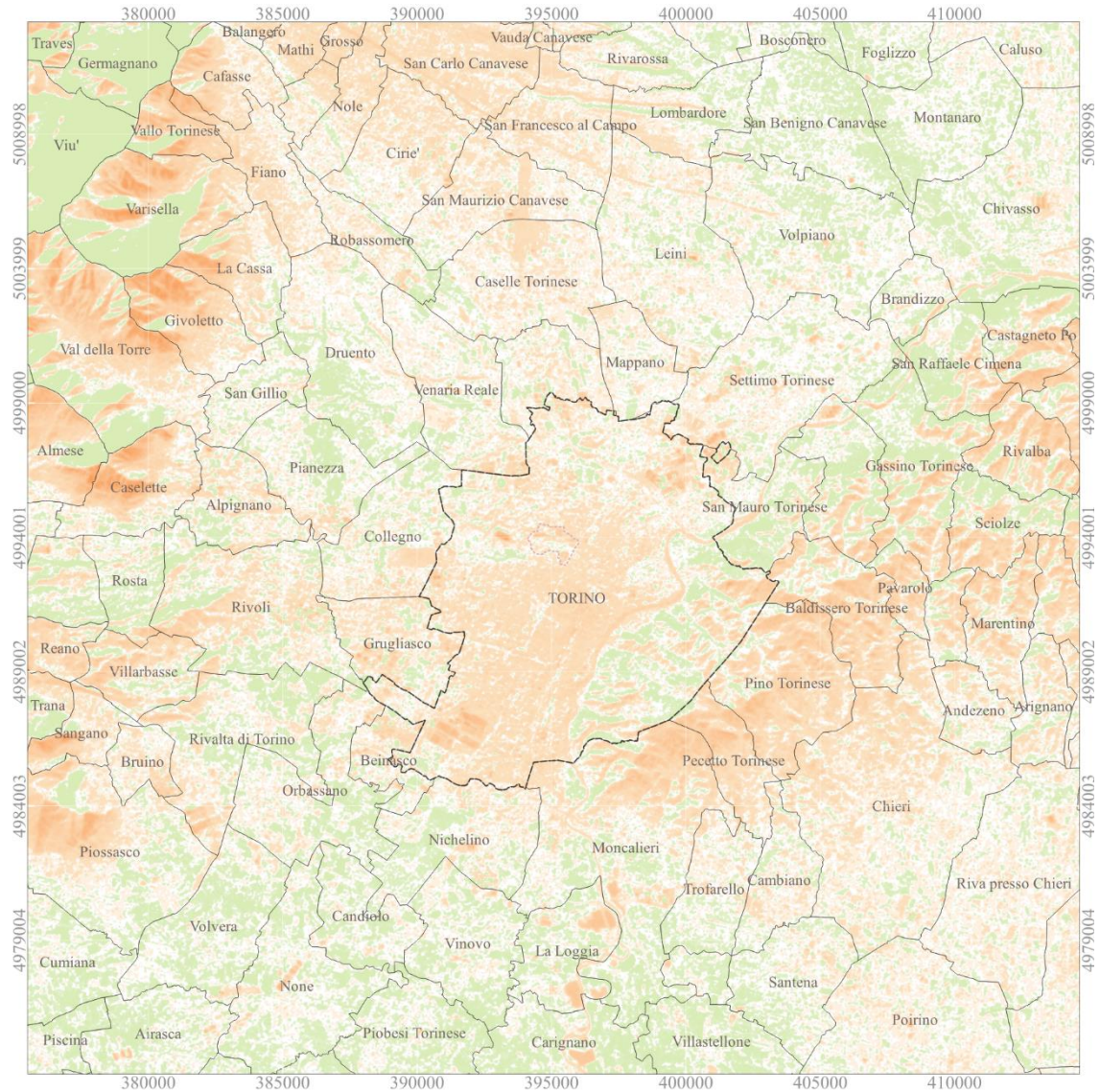
UHI Intensity 22/08/2018 - SUMMER



Notes:

In 2018, Turin's Urban Heat Island (UHI) effect exhibited increased balance and cooling areas attributed to changes in land use. This citywide improvement marked a reduced UHI impact. In the provincial region, compared to 2001, we observe more balanced and cooling areas, indicating a less prominent UHI effect overall. However, areas with historical industrial presence (southern part of the region) continue to experience exacerbated UHI effects.

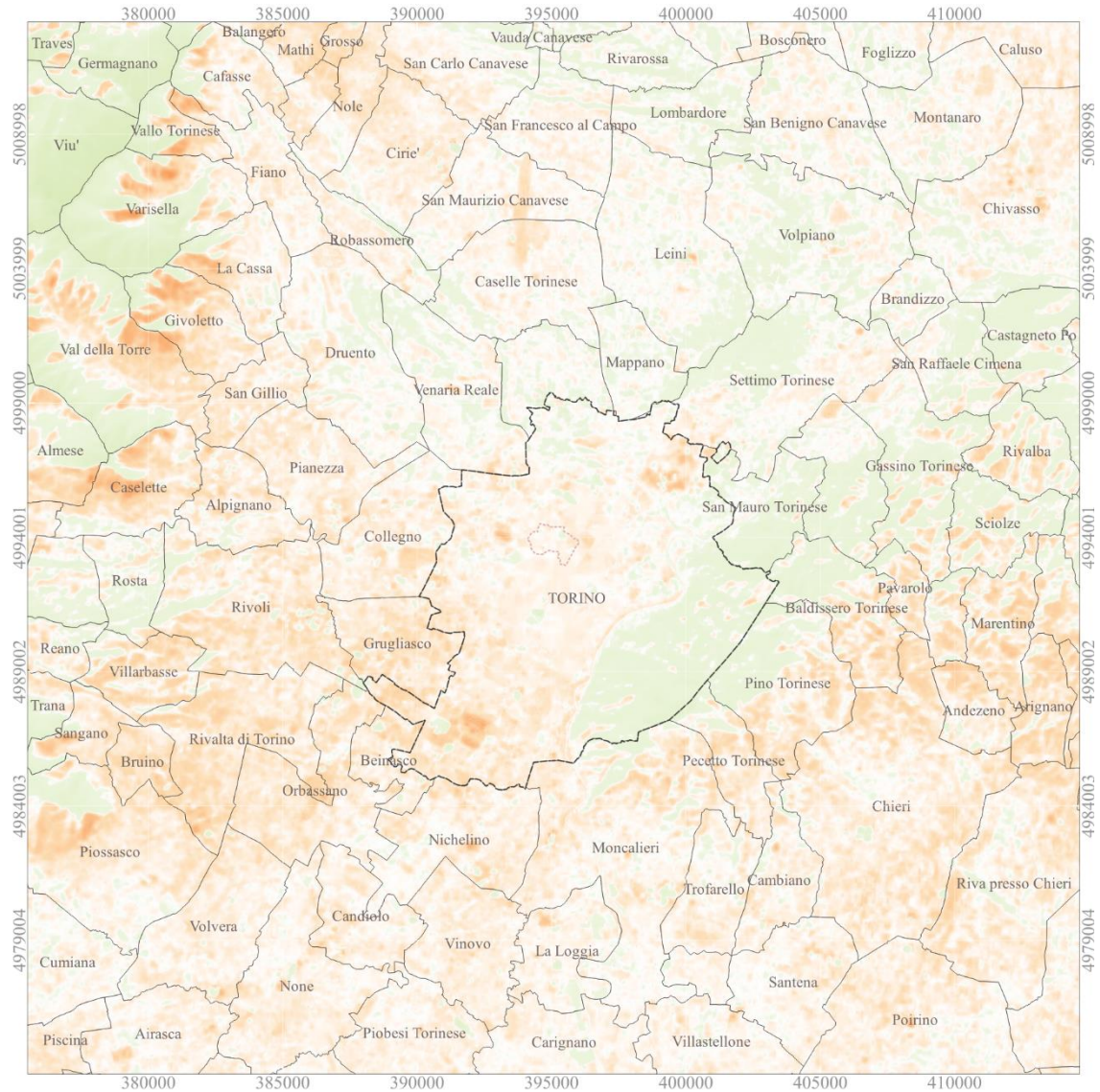
UHI Intensity 18/12/2000 - WINTER



Notes:

In the winter of 2000, Turin exhibited a distinct Urban Heat Island (UHI) phenomenon, characterized by elevated temperatures in its southern areas. This warming was influenced by factors such as industrial and residential concentrations, house heating and heat-absorbing building materials.

UHI Intensity 11/02/2018 - WINTER

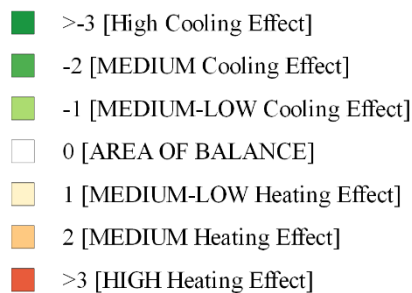
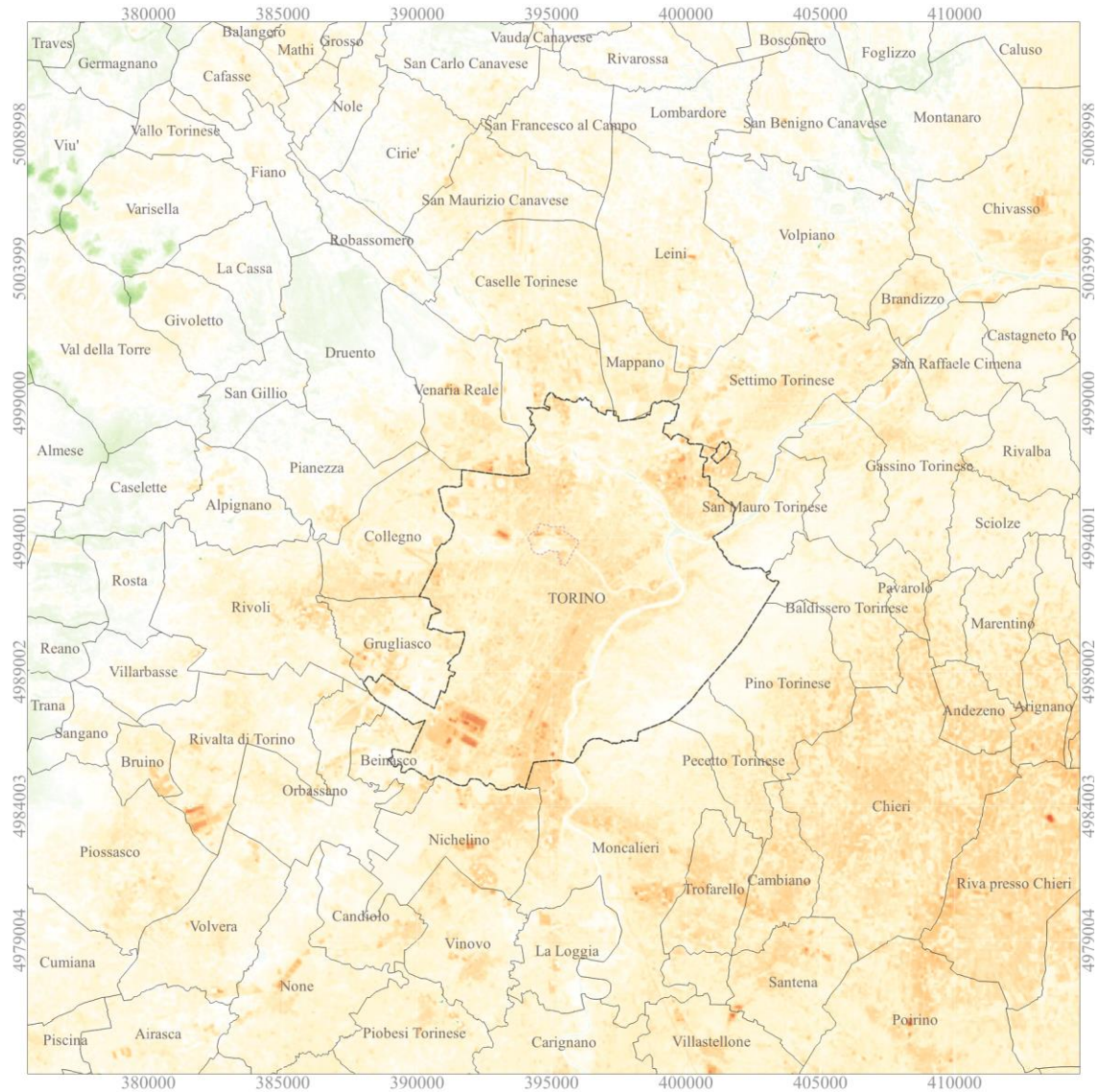


Notes:

Compared to the winter 2000, in 2018 are observable more areas of balance. It is important to notice that the UHI phenomena has been sprawl homogeneously in the whole provincial territory.

- 1 [MEDIUM-LOW Cooling effect]
- 0,5 [LOW Cooling effect]
- 0 [AREA OF BALANCE]
- 1 [MEDIUM-LOW Heating Effect]
- 2 [MEDIUM Heating Effect]

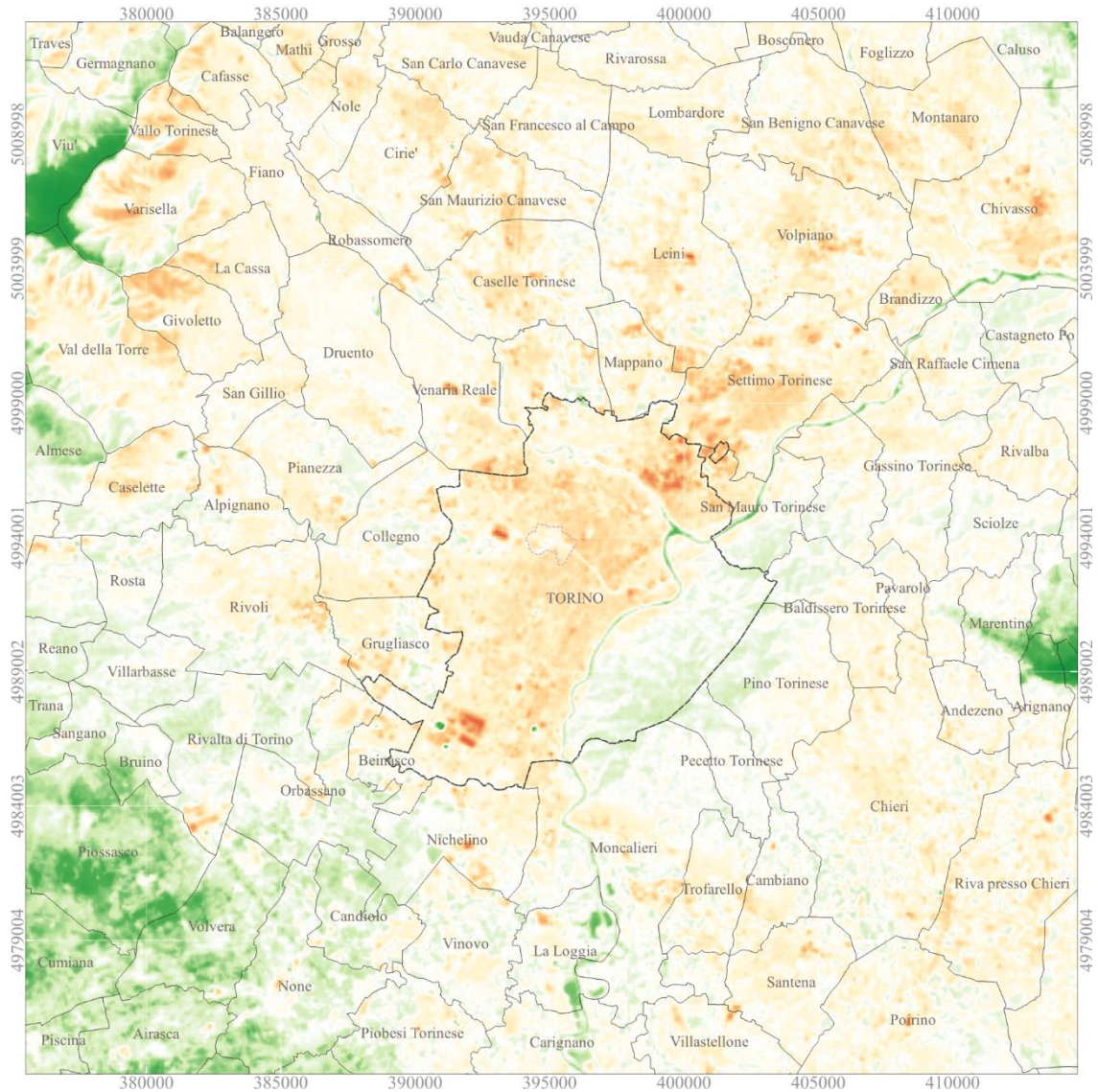
UHI Intensity 27/05/2001 – MID SEASON



Notes:

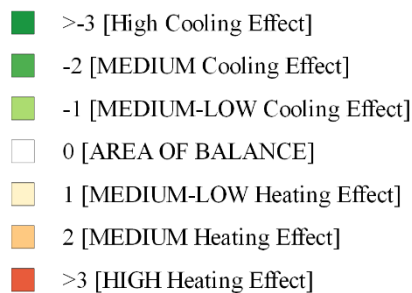
In the mid-season of 2001, it is evident that the UHI phenomenon is concentrated within the city of Turin. In contrast, the provincial territory exhibits a more balanced distribution, with less pronounced intensity. Furthermore, the highest intensity is observed in proximity to industrial areas.

UHI Intensity 16/04/2018 – MID SEASON



Notes:

In 2018, the UHI effect in Turin exhibits a greater balance between warmer and cooler areas, in contrast to the conditions observed in 2001. The significant industrial complexes, owing to their size, materials, and activities, play a substantial role in contributing to the UHI effect, even in their immediate vicinity.



4 Modelling: Variables Influencing Urban Heat Island Effect in Turin

In the context of Urban Heat Island (UHI) research, the modelling phase refers to the crucial step where we develop **computational models to study and simulate the phenomenon of Urban Heat Islands**.

4.1 Hyperparameters tuning

The convergence of machine learning techniques, geospatial data analysis, and hyperparameter optimization has yielded powerful tools for addressing complex environmental and ecological challenges. In this context, the following code snippet presents a holistic process that embodies these concepts.

This process revolves around optimizing a **“Random Forest Classifier”** through intricate hyperparameter tuning using the **“Optuna library”**. Geospatial data, characterized by its multidimensional nature and spatial referencing, plays a pivotal role in environmental analysis. The code commences by engaging with a collection of raster files, each housing crucial environmental information. The **raster’s files** are not uniform in their spatial properties, necessitating a **harmonization process**. This is achieved through **reprojection and resampling**, aligning all raster data with the attributes of a reference raster. The result is a harmonized dataset that can be readily employed for model development.

The synthesized dataset, enriched with simulated labels to emulate a classification task, undergoes partitioning into **training and testing subsets**. This partition is fundamental to the subsequent model evaluation process, ensuring the assessment of model performance on unseen data. It is in this setting that the Optuna library assumes a significant role. Optuna, a library specializing in hyperparameter optimization, orchestrates an intricate exploration of the hyperparameter space. Hyperparameters, integral to the configuration of machine learning algorithms, influence the model's

behaviour and predictive prowess. Through systematic trial and evaluation, Optuna seeks to pinpoint the combination of hyperparameters that maximizes the model's accuracy.

The outcome of the optimization journey culminates in **the identification of optimal hyperparameters**. These parameters, fine-tuned for the specific dataset, are then applied in the creation of a final Random Forest Classifier. This classifier, harnessed with the optimal hyperparameters, is subsequently trained on the designated training dataset. The culmination of this training imbues the model with predictive capabilities, shaping it into a refined tool for classification.

Importantly, the optimization of these parameters serves a larger purpose: the **creation of a robust model that can be effectively integrated into ArcGIS Pro using the Random Forest Regression tool**. By fine-tuning the hyperparameters, the model's performance is optimized, enabling accurate predictions and informed decision-making within the ArcGIS Pro environment.

Below is an illustrative excerpt of the Python code introduced within Data Spell to perform hyperparameter tuning.

The complete code is provided in Section B of the [appendix](#).

```
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna

# List of raster file paths
raster_paths = [
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio 2001/0_CROP/NDWI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio 2001/0_CROP/NDMI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI METEOREOLOGICA/2001_05_27/Turin Tiff/0_RASTER_TIFF/humidity.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio 2001/0_CROP/EMISSIONITY_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI METEOREOLOGICA/2001_05_27/Turin Tiff/0_RASTER_TIFF/wind_sector.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
```



```

Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI_TERMICHE/2001/LANDSAT
7/CALCOLI Maggio 2001/0_CROP/ALBEDO_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/1_BANCA DATI
METEOROLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/temperature.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI_TERMICHE/2001/LANDSAT
7/CALCOLI Maggio 2001/0_CROP/NDVI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI_TERMICHE/2001/LANDSAT
7/CALCOLI Maggio 2001/0_CROP/SRI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/1_BANCA DATI
METEOROLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/wind_speed.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/1_BANCA DATI
METEOROLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/solar_rad.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI_TERMICHE/2001/LANDSAT
7/CALCOLI Maggio 2001/0_CROP/NDBI_TO.tif',
]

reference_raster_path = 'C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI_TERMICHE/2001/LANDSAT
7/LE07_L1TP_195029_20010527_20200917_02_T1/CALCULATION_NEW/UHI/UHI_TO_
27_05_2001.tif'

# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band

# Initialize an empty array to hold all the raster data
raster_data_list = []

from rasterio.warp import reproject, Resampling

# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
            src_transform=src.transform,
            src_crs=src.crs,
            dst_transform=ref_src.transform,
            dst_crs=ref_src.crs,
            resampling=Resampling.bilinear
        )
        raster_data_list.append(resampled_raster.flatten())

# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)

```

```

# Generate synthetic labels for demonstration purposes (replace with
your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array,
labels, test_size=0.2, random_state=42)

# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)

    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )

    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3,
scoring='accuracy'))

# Create an Optuna study
study = optuna.create_study(direction='maximize')

# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train),
n_trials=15, n_jobs=-1)

# Get the best parameters from the study
best_params = study.best_params

# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)

# Fit the final model on the training data
best_rf.fit(X_train, y_train)

# Make predictions
y_pred = best_rf.predict(X_test)

# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)

```

```
# Print best parameters, cross-validation results, and evaluation
metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f"{param}: {value}")

print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)

print(f"Accuracy: {accuracy:.2f}")
print(f"F1-score: {f1:.2f}")
print(f"ROC-AUC: {roc_auc:.2f}")
```

Figure 23 - Data Spell Python code example for hyper parametrization tuning

Here following, is provided the explanation of the code used:

The code use “**Optuna**” library for hyperparameter optimization with a “**Random Forest Classifier**” in the context of remote sensing image analysis.

1. **Importing Libraries:** The necessary libraries are imported at the beginning of the code, including ``numpy``, ``sklearn``, ``rasterio``, and ``optuna``. These libraries are used for data manipulation, machine learning modelling, remote sensing image processing, and hyperparameter optimization.
2. **Raster File Paths:** A list of file paths pointing to different raster image files is provided. These raster images likely contain the explanatory training raster (e.g., wind speed, temperature etc.).
3. **Reference Raster Path:** The path to a reference raster image is provided. This reference raster will serve as the template for reprojecting other raster images to ensure that they all have the same spatial characteristics.
4. **Loading Reference Raster:** The reference raster is loaded using the ``rasterio`` library. The first band of the raster is read and stored in the ``reference_raster`` variable.
5. **Loading Raster Data:** A loop iterates through the list of raster paths. For each path, the corresponding raster data is loaded using ``rasterio``. The raster data is then reprojected to match the spatial characteristics of the reference raster using

bilinear resampling. The reprojected raster data is flattened and added to the ``raster_data_list``.

6. **Combining Raster Data:** All the reprojected and flattened raster data is combined into a single ``raster_data_array`` using ``numpy``.
7. **Generating Labels:** Synthetic labels are generated for demonstration purposes. These labels are randomly generated binary values (0 or 1) and are associated with the raster data samples.
8. **Train-Test Split:** The combined raster data (``raster_data_array``) and the generated labels are split into training and testing sets using the ``train_test_split`` function from ``sklearn``. This is a common step in machine learning to evaluate the model's performance on unseen data.
9. **Defining the Objective Function:** The ``objective`` function is defined, which takes in Optuna's ``trial`` object, the training data (``X_train`` and ``y_train``). Inside this function, different hyperparameters for the ``Random Forest Classifier`` are suggested by the ``trial.suggest_int`` method. Then, a ``Random Forest Classifier`` is created with those hyperparameters, and its cross-validated accuracy is returned.
10. **Creating an Optuna Study:** An Optuna study is created to manage the optimization process. The study is set to maximize the objective function.
11. **Optimizing the Study:** The study's ``optimize`` method is called with the ``objective`` function and the training data. The ``n_trials`` parameter defines how many trials (hyperparameter combinations) Optuna should try to optimize the objective function. The ``n_jobs`` parameter specifies parallelization, with ``-1`` indicating using all available CPU cores.
12. **Getting Best Parameters:** The best set of hyperparameters found during the optimization is retrieved from the study's ``best_params``.
13. **Creating Final Random Forest Classifier:** Using the best hyperparameters, a final ``Random Forest Classifier`` (``best_rf``) is created.
14. **Fitting the Model:** The final model is fitted on the training data using the ``fit`` method.
15. **Making Predictions:** The trained model is used to predict labels for the test data (``X_test``).

16. Printing Results: The best hyperparameters, cross-validation results, and evaluation metrics are printed to the console.

For what concern the parameters we imputed in the system are:

- **Test size:** we allocated 20% of the data for testing purposes.
- **Range of random forest trials:** We defined a range of values for the following parameters in the random forest model: number of estimators, max depth, minimum sample split, and minimum sample leaf. These parameters will be explored systematically to find the best combination that results in optimal performance.
- **Folds in cross validation (cv):** We employed a 3-fold cross-validation approach. This means the dataset is divided into three subsets, and the model's performance is evaluated three times, each time using a different subset for testing and the remaining for training.
- **Number of trials:** We set the number of optimization trials to a specified value, denoted as 'n'.

To have a better understanding of the results is reported an output of the process for a best iteration, performed with the images of 24/08/2001.

```
[I 2023-08-18 09:01:44,559] Trial 28 finished with value: 0.8055555555555555.
trial 6 with value: 0.8055555555555555.
Best Parameters:
n_estimators: 59
max_depth: 50
min_samples_split: 4
min_samples_leaf: 1
Cross-Validation Results:
```

	number	value	...	params_n_estimators	state
0	0	0.694444	...	64	COMPLETE
1	1	0.694444	...	145	COMPLETE
2	2	0.694444	...	183	COMPLETE
3	3	0.694444	...	164	COMPLETE
4	4	0.694444	...	93	COMPLETE
5	5	0.694444	...	165	COMPLETE
6	6	0.805556	...	59	COMPLETE
7	7	0.694444	...	100	COMPLETE
8	8	0.694444	...	111	COMPLETE

Figure 24 - Best parameters output from Python code

What cross validation is? “Cross-validation is a step in the process of building a machine learning model which helps us ensure that our models fit the data accurately and also ensures that we do not overfit [48]”.

▪ **Classification problems, metrics available for machine learning method:**

- **Accuracy**
- Precision (P)
- Recall (R)
- F1 score (F1)
- Area under the ROC (Receiver Operating Characteristics) curve or simply AUC (AUC)
- Log loss
- Precision at k (AP@k)
- Mean average precision at k (MAP@k)

In this case, we used accuracy as classification problems metric, “**Accuracy** is one of the most straightforward metrics used in machine learning. It defines how accurate your model is. To do an example, if we build a model that classifies 90 images accurately, our accuracy is 90% or 0.90. If only 83 images are classified correctly, the accuracy of our model is 83% or 0.83 [49]”.

For our analysis we chose those parameters that presents an **accuracy value greater than 0.70**, once the best parameters were estimated, we inserted them into ArcGIS Pro to launch the random forest regression model.

[48] Thakur, Abhishek. *Approaching (Almost) Any Machine Learning Problem* (English Edition) (p.14). Abhishek Thakur. Kindle Edition.

[49] Thakur, Abhishek. *Approaching (Almost) Any Machine Learning Problem* (p.32). Abhishek Thakur. Edition Kindle.

4.2 Machine Learning algorithm [RF]

- **What Random Forest is [50]** - Random Forest [RF] is an ensemble learning method, which means it combines the predictions of multiple individual models (in this case, decision trees) to make a final prediction. The idea behind ensemble learning is that by combining the predictions of several models, we can improve the overall accuracy and robustness of the final prediction.
 - **Decision Trees:** A decision tree is a simple tree-like structure that breaks down a dataset into smaller and smaller subsets by making decisions based on the values of input features. Each internal node in the tree represents a decision or a test on a specific feature, and each *leaf node* represents the final prediction or classification.
 - **Training on Random Subsets:** In a Random Forest, multiple decision trees are created, and each tree is trained on a random subset of the original data. This randomness ensures that each decision tree is trained on slightly different data, introducing diversity into the ensemble.
 - **Using Random Subsets of Features:** Not only is the data sampled randomly, but also the features (input variables) used to split the data at each node of the decision tree are randomly chosen. This means that at each node, the decision tree considers only a subset of features to make the best split. The number of features considered at each node is typically less than the total number of features in the dataset.
 - **Aggregating Predictions:** After training multiple decision trees on their respective random subsets, the predictions from each individual tree are combined to make the final prediction. In the case of classification tasks (e.g., predicting whether an area has a high, medium, or low UHI effect), the final prediction is determined by majority voting. Each tree "votes" for a specific

[50] Esri. (2023). Random Forest guide in ArcGIS. Retrieved July 31, 2023, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/forestbasedclassificationregression.htm>

class, and the class with the most votes becomes the final prediction. In regression tasks (e.g., predicting the exact temperature increase), the final prediction is the average of the predictions from all the trees.

- **Overfitting Reduction and Generalization Improvement:** The randomness introduced during both data sampling and feature selection helps to reduce overfitting. Overfitting occurs when a model becomes too complex and fits the training data too closely, leading to poor performance on new, unseen data. By using random subsets of data and features, Random Forest prevents individual trees from memorizing the training data and encourages them to capture more general patterns in the data. This improves the model's ability to generalize and make accurate predictions on new data.

- **Usefulness for UHI Evaluation**

- ***High Accuracy:*** Random Forest tends to deliver high accuracy in predictions, making it suitable for complex and non-linear relationships often present in UHI studies. It can efficiently capture the intricate interactions between various factors contributing to UHI, such as land cover, urban morphology, meteorological conditions, and anthropogenic activities.
- ***Feature Importance (sensitivity):*** Random Forest can provide information about the importance of different features in predicting UHI intensity. This feature importance analysis helps researchers identify the most critical variables influencing UHI, guiding urban planners and policymakers in making informed decisions to mitigate UHI effects.
- ***Robustness to Outliers and Noise:*** UHI evaluation often involves dealing with noisy or imperfect data. Random Forest is robust to outliers and noise, reducing the risk of the model being overly influenced by data irregularities.

- **Handling Missing Data:** Random Forest can handle missing data in a dataset effectively. This is crucial in real-world UHI studies, where data gaps or sensor failures can be common.
- **Non-linearity:** UHI is a complex phenomenon with non-linear relationships between different variables. Random Forest can capture these non-linear patterns, making it more suitable than traditional linear regression models.

▪ **Benefits from a Statistical Point of View**

- **Bias-Variance Trade-off:** Random Forest mitigates the bias-variance trade-off, a common challenge in machine learning. It reduces overfitting by averaging predictions from multiple trees, striking a balance between model complexity and generalization performance.
- **Consistent Performance:** Random Forest tends to have stable and consistent performance on different datasets. This consistency is particularly beneficial in UHI evaluation, where generalization to various urban environments is essential.
- **Model Interpretability:** While Random Forest is not as interpretable as linear models, it still provides insights into feature importance (sensitivity). This information aids in understanding the relative impact of various factors on UHI formation.
- **Handling Multicollinearity:** UHI variables may be correlated, leading to multicollinearity issues in traditional regression models. Random Forest is more robust to multicollinearity, allowing it to handle such situations effectively.

4.3 Training & Prediction model of the UHI in the case study

Within this section, we showcase the findings derived from the implementation of the random forest regression model. The model was executed within ArcGIS Pro, utilizing the "Forest-Based Classification and Regression" tool. The input parameters for the tool were established through the process of hyperparameter tuning, as elucidated in [paragraph 4.1](#), in the image has been highlighted with a red box.

Forest-based Classification and Regression

Parameters Environments

Prediction Type
Predict to raster

* Input Training Features

* Variable to Predict

☐ Treat Variable as Categorical

* Explanatory Training Rasters ☐ Categorical

* Output Prediction Surface

Match Explanatory Rasters
Prediction Training

Additional Outputs
Output Trained Features

Output Variable Importance Table

Advanced Forest Options
Number of Trees
100
Minimum Leaf Size
2
Maximum Tree Depth
34
Data Available per Tree (%)
100
Number of Randomly Sampled Variables

Validation Options
Training Data Excluded for Validation (%) 20
Number of Runs for Validation
2
Output Validation Table

☐ Calculate Uncertainty

Figure 25 - Forest-based Classification and Regression (ArcGis Pro)

This versatile tool fulfils a twofold role: it simplifies both dataset training and validation, while also empowering the development of a predictive model. Subsequently, this predictive model will undergo a comparison with observed Urban Heat Island (UHI) data maps obtained from satellites. This comparative analysis aims to unveil any disparities between the observed and calculated values, commonly referred to as residuals—a metric conveniently supplied by the tool itself.

To offer a comprehensive guide on completing the tool's fields effectively, the following explanations are provided:

- **Input Training Feature:** Within this section, it's essential to input a punctual-shaped shapefile containing the recorded UHI values from the observation campaign.
- **Variable to Predict:** In this segment, it's crucial to select the field associated with UHI values. The field's name corresponds to the attribute table entry in the shapefile.
- **Explanatory Training Raster:** This field necessitates the inclusion of raster files that correspond to variables contributing to model creation (e.g., wind speed.tif, temperature.tif, NDBI.tif, etc.).
- **Output Prediction Surface:** In this category, you must specify the path to save the raster prediction surface.
- **Match Explanatory Raster:** This section will be automatically populated by the tool.
- **Additional Output:** These fields require you to define the paths for saving both the trained feature and the variable importance table.
- **Advanced Forest Option:** In this section, enter the parameters derived from hyperparameter optimization using the Data Spell software.

After initiating the random forest process, its duration will vary based on the performance of the utilized machine. In this analysis, the following machine configuration was employed:

- **Processor:** 11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80GHz 2.80 GHz
- **RAM:** 16,0 GB
- **SSD:** 1TB
- **GPU Memory:** Intel IRIS Xe Graphics 8 GB

In this scenario, all the operations conducted within the Random Forest framework proved to be consistently time efficient. The duration of these operations ranged from a minimal time of one minute to a maximum of 25 minutes, as indicated by the table below:

	Elapsed time
Summer 2018	13 mins - 2 seconds
Mid-season 2018	13 mins - 2 seconds
Winter 2018	1 min - 42 seconds
Summer 2001	3 mins - 59 seconds
Mid-season 2001	25 mins - 36 seconds
Winter 2000	7 mins - 22 seconds

Table 12 - Elapsed time for different simulations

▪ **Hyper parameters settings:**

For a comprehensive perspective, all six simulations are presented collectively (encompassing summer, mid-season, and winter for both 2018 and 2001). The delineated process adheres to a sequential order, with the initial table detailing the hyperparameter settings:

	HYPER PARAMETERS						
	Number of Trees	Leaf size	Tree depth range	Mean Tree depth	% of training available per Tree	N° randomly sampled variables	% of training data excluded for validation
Summer 2018	200	1	31-31	31	100	5	20
Mid-season 2018	59	3	13-13	13	100	5	20
Winter 2018	60	5	25-25	25	100	5	20
Summer 2001	59	1	39-50	44	100	4	20
Mid-season 2001	164	1	38-44	43	100	4	20
Winter 2000	97	2	16-16	16	100	4	20

Table 13 - Hyper parameters

As mentioned earlier, the hyperparameters are derived from the outcomes obtained through Data Spell.

▪ **OBB model's performance:**

When it comes to regression tasks, there are several commonly used evaluation metrics to assess the performance of the predictive models. Some of the most frequently used regression evaluation metrics include:

- Mean absolute error (MAE)
- **Mean squared error (MSE)**
- **Root mean squared error (RMSE)**
- Root mean squared logarithmic error (RMSLE)
- Mean percentage error (MPE)
- Mean absolute percentage error (MAPE)
- **R²**

The choice of evaluation metrics falls into MSE, RMSE and R².

MSE = (True Value – Predicted Value)²;

RMSE = SQRT (MSE), [51] both allows for a comprehensive assessment of the model's predictive capabilities, considering both the magnitude and distribution of errors. The magnitude of the score errors remains within an acceptable range across all seasons.

	OBB MODEL (Out of Bag Errors)	
	MSE	RMSE
Summer 2018	0.023	0.152
Mid-season 2018	0.08	0.283
Winter 2018	0.021	0.145
Summer 2001	0.045	0.212
Mid-season 2001	0.017	0.130
Winter 2000	0.011	0.105

Table 14 - OBB model's performance (MSE and RMSE)

[51] Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.67). Abhishek Thakur. Edition Kindle.

The presented table showcases the model's varying accuracy across different seasons and years in predicting UHI phenomena. Notably, the model's performance is influenced by temporal factors, with varying levels of accuracy observed.

▪ **Coefficient of determination R^2 in training and validation phase**

“R-squared says how good your model fits the data. R-squared closer to 1.0 says that the model fits the data quite well, whereas closer 0 means that model isn't that good [52]”.

The formula for R-squared is shown here below:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_{ti} - y_{pi})^2}{\sum_{i=1}^N (y_{ti} - y_{tmean})^2} \quad [53]$$

Training	R^2	Statistical significance [p-value]	Standard error
Summer 2018	0.997	0.000	0.000
Mid-season 2018	0.948	0.000	0.001
Winter 2018	0.991	0.000	0.000
Summer 2001	0.994	0.000	0.000
Mid-season 2001	0.990	0.000	0.000
Winter 2000	0.853	0.000	0.001

Table 15 - Coefficient of determination R^2 in training phase

The analysis of the training phase reveals noteworthy R^2 coefficients, highlighting the model's effectiveness in explaining the variance within the dependent variable. The consistent statistical significance, represented by p-values consistently below 0.001, firmly establishes the robustness of the identified associations. Moreover, the modest standard errors underline the precision of the coefficient estimates. Importantly, the model consistently maintains a high level of precision across diverse seasonal contexts,

[52] Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.69). Abhishek Thakur. Edition Kindle.

[53] Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.69). Abhishek Thakur. Edition Kindle.

spanning from an impressive R^2 of 0.997 observed during the summer of 2018 to a substantial value of 0.853 in the winter of 2000.

Now, is shown the coefficient of determination in the validation phase.

Validation	R^2	Statistical significance [p-value]	Standard error
Summer 2018	0.980	0.000	0.001
Mid-season 2018	0.932	0.000	0.001
Winter 2018	0.955	0.000	0.001
Summer 2001	0.954	0.000	0.001
Mid-season 2001	0.925	0.000	0.002
Winter 2000	0.714	0.000	0.002

Table 16 - Coefficient of determination R^2 in validation phase

The validation phase table elucidates significant insights into the predictive performance of the model. The recorded R^2 coefficients serve as indicators of the model's aptitude in elucidating the variance within the response variable. Notably, the model achieves a commendable explanatory capability across various seasons, as evidenced by the R^2 values. The statistical significance of these associations, denoted by p-values consistently at or below 0.001, reinforces the robustness of the identified relationships. The marginal standard errors support the reliability of the coefficient estimations. It is noteworthy that the model exhibits consistent predictive precision across distinct temporal segments. This spans from a noteworthy R^2 of 0.980 in the summer of 2018 to the substantial value of 0.714 encountered during the winter of 2000.

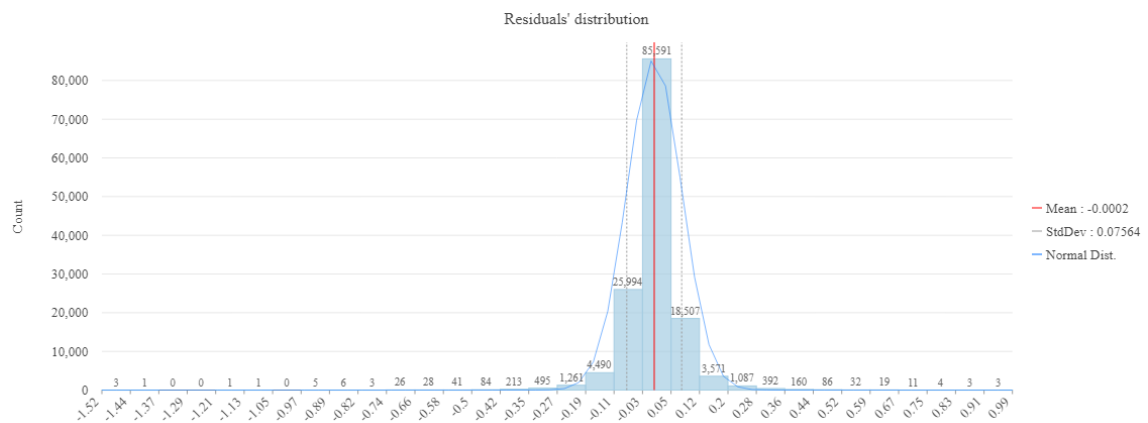
▪ **Residuals analysis:**

	Mean Residuals	St. dev of Residuals
Summer 2018	-0.00041	0.07864
Mid-season 2018	-0.00004	0.23776
Winter 2018	-0.0002	0.07564
Summer 2001	-0.00082	0.10517
Mid-season 2001	-0.00026	0.06821
Winter 2000	-0.00013	0.07934

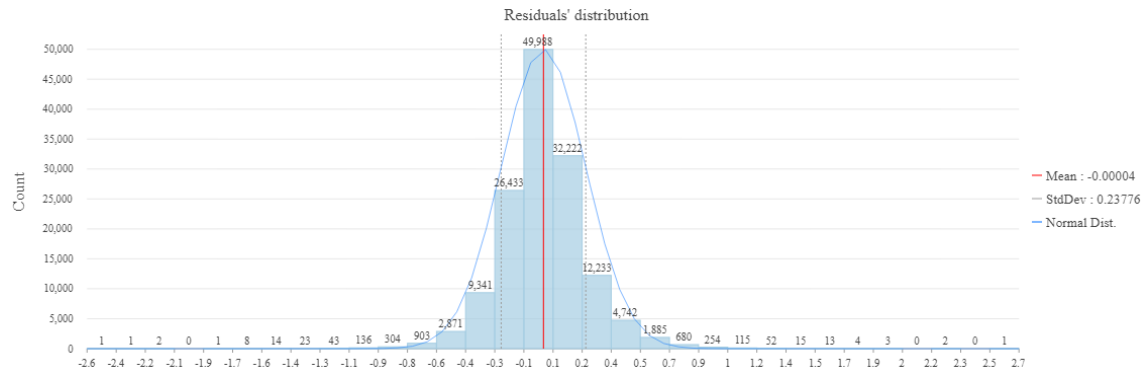
Table 17 - Residuals analysis (Mean and St.deviation)

Residuals, representing the differences between the observed and predicted values, are essential in assessing the effectiveness of a model's predictions. The calculated mean residuals indicate that, on average, the model's predictions closely align with the actual observed values. The proximity of mean residuals to zero underscores the model's capability to mitigate systematic biases, while the standard deviations of residuals provide a measure of the variability of errors around the mean. Notably, the small standard deviations of residuals across various seasons signify that the model's predictions maintain a relatively stable level of accuracy. This is evident in the marginal deviations from zero for mean residuals, suggesting that the model consistently achieves a balanced prediction performance across different timeframes. It is evident that the model's predictions are not only accurate on average but also exhibit consistent precision, as indicated by the narrow dispersion of residuals around their respective means.

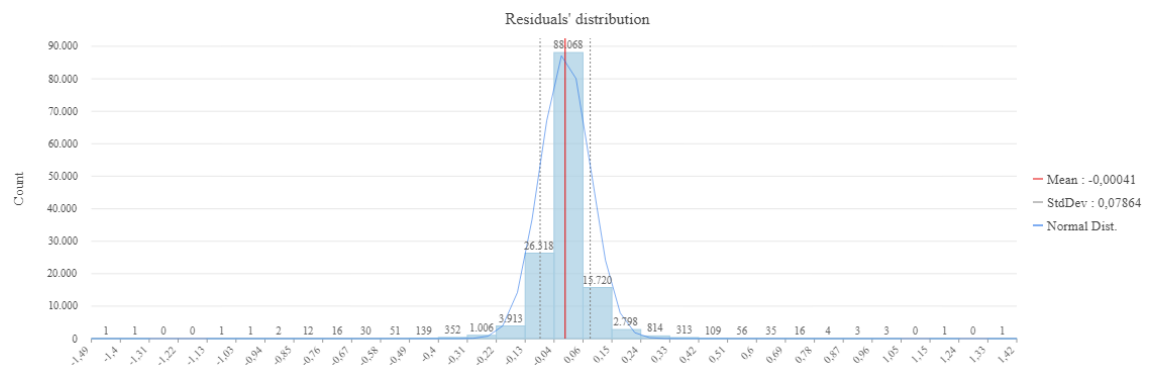
Winter 2018 Residuals' distribution graph:



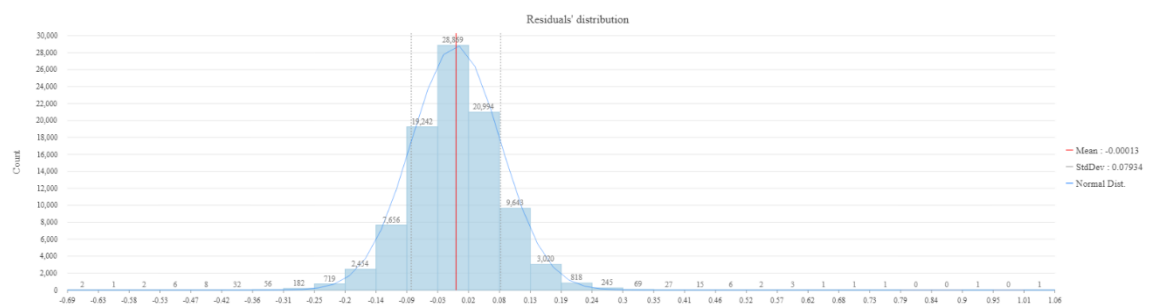
Mid-season 2018 Residuals' distribution graph:



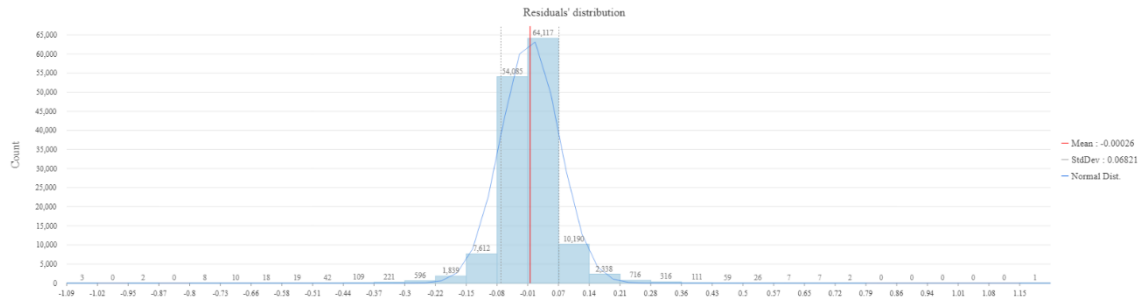
Summer 2018 Residuals' distribution graph:



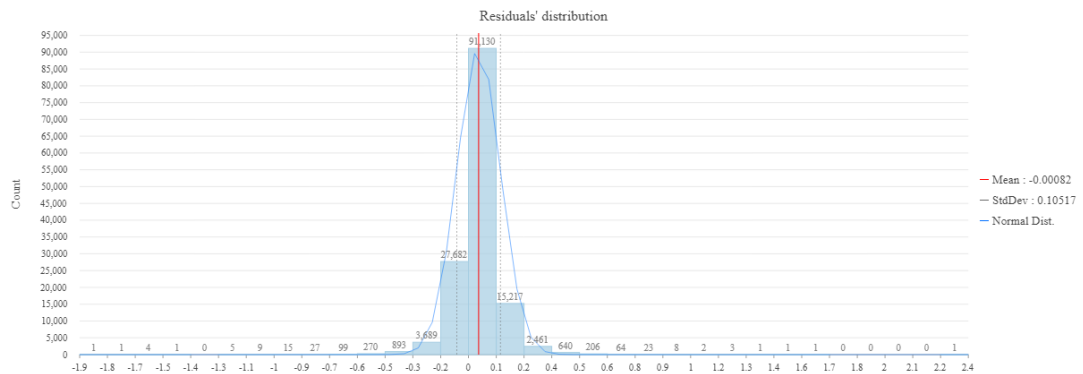
Winter 2001 Residuals' distribution graph:



Mid-season 2001 Residuals' distribution graph:



Summer 2001 Residuals' distribution graph:



The *residual values* obtained from the analysis are relatively small, indicating a *close alignment between the model's predictions and the observed data*. The mean residuals, ranging from -0.001 to 0.001, signify that, *on average, the model's predictions are highly accurate and closely match the actual observed values*. These small deviations from zero suggest that the model captures the underlying patterns of the urban heat island (UHI) effect with precision. Furthermore, the *standard deviations of residuals, typically falling within the range of 0.05 to 0.25*, are also relatively small. This suggests that the *differences between the predicted and observed values tend to cluster closely around the mean residual*. The *smaller standard deviations indicate consistent model*

performance across different seasons and years. In essence, the small residual values underscore the robustness of the UHI model's predictions.

▪ **Sensitivity analysis:**

Summer 2018	Variable	Importance	%
	NDMI	27652,59	25
	Buildings (geometry and human activities)	22908,52	21
	NDWI	10005,72	9
	NDVI	8896,15	8
	Wind speed	6756,35	6
	DTM	6140,77	6
	Emissivity	5940,83	5
	Solar radiation	3543,30	3
	Wind direction	3359,64	3
	Sky View Factor	3304,64	3
	SRI	3219,35	3
	S/V Ratio	3075,77	3
	Humidity	2316,41	2
	Temperature	2104,24	2
	Albedo	988,68	1
Mid-season 2018	Variable	Importance	%
	Buildings (geometry and human activities)	26503,93	25
	DTM	22336,14	21
	Sky View Factor	8598,10	8
	Emissivity	7220,27	7
	NDMI	4958,73	5
	NDVI	4937,59	5
	S/V Ratio	4575,53	4
	SRI	3857,27	4
	Albedo	3737,53	4
	Temperature	3662,86	3
	Wind speed	3430,94	3
	NDWI	3019,41	3
	Wind direction	2973,04	3
	Humidity	2891,29	3
	Solar radiation	2328,75	2
Winter 2018	Variable	Importance	%
	DTM	7934,46	20
	Sky View Factor	5801,25	15
	Buildings (geometry and human activities)	4659,05	12
	Humidity	3065,70	8
	Wind speed	2918,58	7
	NDMI	2590,64	7
	Wind direction	2294,09	6
	Temperature	2273,02	6
	Solar radiation	1937,59	5
	S/V Ratio	1869,04	5
	Albedo	1728,76	4
	SRI	639,46	2
	NDVI	469,13	1
	NDWI	439,92	1
	Emissivity	368,12	1

	Variable	Importance	%		Variable	Importance	%		Variable	Importance	%
Summer 2001	DTM	16748,36	19	Mid-season 2001	NDBI	4792,56	21	Winter 2000	Wind speed	430,58	13
	Solar radiation	11975,42	14		NDMI	3576,73	16		DTM	363,73	11
	NDBI	11733,79	13		SRI	2434,22	11		Albedo	308,42	10
	NDMI	8869,92	10		NDVI	2218,11	10		Humidity	306,32	10
	Albedo	8588,63	10		Emissivity	1854,66	8		SRI	287,59	9
	Temperature	7283,06	8		NDWI	1539,75	7		Temperature	259,66	8
	NDVI	5692,54	7		DTM	1194,88	5		Wind direction	250,93	8
	Emissivity	3840,87	4		Temperature	929,91	4		Solar radiation	234,62	7
	Wind direction	3713,77	4		Albedo	856,33	4		NDVI	179,52	6
	SRI	2682,92	3		Wind speed	819,37	4		NDWI	156,52	5
	NDWI	2149,92	2		Solar radiation	773,62	3		Emissivity	155,58	5
	Humidity	2052,98	2		Wind direction	762,47	3		NDMI	147,08	5
	Wind speed	2021,69	2		Humidity	747,69	3		NDBI	139,58	4

Table 18 - Sensitivity analysis for the six simulations (2001 and 2018)

The provided tables offer insights into the variable importance for each simulation within a given season. However, gaining a comprehensive understanding of the Urban Heat Island (UHI) phenomenon and identifying the most influential variables necessitates calculating the “*average aggregate importance*” of these variables across all six models (three models for 2018 and three for 2001). To achieve this, a Python code executed within the Data Spell environment computes and analyzes the “*mean aggregate importance*”, providing a more comprehensive view of the factors impacting UHI intensity:

```
import pandas as pd
import matplotlib.pyplot as plt

# Load CSV files for each simulation with semicolon separator
simulation1_df = pd.read_csv('C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/13_MODELLI_FINALI/24_08_2001/Sens
itivity_analysis/Sensitivity_summer_2001.csv', sep=';')
simulation2_df = pd.read_csv('C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/13_MODELLI_FINALI/27_05_2001/Sens
itivity_analysis/Sensitivity_midseason_2001.csv', sep=';')
simulation3_df = pd.read_csv('C:/Users/aless/Dropbox (Politecnico Di Torino
Studenti)/Tesi_HUI/1_PARTE_OPERATIVA/13_MODELLI_FINALI/18_12_2000/Sens
itivity_winter_2000/sensitivity_winter_2000.csv', sep=';')
```



```
# Combine DataFrames into a single DataFrame
all_simulation_df = pd.concat([simulation1_df, simulation2_df,
simulation3_df],
                              ignore_index=True)

# Calculate aggregate importance scores across simulations for each
variable
agg_importance =
all_simulation_df.groupby('Variable')['Importance'].mean().reset_index
()

# Rank variables based on aggregate importance scores
ranked_variables = agg_importance.sort_values(by='Importance',
ascending=True)

# Create a bar plot to visualize the results
plt.figure(figsize=(10, 6))
plt.barh(range(1, len(ranked_variables) + 1),
ranked_variables['Importance'])
plt.yticks(range(1, len(ranked_variables) + 1),
ranked_variables['Variable'])
plt.xlabel('Aggregate Importance')
plt.ylabel('Variable')
plt.title('Most Important Variables Across Simulations')
plt.tight_layout()

# Save the plot as an image file
plt.savefig('ranked_variables_plot.png')

# Display the plot
plt.show()
```

The values in the table are arranged in descending order, with the highest importance values presented first.

2018			2001		
Rank	Variable	Score	Rank	Variable	Score
1	Buildings (geometry and human activities)	18023.83	1	DTM	6102.323
2	DTM	12137.12	2	NDBI	5555.31
3	NDMI	11733.99	3	Solar radiation	4327.887
4	Sky View Factor	5901.33	4	NDMI	4197.91
5	NDVI	4767.623	5	Albedo	3251.127
6	Emissivity	4509.74	6	Temperature	2824.21
7	NDWI	4488.35	7	NDVI	2696.723
8	Wind speed	4368.623	8	Emissivity	1950.37
9	S/V Ratio	3173.447	9	SRI	1801.577
10	Wind direction	2875.59	10	Wind direction	1575.723
11	Humidity	2757.8	11	NDWI	1282.063
12	Temperature	2680.04	12	Wind speed	1090.547
13	Solar radiation	2603.213	13	Humidity	1035.663
14	SRI	2572.027			
15	Albedo	2151.657			

Table 19 - Aggregate mean importance of the variables

Sensitivity Analysis Comparison: Urban Heat Island (UHI) Model for 2001 and 2018:

Understanding the sensitivity of urban heat island (UHI) dynamics across different years provides crucial insights into the evolving relationship between urbanization, environmental variables, and local temperature variations. In this analysis, we explore the UHI sensitivity scores for the city of Turin during *two distinct periods: 2001 and 2018*.

These scores are based on the average aggregate points obtained from three simulations conducted across various seasons, including summer, winter, and mid-season.

2001 Analysis:

The UHI sensitivity analysis for 2001 provides an intriguing snapshot of UHI dynamics in the earlier urban landscape of Turin. Notably, "*DTM*" (*Digital Terrain Model*) takes a prominent place among the top-ranked variables. This underscores the

long-standing influence of elevation and topography in shaping temperature patterns within the city.

Furthermore, "*Solar Radiation*" emerges as a significant factor, signifying the role of energy input in driving temperature disparities.

The inclusion of "*NDBI*" (Normalized Difference Built-Up Index) in the top ranks points to the increasing impact of built-up areas on UHI during this period. This suggests evolving urbanization patterns and their implications for localized heat intensification.

Variables such as "*NDMI*," "*NDVI*," and "*Albedo*" emphasize the enduring importance of vegetation and surface properties in mitigating UHI effects. "*Temperature*" holds a significant place, but its position within the rankings suggests that other variables are equally instrumental in shaping UHI.

Meteorological variables like "*Wind Direction*," "*Wind Speed*," and "*Humidity*" underline their consistent role in UHI modulation. "*Emissivity*" and "*SRI*" reflect the ongoing impact of surface characteristics on heat retention and reflection.

2018 Analysis:

Comparatively, the UHI sensitivity analysis for 2018 offers insights into the UHI dynamics in a more recent urban context. Here, "*Buildings (geometry and human activities)*" emerges as the top-ranked variable, signifying the increasing influence of urban structural characteristics on UHI. This underscores the growing importance of urban planning and design in managing UHI effects.

Similar to 2001, "*DTM*" maintains its significance, showcasing the persistent role of topography in governing temperature variations. "*NDMI*" and "*NDVI*" continue to highlight the importance of vegetation and moisture content in mitigating UHI, aligning with the understanding of green spaces' impact on temperature moderation.

Variables like "*Emissivity*," "*Albedo*," and "*Solar Radiation*" underscore the role of surface properties and energy exchange processes in UHI development. Meteorological factors such as "*Wind Speed*," "*Wind Direction*," "*Humidity*," and "*Temperature*" reflect the intricate interplay between atmospheric conditions and localized temperature patterns.

Comparative Insights:

A compelling observation from these analyses is the shared importance of factors such as **vegetation, surface properties, and meteorological conditions in both years**. This suggests that certain underlying mechanisms shaping UHI effects remain relatively consistent across different urban contexts.

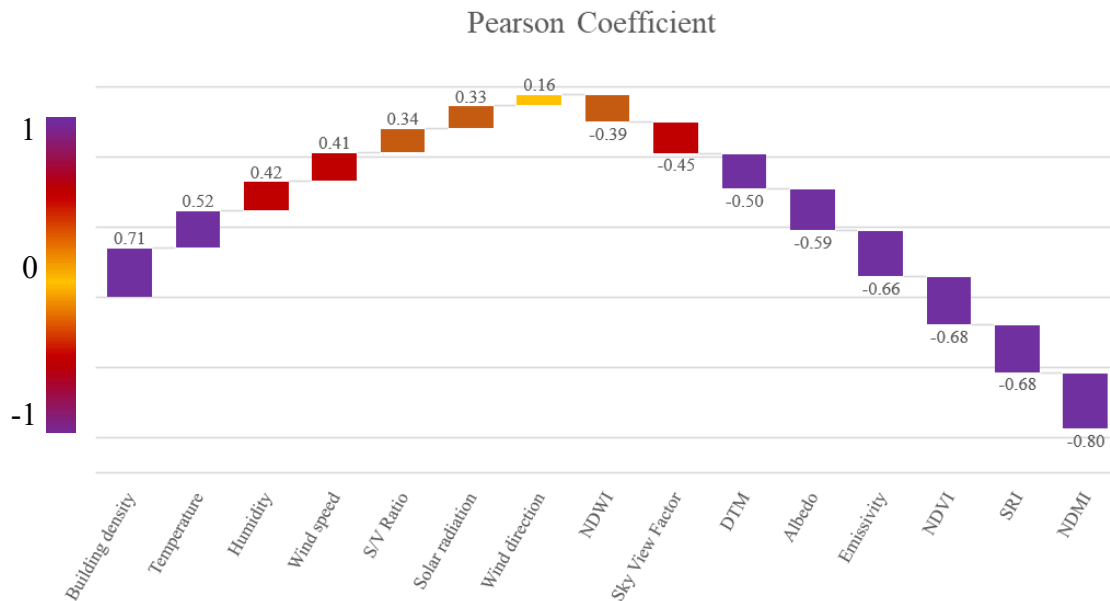
In conclusion, these sensitivity analyses underscore the intricate interactions among urban form, land use, vegetation, surface characteristics, and atmospheric conditions that collectively drive UHI. As cities evolve, these insights can guide strategies to mitigate UHI effects, fostering more sustainable and resilient urban environments.

▪ **Predicted vs Calculated [graphical results]**

In our study, we employed a Random Forest algorithm to **calculate and predict UHI intensity**, yielding intriguing results. First, we can observe in the following maps that the magnitude of UHI intensity remained consistent across calculations and predictions, indicating robustness in the algorithm's performance. However, a slight divergence in the classification emerged, leading to **variations in spatial delineation**. This discrepancy implies that while the overall UHI effect remains consistent, the **algorithm may capture subtle distinctions in UHI distribution**.

Furthermore, a **pronounced seasonality in UHI intensity was identified**. The warmer months displayed heightened UHI intensity, whereas midseason transition periods experienced a reduction in intensity. Surprisingly, UHI intensity exhibited a notable decrease during the winter months, particularly within built-up areas. This observation suggests complex interactions between urban morphology, land cover characteristics, and climatic influences (as snow), beyond just heating-related effects.

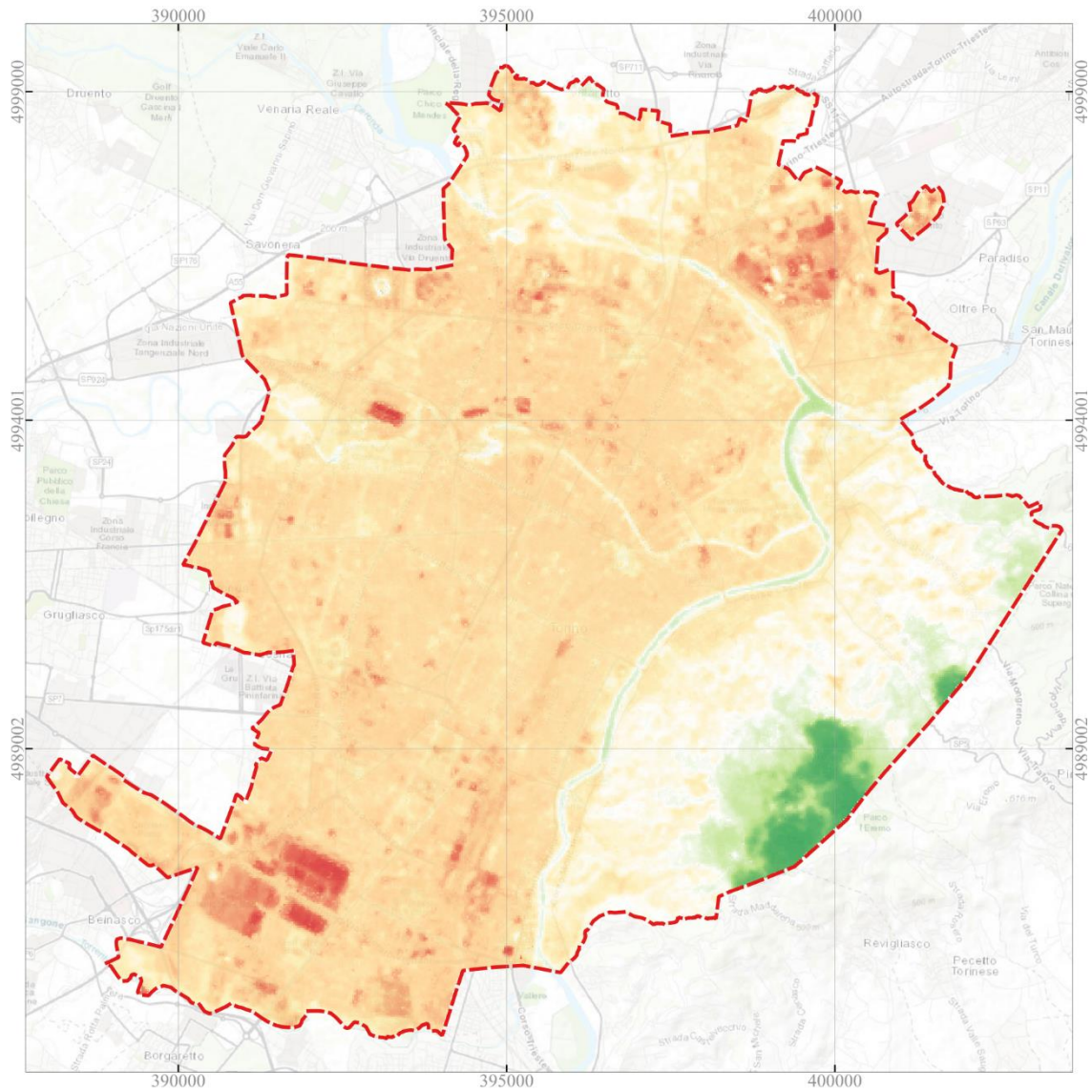
Additionally, the consistent UHI intensity values observed on the winter map, approximating unity, **raise questions about the role of snow cover for December 2000**.



The Pearson correlation coefficients presented in the provided graph represent the degree and direction of linear relationships between various environmental variables and the **Urban Heat Island (UHI) effect**. These coefficients quantify the strength and direction of linear relationships between each variable and the UHI effect. Positive coefficients indicate a positive linear correlation, while negative coefficients indicate a negative linear correlation. The magnitude of the coefficient reflects the strength of the relationship. Pearson correlation coefficients are specifically designed to measure linear correlations between two variables, providing a single numerical value representing the strength and direction of that linear relationship. In contrast, random forest importance scores are used in machine learning to assess the relative importance of variables for predictive modelling, considering both linear and nonlinear relationships and not providing a measure of correlation strength. Integrating the information from both tables, the Pearson correlation coefficients graph and the Random Forest importance scores table, serves a dual purpose. Firstly, it aids in prioritizing variables that are not only significantly correlated with the Urban Heat Island (UHI) effect but also exhibit high importance scores in predictive modelling. Such variables are of particular interest because they are likely influential in driving the UHI effect and hold potential for targeted interventions.

Secondly, this integration offers valuable contextual insights. It helps researchers delve beyond mere correlation strength and considers how variables interact in complex, nonlinear ways. For instance, a variable might display a weak linear correlation with the UHI effect according to Pearson coefficients but receive a high importance score in the random forest model. This suggests that although the linear relationship is weak, the variable plays a critical role in predicting the UHI effect due to intricate nonlinear interactions.

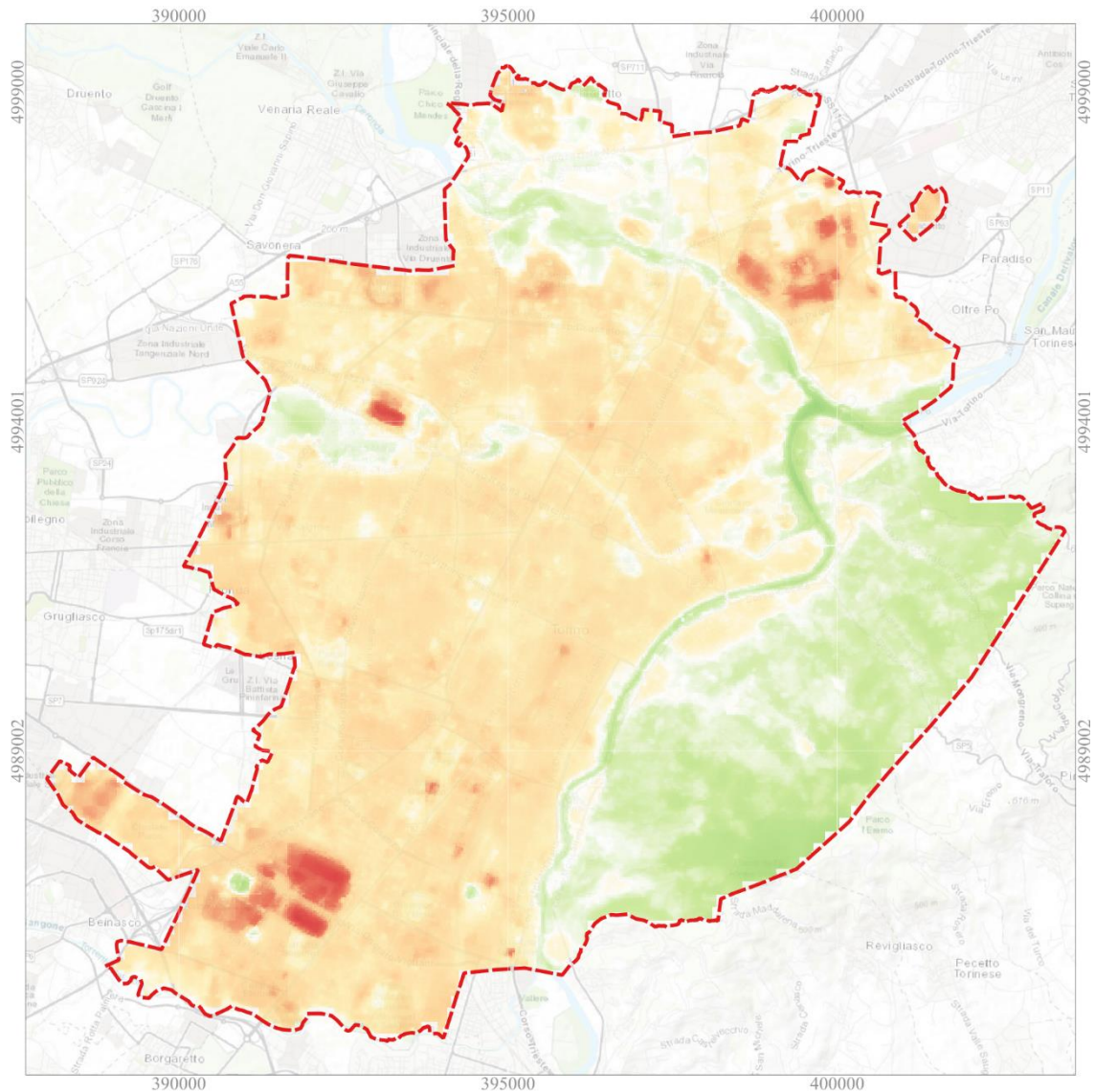
UHI Intensity - RF prediction – [24/08/2001 – SUMMER]



UHI Prediction [24/08/2001]

- 2 [MEDIUM-LOW Cooling Effect]
- -1 [LOW Cooling Effect]
- 0 [AREA OF BALANCE]
- 1 [LOW Heating Effect]
- 2 [MEDIUM-LOW Heating Effect]
- 3 [HIGH Heating Effect]
- >4 [VERY-HIGH Heating Effect]

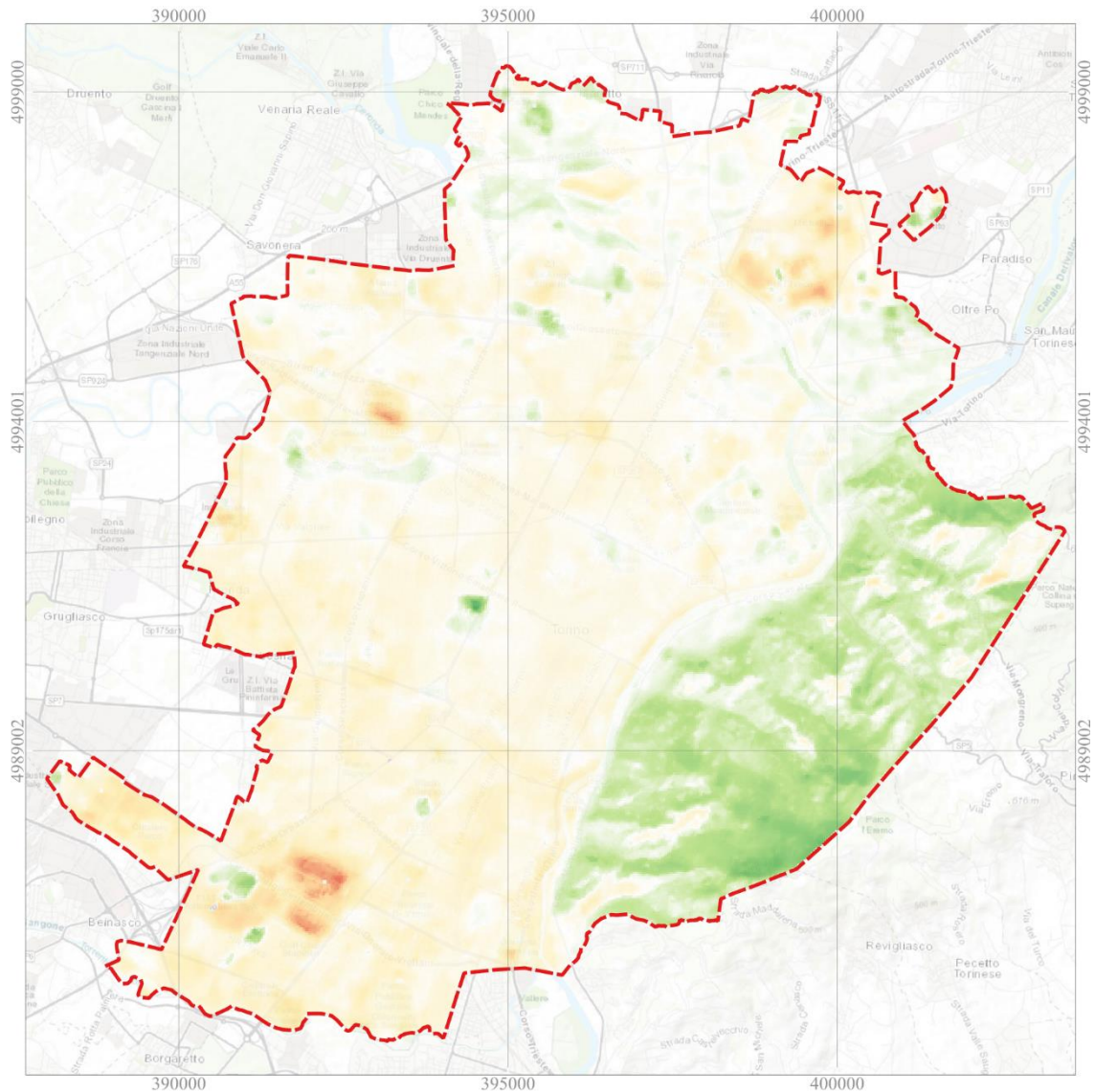
UHI Intensity - RF prediction – [22/08/2018 – SUMMER]



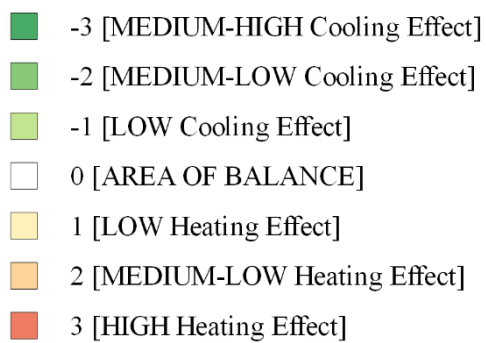
UHI Prediction [22/08/2018]

- 2 [MEDIUM-LOW Cooling Effect]
- -1 [LOW Cooling Effect]
- 0 [AREA OF BALANCE]
- 1 [LOW Heating Effect]
- 2 [MEDIUM-LOW Heating Effect]
- 3 [HIGH Heating Effect]
- >4 [VERY-HIGH Heating Effect]

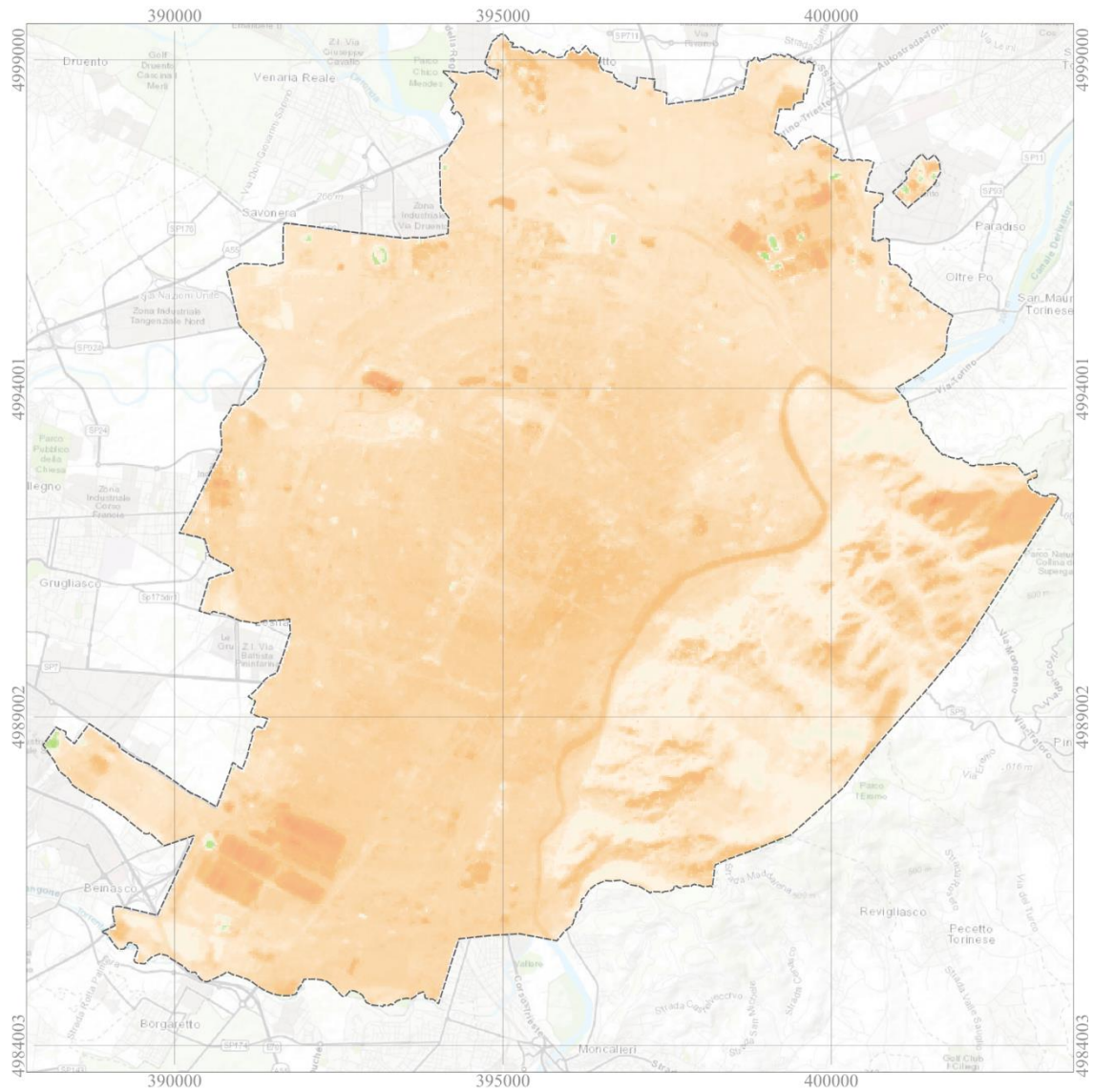
UHI Intensity - RF prediction – [11/02/2018 –WINTER]



UHI Prediction [11/02/2018]



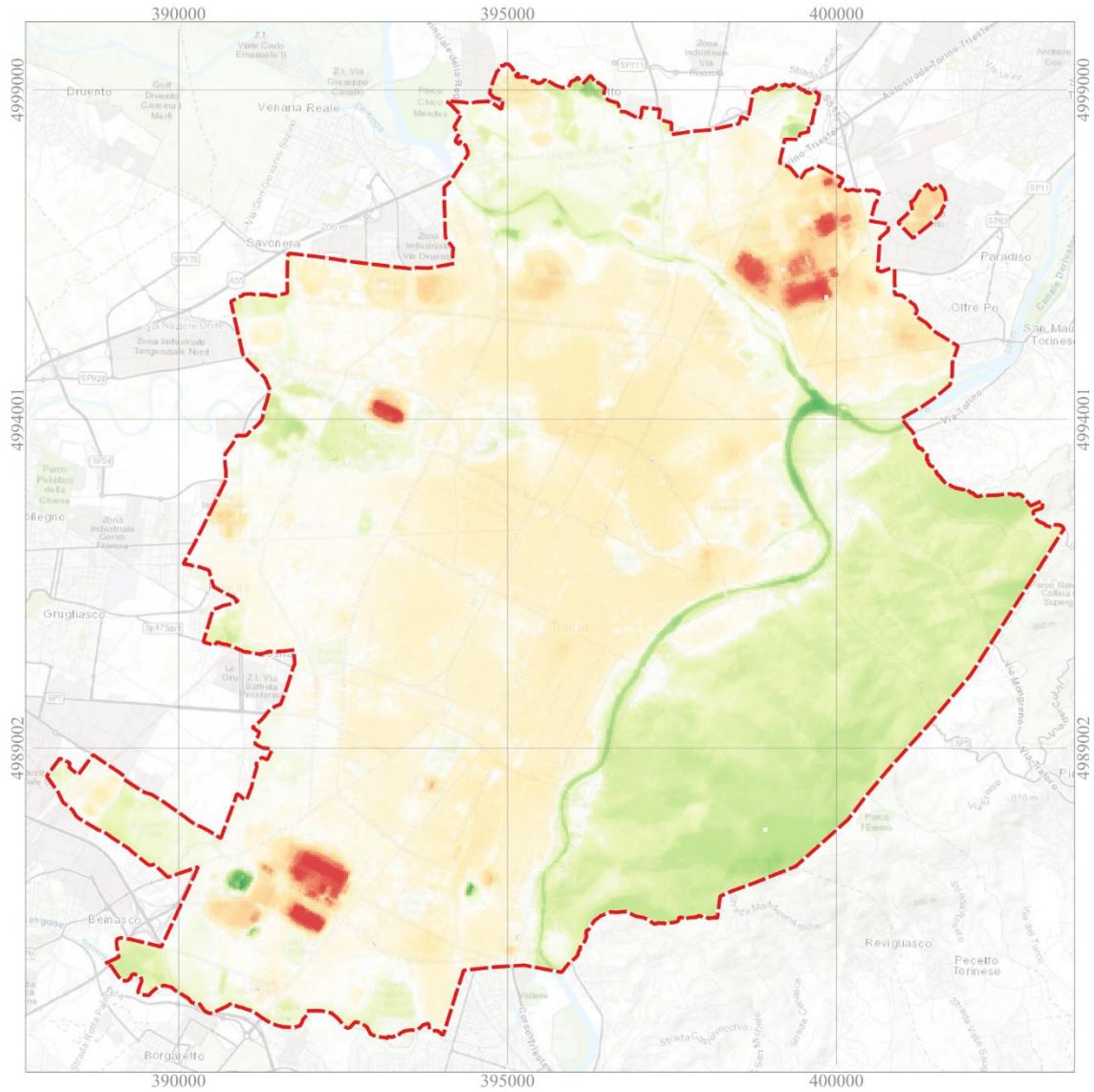
UHI Intensity - RF prediction – [18/12/2000 – WINTER]



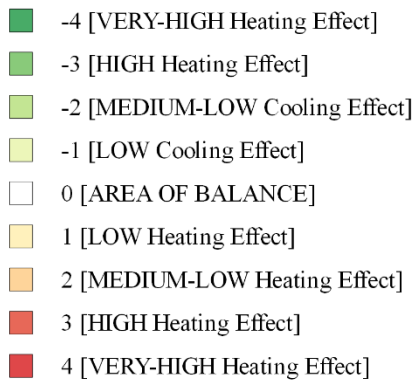
UHI Predicted [18/12/2000]

- 1 [MEDIUM-LOW Cooling effect]
- 0,5 [LOW Cooling effect]
- 0 [AREA OF BALANCE]
- 1 [MEDIUM-LOW Heating Effect]
- 2 [MEDIUM Heating Effect]

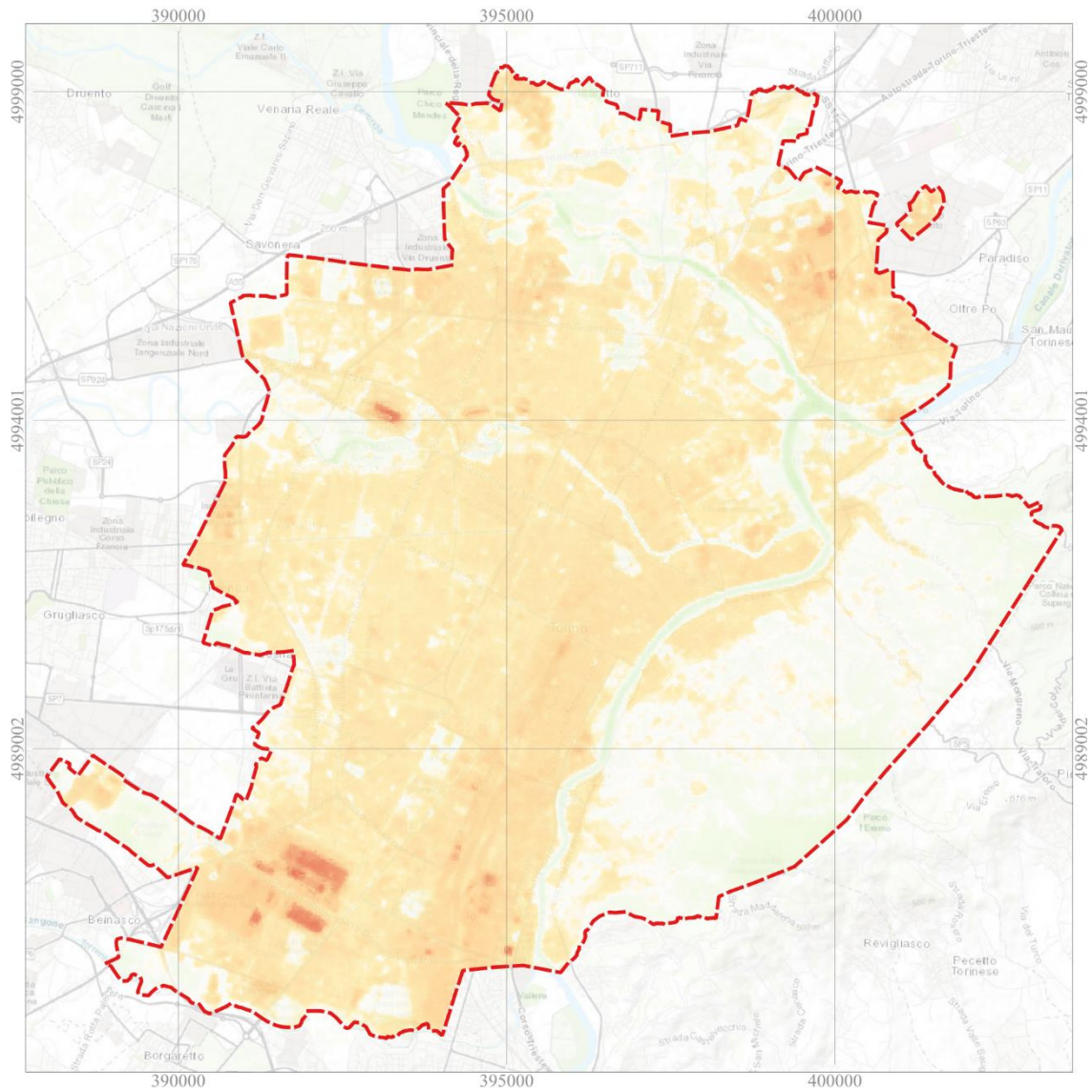
UHI Intensity - RF prediction – [16/04/2018 – MIDSEASON]



UHI Intensity [16/04/2018]



UHI Intensity - RF prediction – [2705/2001 – MIDSEASON]



UHI Intensity [27/05/2001]

- 0,5 [VERY-LOW Cooling Effect]
- 0 [AREA OF BALANCE]
- 1 [LOW Heating Effect]
- 2 [MEDIUM-LOW Heating Effect]
- 3 [HIGH Heating Effect]

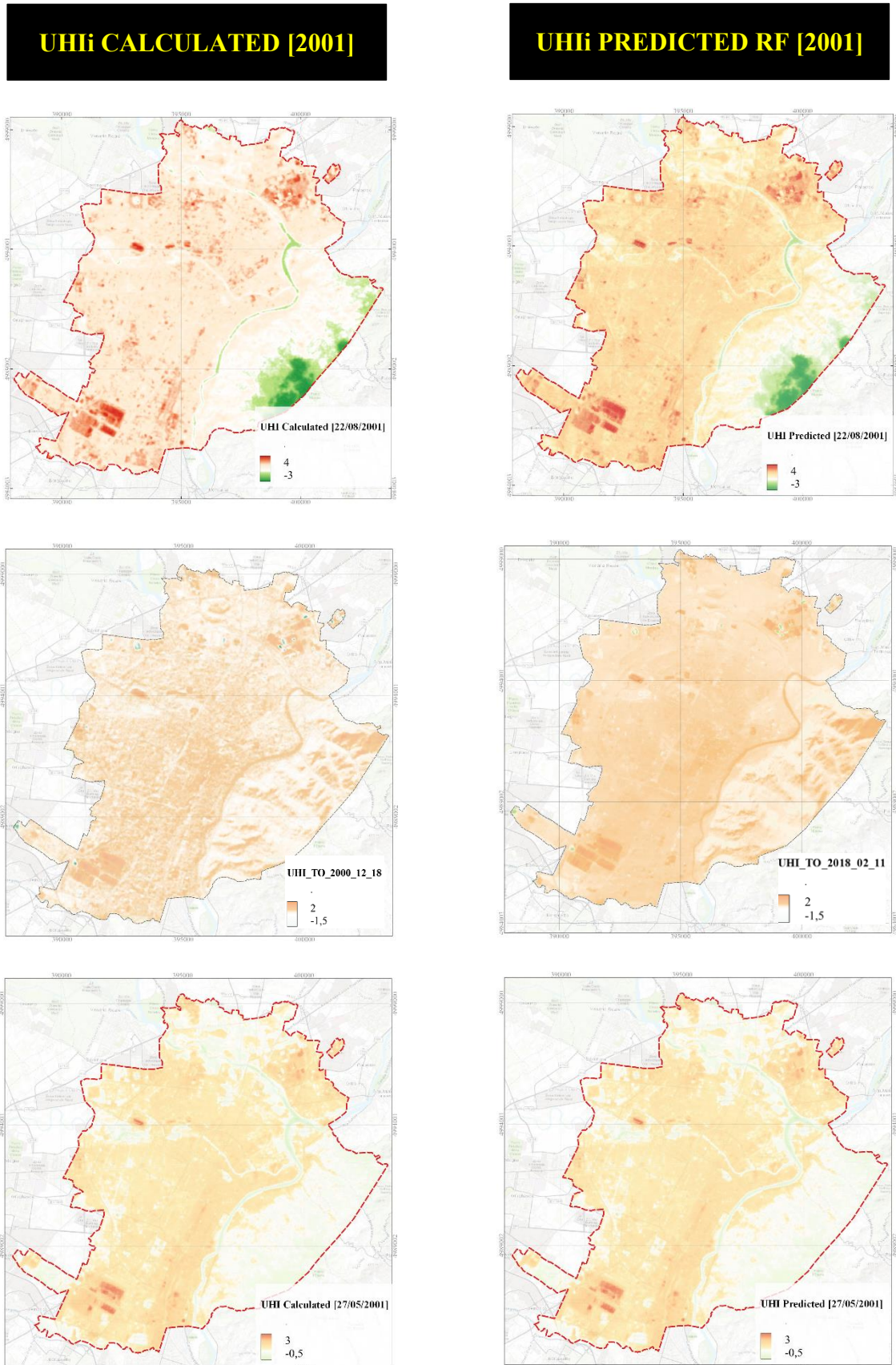


Figure 26 - Calculated vs Predicted RF UHI [2001]

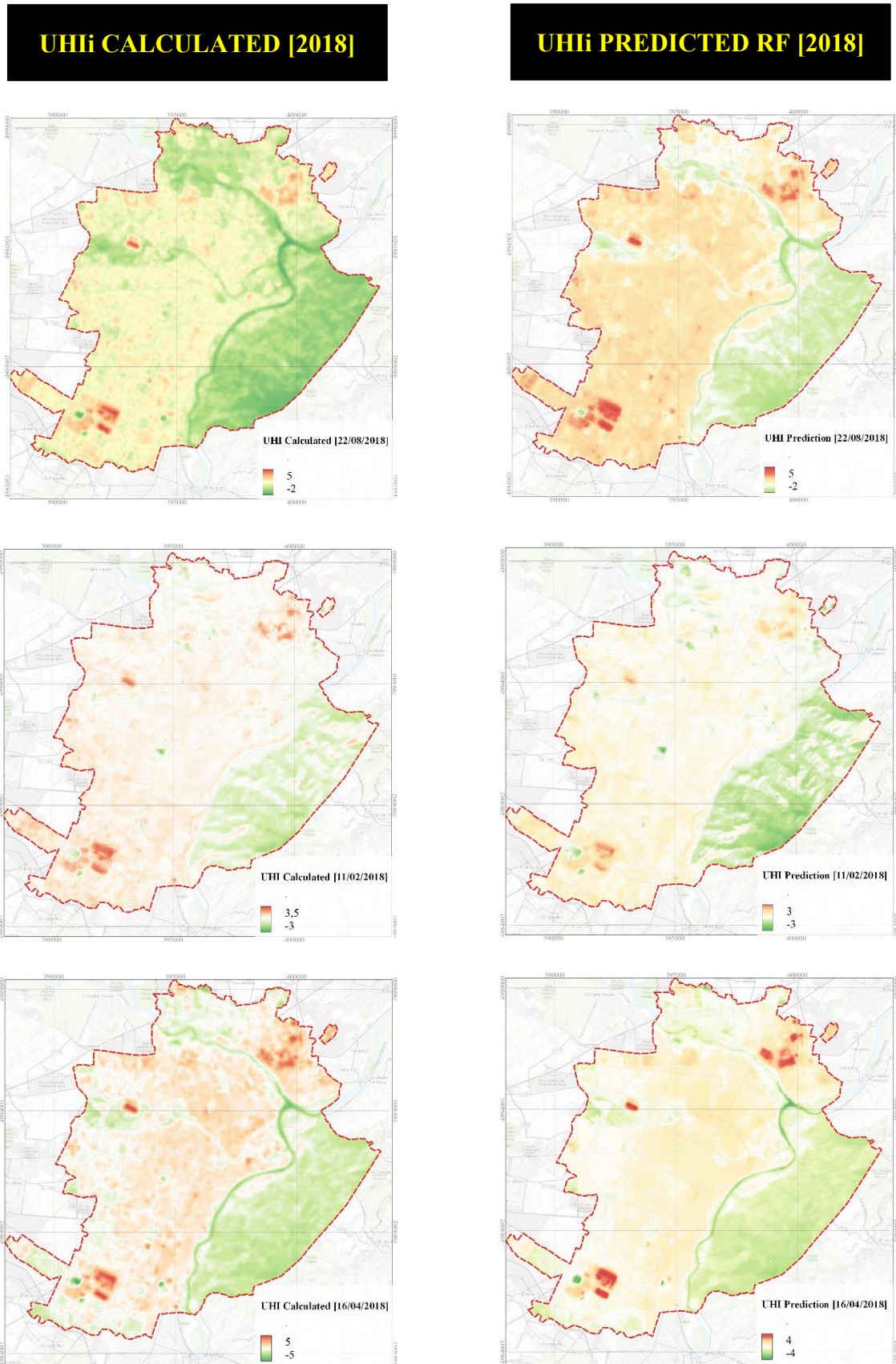


Figure 27 - Calculated vs Predicted RF UHI_i [2018]

5 Urban regeneration & UHI

The main thrust of our research was to harness machine learning techniques to quantify Urban Heat Island (UHI) intensity by deploying a multifactorial model encompassing variables recognized as significant contributors to this climatic phenomenon within the urban context of Turin. Building upon foundational studies in urban climate science (Oke, 1982) [54], our investigation sought to illuminate the complex interplay of factors influencing UHI. Through meticulous evaluation of UHI dynamics across multiple seasons during the pivotal years of 2001 and 2018, we aimed to provide a comprehensive spatiotemporal understanding of UHI variation within the Turin metropolitan area. This research also grappled with a pivotal **second research question, *probing the potential of extensive urban regeneration interventions to ameliorate UHI intensity in the locales where they are implemented.*** Drawing upon seminal work in urban planning and sustainability, our inquiry into the mitigative capacity of urban regeneration interventions holds promise in enhancing our comprehension of UHI management strategies and their broader implications for sustainable urban development.

Turin, with its unique history of **transitioning from an industrial powerhouse to a post-industrial city**, provided an intriguing case study. The city's industrial past, marked by heavy manufacturing and associated urban infrastructure, underwent a significant transformation starting in the early **2000s as industries were disinvested and deindustrialization processes began.**

Our research, therefore, explored how these dynamic urban changes, including the dismissal of industry, may have influenced UHI dynamics.

The case, for the assessment of the changes in UHI magnitude, is **the Parco Dora area [EX-Teksid]**, in the following paragraph we will delve into the **results obtained by the model to confirm the research question.**

[54] Oke, T. R. (1982). The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society, 108(455), 1-24

5.1 UHI magnitude before and after the great intervention

Urban regeneration, within the context of UHI mitigation, represents a multifaceted urban planning and development approach underpinned by scientific principles. The Urban Heat Island (UHI) effect, characterized by elevated temperatures within urban areas, is primarily driven by changes in land use, land cover, and heat-absorbing surfaces.

When comparing the impact of a great urban regeneration intervention on UHI between two years, 2001 and 2018, **it's essential to choose seasons that are representative and can highlight potential changes**. Summer is typically the season when UHI intensity is most pronounced due to higher temperatures and increased energy consumption.

Climate change, a prominent global driver, **has the potential to influence local climate patterns significantly** (Intergovernmental Panel on Climate Change, 2014) [55].

Concurrently, urbanization and the expansion of impervious surfaces can intensify the Urban Heat Island (UHI) effect, contributing to elevated LST values within urban areas (Oke, 1982) [56]. Changes in land use, surface materials, can further exacerbate UHIs, affecting the local thermal environment (Voogt & Oke, 2003; Shashua-Bar & Hoffman, 2009) [57].

These complex interactions illustrate the importance of considering not only air temperature but also broader climatic and urbanization factors when analysing LST variations over time.

[55] Intergovernmental Panel on Climate Change (IPCC). (2014). Climate Change 2014: Synthesis Report. Retrieved from <https://www.ipcc.ch/report/ar5/syr/>

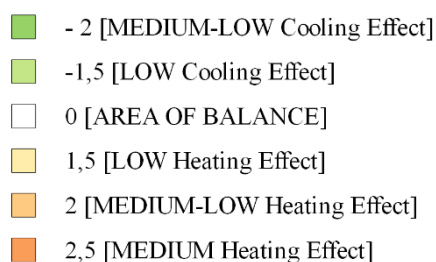
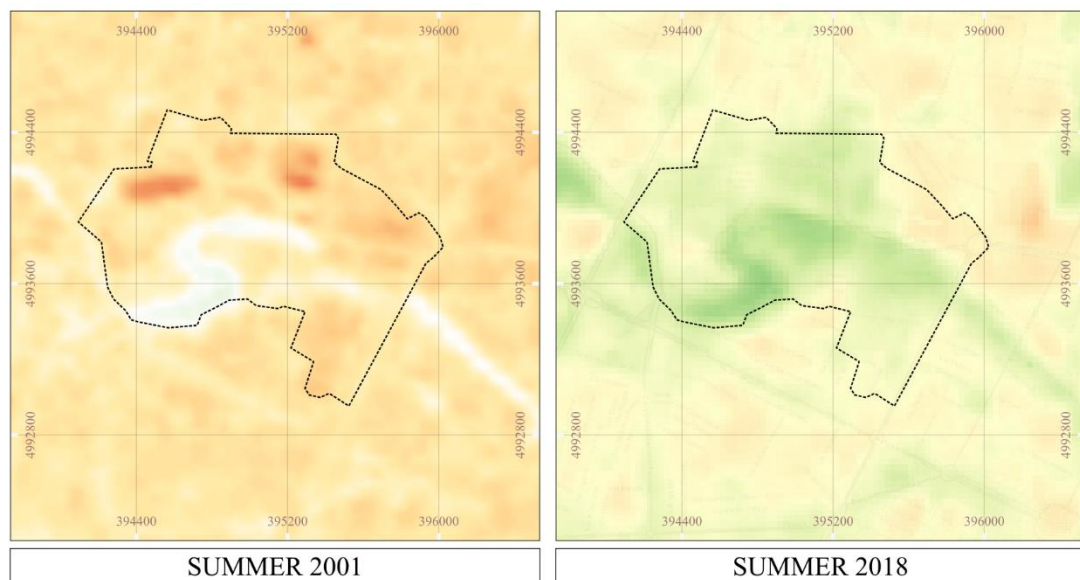
[56] Oke, T. R. (1982). The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society, 108(455), 1-24

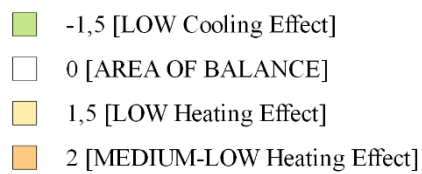
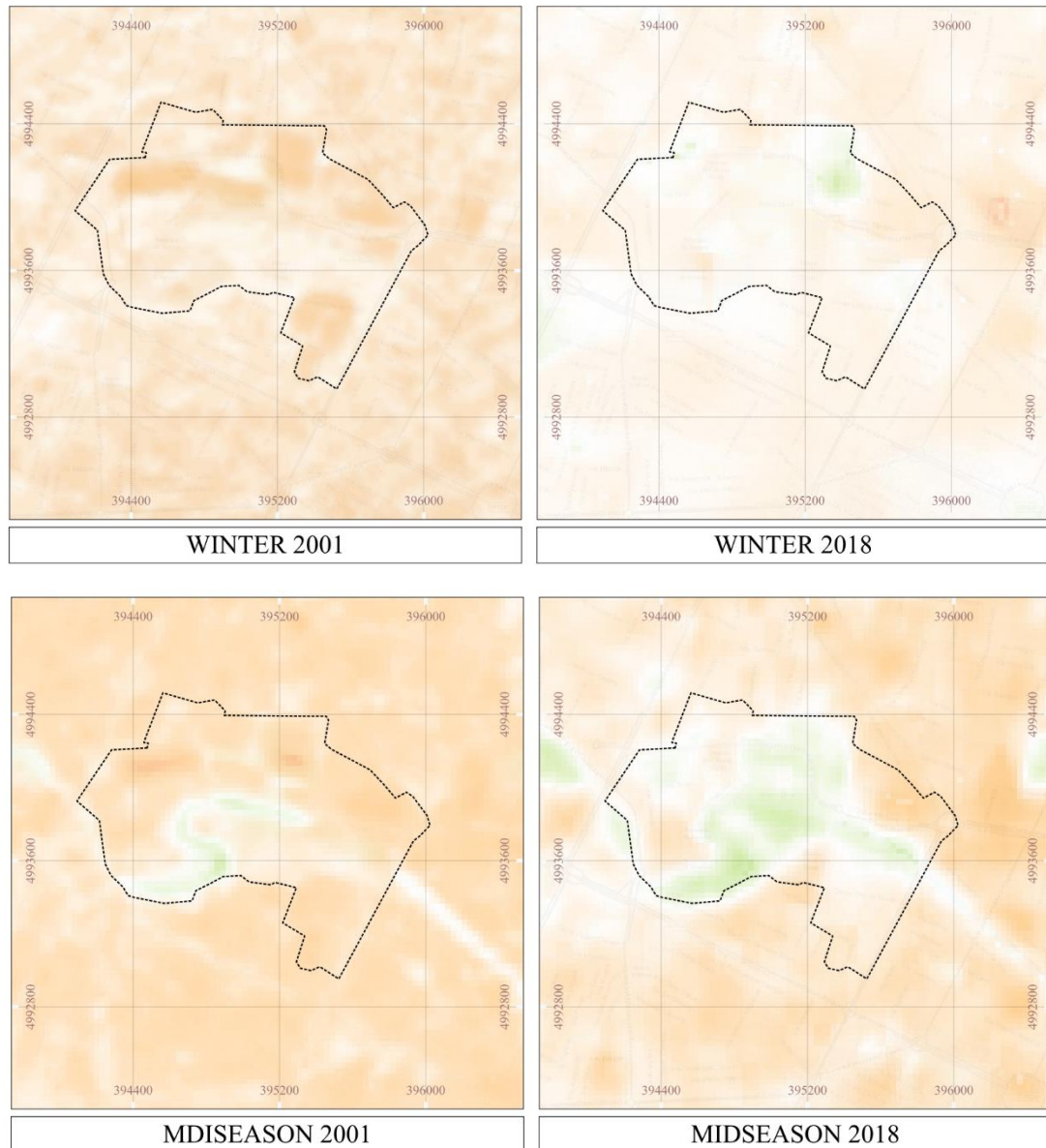
[57] Ibid [53]

The evaluation of Urban Heat Island (UHI) intensity serves as a vital metric for comprehending the thermal dynamics of urban environments. In the context of Turin's EX-area Teksid, a region that underwent substantial urban regeneration efforts between 2001 and 2018, understanding changes in UHI intensity provides critical insights into the efficacy of urban planning interventions.

By quantifying this intensity before and after urban regeneration, we aim to ascertain whether these efforts have led to a reduction in temperature disparities and a more thermally sustainable urban environment. Such assessments are instrumental in gauging the success of urban revitalization projects in mitigating the UHI effect and fostering more comfortable urban spaces.

“What alterations have occurred in the Urban Heat Island (UHI) intensity from 2001 to 2018?”




















Variable	Area of analysis	Distance's radius from the intervention [m]	Intensity Variation Summer (2001-2018)	Intensity Variation Spring (2001-2018)	Intensity Variation Winter (2001-2018)	Total mean [absolute value]
UHI (Intervention Area statistics)		0	 -0.944	 -1.037	 -0.544	 -0.841
UHI (District statistics)		500	 -0.413	 -0.524	 -0.476	 -0.471
UHI (Municipal statistics)		All Turin's municipal area	 -0.500	 -0.849	 -0.492	 -0.614

Table 20 - UHI Variation 2001-2018

The evaluation of Urban Heat Island (UHI) intensity within Turin's EX-area Teksid over the period from 2001 to 2018 reveals a **noteworthy and positive transformation**. The data analysis indicates a **substantial decrease in UHI intensity during this time frame**, suggesting the effectiveness of urban regeneration initiatives in mitigating thermal disparities. Notably, this reduction in UHI intensity is not uniform across all seasons. During summer months, when UHI effects are typically most pronounced, the decline is particularly noteworthy, signifying a more substantial cooling effect in response to these urban interventions. Conversely, winter and midseason (spring) show milder decreases, highlighting the complexity of UHI dynamics in different seasons and the potential for further strategies tailored to specific climatic conditions. These findings underscore the importance of considering seasonality in urban planning and climate resilience efforts within dynamic urban environments.

5.2 Potential insights: Combating UHIs through Urban Planning for Climate-Resilient Futures

As urbanization accelerates, cities face the Urban Heat Island (UHI) effect—a rise in temperatures within urban areas. This *phenomenon has far-reaching implications for health, energy, and city life*. Urban planning is essential to combat UHI, as seen in sensitivity analyses of Turin for 2001 and 2018. By understanding key variables, we unveil actionable strategies for cooler cities. Climate change is intertwined with UHI, emphasizing the link between global climate shifts and local temperature disparities.

Urban planning's role aligns cities with climate action, fostering sustainable, adaptable, and vibrant urban environments.

1. Urban Planning Strategies to Mitigate Urban Heat Island (UHI) Effect:

1. ***Increase Green Spaces and Vegetation:*** The consistent prominence of variables like "NDVI" and "NDMI" in both years underscores the *importance of green spaces and vegetation*. Urban planners should prioritize incorporating parks, green roofs, and trees to provide shade, reduce heat absorption, and encourage evaporative cooling (as we see in the case of Ex-TEKSID Area).



Figure 28 - Parco Dora renewal, Ex Area TEKSID [58]

[58] Comune di Torino: Parco Dora: aperti al pubblico il lotto Valdocco Nord e Iron Valley – Verde Pubblico (comune.torino.it)

2. **Promote Sustainable Land Use:** The emergence of "*Buildings (geometry and human activities)*" as a top-ranking variable in 2018 indicates the *impact of urban form*. Urban planning should *encourage mixed land use*, compact development, and high-density areas to minimize heat-retaining surfaces while optimizing energy efficiency.
3. **Design Reflective Surfaces:** "*Albedo*" and "*SRI*" highlight the role of *surface reflectivity*. Urban planners can encourage the use of reflective and cool roofing materials, light-coloured pavements, and other strategies to reduce heat absorption.

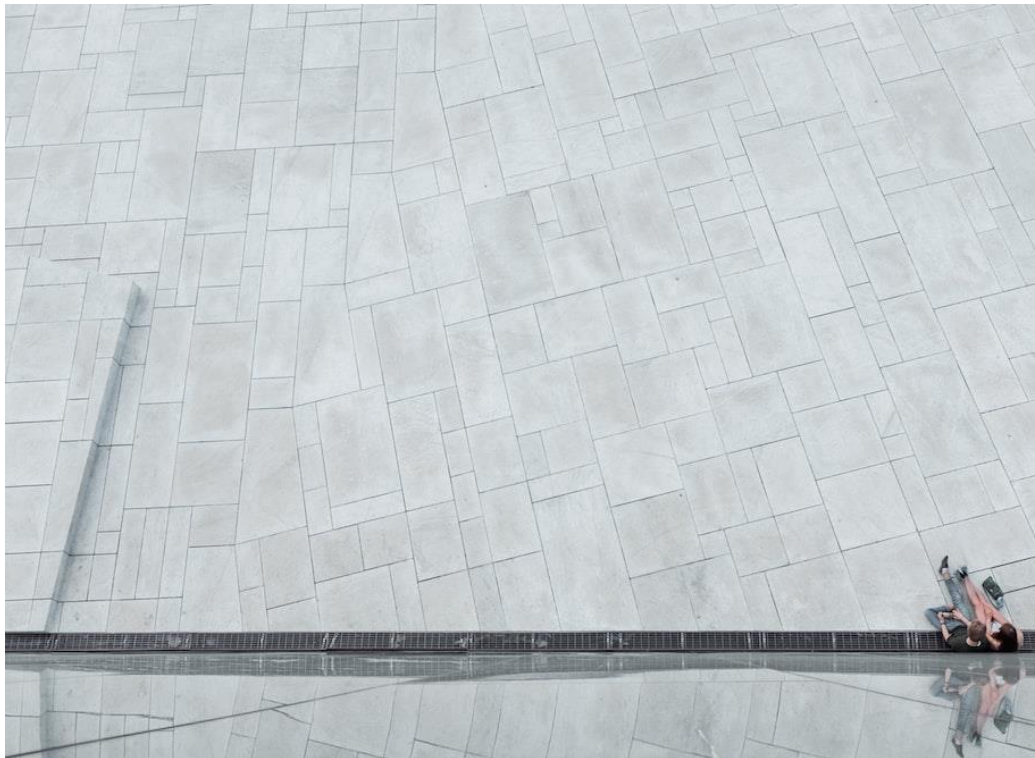


Figure 29 – Oslo Opera House [59]

4. **Implement Cool Roofs and Pavements:** Building codes and regulations can require cool roofs and pavements that reflect more sunlight and absorb less heat, contributing to localized cooling effects.

[59] Oslo Opera House, Oslo, Norway, Daniel Ghio, Unsplash:
<https://unsplash.com/it/foto/LHf1WVqdmCI>

5. **Promote Energy-Efficient Building Design:** Incorporate energy-efficient building designs that reduce the need for air conditioning, thus decreasing local heat generation.
6. **Urban Forestry Management:** Invest in urban forestry management to ensure healthy tree populations, strategically placed to provide shade and cooling effects.



Figure 30 - Yuan Ching Road, Jurong Lake Gardens, Singapore [60]

2. Importance of Climate Change in Relation to UHI:

- **Amplification of UHI:** As global temperatures rise due to climate change, UHI effects are amplified. Urban areas experience higher baseline temperatures, intensifying UHI impacts and making urban residents more vulnerable to extreme heat events.
- **Positive Feedback Loop:** UHI contributes to a positive feedback loop where higher temperatures in urban areas lead to increased energy consumption for cooling, further contributing to greenhouse gas emissions.

[60] Unsplash: <https://unsplash.com/it/foto/bIpKSEsaN6Q>

- ***Adaptation and Resilience:*** Urban areas need to adapt to changing climate patterns. Strategies to counter UHI effects align with broader climate adaptation efforts, ensuring cities are resilient in the face of rising temperatures.
- ***Health and Equity Concerns:*** UHI poses health risks, particularly to vulnerable populations. Elderly, children, and low-income communities are more susceptible to heat-related illnesses. Addressing UHI is part of promoting equitable urban environments.
- ***Mitigation and Global Efforts:*** Mitigating UHI contributes to broader global climate mitigation efforts. Cooler cities reduce energy consumption, which in turn reduces greenhouse gas emissions from power generation.
- ***Collaborative Approach:*** Combating UHI due to climate change requires a collaborative effort among urban planners, policymakers, researchers, and communities. Local policies and initiatives need to align with regional and global climate goals.

In conclusion, addressing the urban heat island effect through urban planning strategies is essential for creating liveable, sustainable cities. Climate change intensifies UHI impacts, making adaptation and mitigation crucial components of urban development strategies. By implementing innovative urban planning measures and understanding the interconnectedness between UHI and climate change, cities can contribute to a more resilient and sustainable future.

3. **Similarities and Mitigation Strategies in Singapore:**

One city that has effectively mitigated the urban heat island (UHI) effect and faces similar challenges as Turin is Singapore. *Singapore is known for its tropical climate, rapid urbanization, and efforts to create a sustainable and liveable urban environment.* Despite its distinct climate, Singapore's strategies to address UHI can offer valuable insights for Turin and other cities facing similar issues.

1. ***Green Spaces and Vegetation:*** Like Turin, Singapore has prioritized the incorporation of green spaces, parks, and rooftop gardens. The city's "*City in a Garden*" initiative focuses on enhancing urban greenery to provide shade and reduce heat absorption.
2. ***Cool Roof Initiatives:*** Singapore has promoted the use of cool roofs, which reflect sunlight and absorb less heat. This helps to maintain lower rooftop temperatures and reduce overall UHI impact.
3. ***Urban Planning and Mixed Land Use:*** Both cities emphasize sustainable land use planning to optimize energy efficiency and reduce the heat island effect. Compact development and mixed land use encourage efficient use of space and reduce heat-absorbing surfaces.
4. ***Vegetated Infrastructure:*** Singapore's innovative approach includes vegetated facades, vertical gardens, and green corridors. These elements enhance urban aesthetics, biodiversity, and cooling effects.
5. ***Water Features and Blue-Green Infrastructure:*** Both cities have incorporated water features, such as ponds, lakes, and canals, to cool the environment through evaporative cooling. Singapore's "ABC Waters" program integrates water management with urban planning.
6. ***Climate-Responsive Architecture:*** Singapore's architecture incorporates climate-responsive designs that optimize natural ventilation, shading, and energy efficiency, aligning with UHI mitigation.
7. ***Research and Innovation:*** Singapore's research institutions collaborate with the government and private sector to develop cutting-edge solutions for UHI mitigation. This approach encourages ongoing innovation in addressing urban climate challenges.

While Singapore's climate differs from Turin's, its emphasis on green infrastructure, energy-efficient designs, and integrated planning offers transferable lessons for UHI mitigation.



Figure 31 - Singapore National Parks Board (NParks) [61]

While Singapore pursued such a strategy, the city of *Turin* ***opted for a similar approach, implementing two distinct strategic plans to enhance its resilience in the face of climate change.***

1. [Il Piano Strategico dell'Infrastruttura Verde Torinese](#)
2. [Piano di Resilienza Climatica](#)

[61] Singapore National Parks Board (NParks): [Parks & Nature Reserves - Gardens, Parks & Nature - National Parks Board \(NParks\)](#)

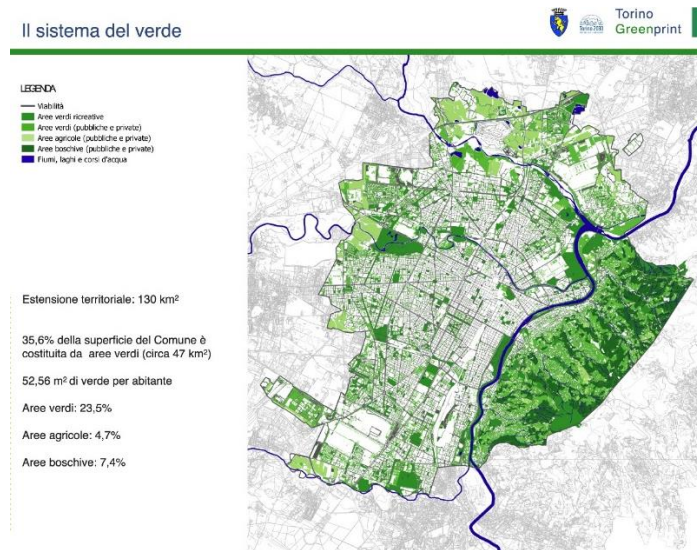


Figure 32 - Green system in the city of Turin

In line with the earlier mentioned plans, the city of Turin has already implemented proactive solutions aimed at mitigating the effects of the Urban Heat Island (UHI) phenomenon.

Chapter six of the “*Piano di Resilienza Climatica*” titled “*Torino a Rischio*”, addresses all aspects pertaining to the urban heat island risk. Information on the implemented nature-based solutions and ongoing initiatives can be found in Chapter eleven.

Below, we provide an example of an action that has already been implemented present in chapter eleven of the report:

3. TETTI VERDI	
Descrizione	La sovrapposizione delle ondate di calore e delle isole di calore, generate dalla tessitura urbana, impone una revisione della progettazione delle città che includa modelli e sistemi resilienti. Non sempre è, però, possibile intervenire sulle fragilità urbane attraverso l'utilizzo di grandi alberature stradali; esiste, invece, la concreta possibilità di poter integrare la funzione del verde pubblico mediante le coperture a verde (tetti verdi, giardini pensili, pareti verdi), che svolgono questo ruolo di adattamento attraverso la cattura della radiazione solare per la trasformazione energetica in processi evapo-traspirativi. I tetti verdi sono in grado di svolgere l'effetto mitigante e possono ridurre la temperatura superficiale del tetto e quella dell'aria circostante ottenendo importanti risparmi sui consumi energetici. Le coperture a verde degli edifici non sostituiscono ma integrano gli effetti della vegetazione urbana apportando notevolissimi contributi in termini microclimatici ed energetici e riducendo i volumi di acqua defluiti in rete. È, pertanto, importante favorire ed incentivare la diffusione dei tetti verdi anche a fini di calmierazione dei fenomeni estremi di calore estivo.
Interventi già realizzati/in corso	Il più importante esempio di tetto verde presente in città è quello dell'Environment Park, che ha permesso di ridurre l'impatto visivo della struttura del Parco tecnologico e la compenetrazione dello stesso nell'area verde del nuovo Parco Dora, assicurando comfort interno agli uffici grazie all'azione di filtro del calore esterno durante l'estate e un maggiore isolamento interno in inverno. Anche la Città ha già realizzato progetti che hanno visto la realizzazione di tetti verdi su edifici di proprietà comunale. In particolare: la Casa del Parco - Parco Colonnetti (1) la ludoteca Il Paguro in Via Oropa 48 (2) il bocciodromo "La Tesorina" di Corso Moncalieri (3) il Parco Arte Vivente in via Giordano Bruno.
	
Sono attualmente in corso, nella zona di Mirafiori sud nell'ambito del Progetto Progireg, la realizzazione del tetto verde estensivo con riciclo dell'acqua piovana per irrigazione presso l'edificio, di proprietà della Città, di via Onorato Vigliani 102, nonché la realizzazione di una parete verde presso l'Istituto comprensivo G. Salvemini di Via Negarville 30/6. È, inoltre, in corso, nell'ambito del Progetto europeo CWC (City Water Circle), la progettazione del tetto verde presso un altro edificio di proprietà comunale, l'Open 011 - Casa della Mobilità Giovanile e dell'intercultura - in c.so Venezia 11, che sarà completato entro il 2021 e prevede anche la realizzazione di una serra per testare i risultati della coltivazione aeroponica basata sull'utilizzo di acqua piovana e un ciclo di produzione alimentare chiuso.	
Soggetti responsabili	Città di Torino
Altri soggetti coinvolti	Soggetti privati
Tempi previsti di realizzazione	Fine lavori per le due sperimentazioni: 2021
Costi previsti e risorse disponibili	La realizzazione di questo tipo di interventi sarà finanziata attraverso diverse tipologie di risorse: - fondi di soggetti privati - fondi dei Progetti europei - contributi statali (es. bonus verde) - fondi di bilancio
Monitoraggio	L'implementazione di questa azione sarà monitorata attraverso i seguenti indicatori: superficie di coperture con verde pensile riduzione consumi energetici

Figure 33 - Piano di Resilienza Climatica - Torino 2030

6 Conclusions: Integrated Reading and Interpretation of Results

In conclusion, our comprehensive analysis, utilizing a machine learning model that incorporated a range of variables, has provided valuable insights into the factors influencing the phenomenon under investigation of the UHI effect in our case study Turin. Through a systematic exploration of these variables, we have discerned their **individual impacts**, shedding light on the **key determinants at play**, as for example buildings. In fact, we need to consider that, while the geometric characteristics of buildings themselves may not bear all the primary determinant of UHI intensity, the human activity taking place within/around these structures plays a pivotal role in shaping urban microclimates. The energy consumption, heating, cooling, and industrial processes occurring within buildings significantly contribute to the heat generated in urban areas. These activities release waste heat into the environment, which, when aggregated across a densely urban landscape, can substantially elevate local temperatures. In fact, if we consider the graphical results obtained in the UHI maps (see chapter 3, [paragraph 3.2.4](#)), for winter, we expect no UHI phenomena, instead, we found out that the phenomenon was indeed present, even with low magnitude, this probably related to the heating systems-on in that period.

It is important to note that, these variables should be considered, although they are often analysed more in-depth in other research efforts. As such, understanding the interplay between building-related human activities and UHI dynamics is indispensable for devising effective strategies to manage and mitigate the urban heat island phenomenon.

Instead, our approach has enabled us to identify which of the fifteen variables considered hold the most significant influence over the outcome in question (Buildings, moisture, green, etc.), by allowing us to obtain a more **intercalary perspective of the phenomenon (urban, district and intervention scale)**, across two distinct time periods

(2001 and 2018), this temporal comparison spanning nearly two decades has allowed us to discern the transformations that have transpired over this significant timeframe, as the **deindustrialisation process** ([see chapter 2.3](#)) that occurred in Turin in 2000s.

Furthermore, the evaluation of variables (almost fifteen in total, [see chapter 3.2](#)), morphological, climate and land cover one, helped us to detect the main changes that interoccurred in these last twenty-years, assuming the changes in climate conditions, as: the reduction of precipitations regime (-312,84 mm since 2010 to today) and the increase of air temperature (+1,07 °C, since 2010 to today); a general overview of the phenomenon can be found in [chapter 2, paragraph 2.2](#), where, with a **broad climate trend analysis** (weather station data), we directly could appreciate the differences between this time span.

6.1 Assessing the Impact of Climate Change: Analysis of Variables Across Seasons in 2001 and 2018

“Climate change is not a mere hypothesis; it's an undeniable scientific truth, and the very cities we inhabit stand as prime architects fuelling this planetary crisis”.

This assumption marked a crucial starting point because **cities, with their anthropic activities, represent the primary focus of our investigation**, into factors that amplify the adverse impacts of climate change. It's not just the planet at stake; it's also the well-being of the individuals residing in these urban centres, as we saw in the literature analysed in [chapter 1, paragraph 1.4](#). Among the phenomena posing a potential threat to cities, the **Urban Heat Island (UHI) effect** (see paragraphs [1.1](#), [1.2](#), [1.3](#)) in dense urban areas (as Turin) stands out as a significant concern. This phenomenon it is important enough to warrant inclusion in **international environmental protocols** as listed in paragraph [1.6](#).

Since dense cities are the core of the phenomenon of the UHI, our main research question was:

"Can good planning practices demonstrate tangible microclimatic changes?"

The answer to this question has been retrieved by the **elaboration of a machine learning model** (see [chapter 4](#)) that leverage algorithms and statistical techniques to identify patterns and relationships within the data and provide insights into **how various factors influence the UHI phenomenon in Turin**.

Modelling phase was preceded by a pre-modelling phase where **all the data were collected in a comparable way** between 2001 and 2018 (see paragraph [3.2](#)), the variables were processed with the algorithms (paragraphs [4.1](#), [4.2](#), [4.3](#)) and finally, a rank of variables has been retrieved by a significance score.

In fact, into this chapter we will delve into the **main comparisons to resume the results obtained**.

- **Compared variables:** These tables are the results from the **average aggregate importance** of the variables, across summer, winter, and midseason (2001 and 2018), that resulted having a significant influence on our UHI model. The specific significance of each season can be consuler at paragraph [4.3](#).

2018			2001		
Rank	Variable	Score	Rank	Variable	Score
1	Buildings (geometry and human activities)	18023.83	1	DTM	6102.323
2	DTM	12137.12	2	NDBI	5555.31
3	NDMI	11733.99	3	Solar radiation	4327.887
4	Sky View Factor	5901.33	4	NDMI	4197.91
5	NDVI	4767.623	5	Albedo	3251.127
6	Emissivity	4509.74	6	Temperature	2824.21
7	NDWI	4488.35	7	NDVI	2696.723
8	Wind speed	4368.623	8	Emissivity	1950.37
9	S/V Ratio	3173.447	9	SRI	1801.577
10	Wind direction	2875.59	10	Wind direction	1575.723
11	Humidity	2757.8	11	NDWI	1282.063
12	Temperature	2680.04	12	Wind speed	1090.547
13	Solar radiation	2603.213	13	Humidity	1035.663
14	SRI	2572.027			
15	Albedo	2151.657			

Figure 34 - Retrieved from page 105, chapter 4.3.

From this table what emerges is that the variables that have the heaviest importance are pretty the same for 2001 and 2018. Buildings (NDBI for 2001), as already anticipated at the beginning of this chapter, are the main participants of the UHI phenomenon, together with: DTM, NDMI, NDVI, Emissivity, etc. Moreover, the variables from 2001 have a slightly different rank, due to the: seasonal variations (for ex. Spring in 2018 is less rainy than in 2001) and data quality (ex. Buildings are much more accurate in 2018 than 2001, because we used a very accurate shapefile from the Piemonte region instead, in 2001, we calculated an index using satellite imagery).

But mainly the variables show a consistent presence in both the years, furthermore, to analyse the differences we retrieved some statistics from the most relevant variables:

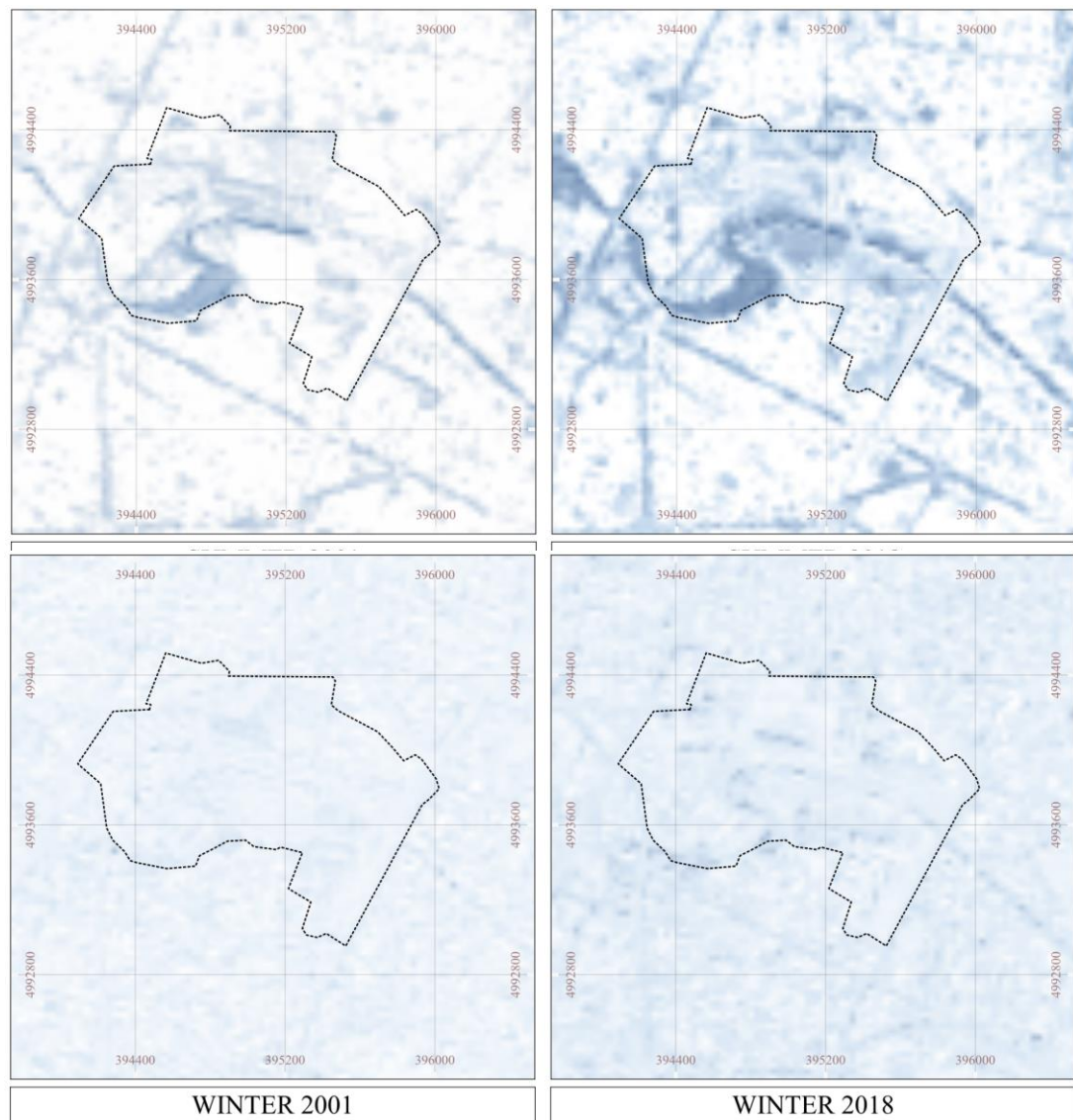
Variable	Scale	Variation Summer (2001-2018)	Variation Spring (2001-2018)	Variation Winter (2001-2018)	Total mean variation (2001-2018)
NDMI	Intervention	↑ 10,07%	↓ -7,41%	↑ 17,89%	↑ 6,85%
	District	↑ 6,77%	↓ -4,27%	↑ 17,69%	↑ 6,73%
	Municipal	↑ 8,44%	↓ -12,43%	↑ 17,39%	↑ 4,47%
NDVI	Intervention	↑ 19,46%	↑ 21,79%	↑ 8,23%	↑ 16,49%
	District	↑ 6,41%	↑ 16,61%	↑ 7,61%	↑ 10,21%
	Municipal	↑ 6,17%	↑ 11,09%	↑ 1,53%	↑ 6,26%
EMISSION	Intervention	↓ -0,02%	↓ -0,08%	↑ 0,00%	↓ -0,03%
	District	↓ -0,03%	↓ -0,10%	↑ 0,00%	↓ -0,04%
	Municipal	↓ -0,03%	↓ -0,10%	↓ -0,01%	↓ -0,05%
ALBEDO	Intervention	↑ 3,09%	↑ 0,63%	↑ 14,86%	↑ 6,19%
	District	↑ 1,18%	↑ 0,11%	↑ 13,46%	↑ 4,92%
	Municipal	↑ 3,27%	↑ 0,23%	↑ 15,40%	↑ 6,30%

Table 21 - Tabular statistics of some of the most relevant variables of UHI

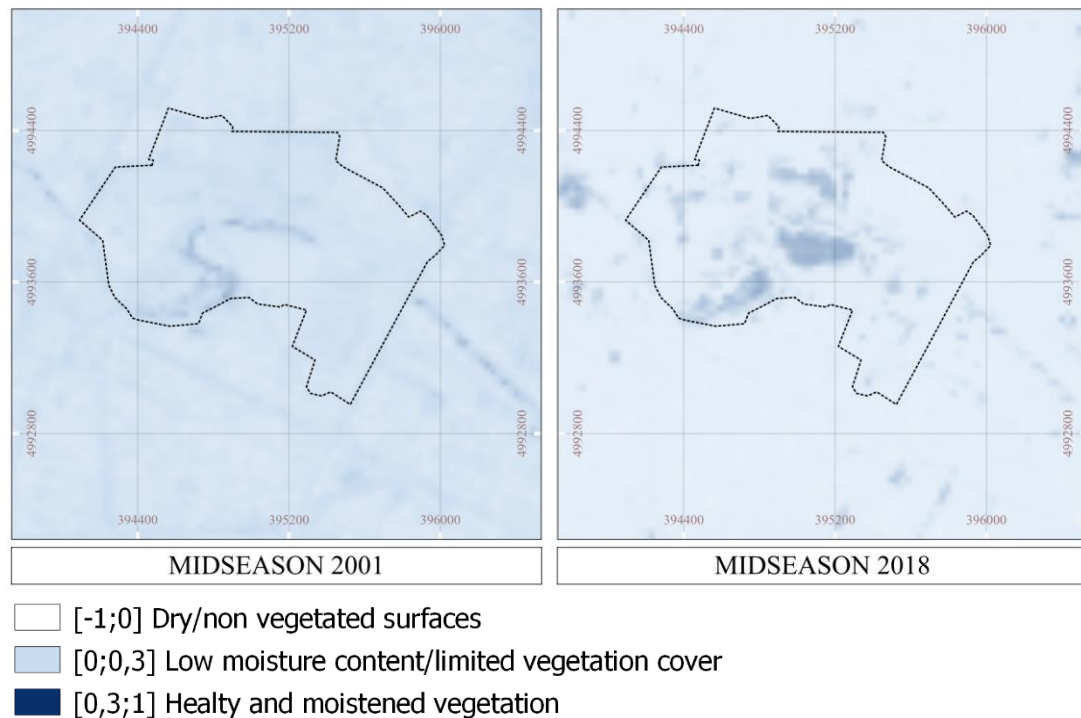
- **Soil moisture:** in urban areas humidity can vary widely due to differences in land use and soil properties. These variations can influence UHI intensity by affecting the thermal properties of the surface, including heat storage and heat transfer processes. Moisture, whether in the form of soil moisture, vegetation, or higher atmospheric humidity, has a **cooling effect that counteracts the heat island effect.**

The reference by Stewart and Oke (2012) [62] highlights the relevance of considering local climate zones in urban temperature studies, emphasizing the importance of soil moisture as a key variable. Moisture levels vary across seasons due to temperature fluctuations and atmospheric conditions. From the table above we can notice that:

- moisture variation levels (2001-2018) in Turin are slightly increased, also in the intervention area the values are higher.
- Spring registers a negative variation, this referred to the fact that the rainy season is characterised by less precipitation in 2018.



[62] Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879-1900

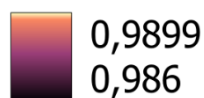


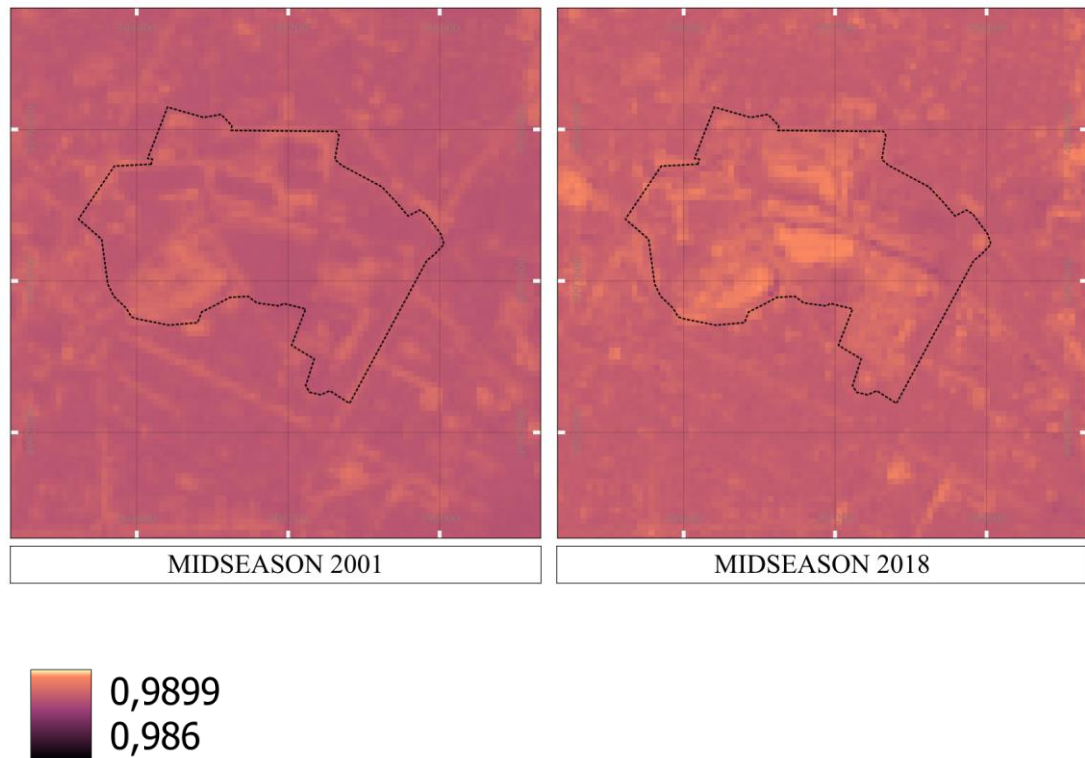
- Increased Vegetation:** Urban regeneration often includes initiatives to enhance green spaces and increase urban vegetation. This scientific approach aligns with findings that vegetation, through processes such as *transpiration and shading*, can significantly reduce surface temperatures (Shashua-Bar & Hoffman, 2009) [63]. From the NDVI calculations for 2001 and 2018 is pretty evident that the vegetation increased in the area where the urban interventions took place. In fact, **the built-up area is coloured with dark blue, instead, the green area is coloured with highlighted green.** Nowadays more areas with highlighted green are present. Note that in winter the classification palette used to enhance the green is darker (dark violet), because the fact that most of the leaves felled. Instead, in mid-season the vegetation is blossoming so lighter colours indicate that there's a major presence of vegetation.

[63] Shashua-Bar, L., & Hoffman, M. E. (2009). Vegetation as a climatic component in the design of an urban street: An empirical model for predicting the cooling effect of urban green areas with trees. *Energy and Buildings*, 41(6), 647-655



- **Emissivity:** Surfaces with higher emissivity values have a greater ability to emit thermal radiation, which means they release heat more efficiently. In contrast, surfaces with lower emissivity values retain heat, contributing to higher temperatures within urban environments. This phenomenon can exacerbate the UHI effect by causing urban areas to become significantly warmer.



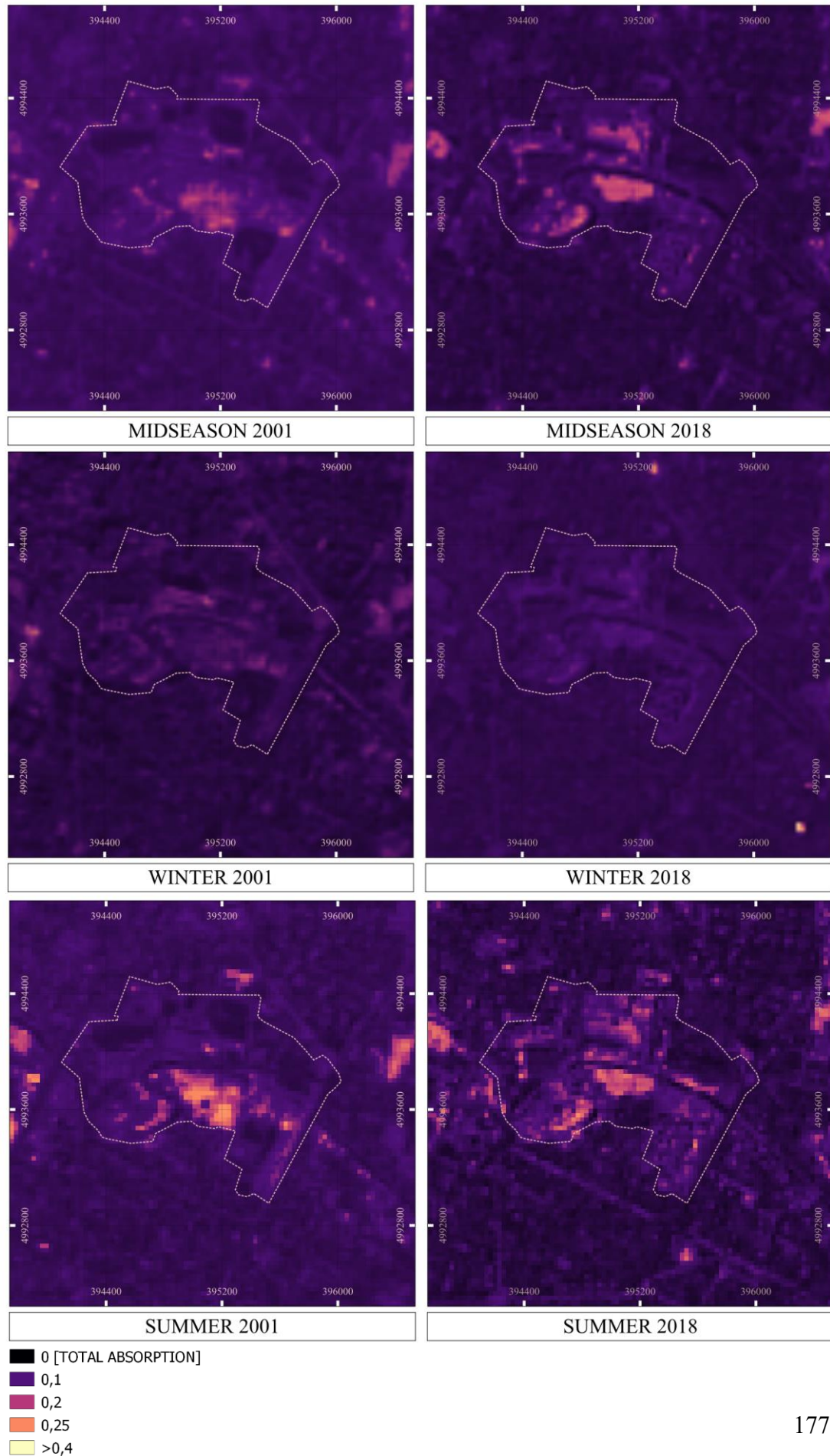


- **Albedo Modification:** Scientific research highlights the importance of albedo modification in UHI mitigation. Urban regeneration initiatives can introduce high-albedo surfaces, such as cool roofs and reflective pavements, which reflect more solar radiation and absorb less heat, thereby mitigating UHI effects (Akbari et al., 2009) [64].

Albedo in 2001 registers areas with **low reflectance capacity (darker colours)**, instead, 2018 has areas where the **reflectance capacity, after intervention, has increased (lighter colours)**.

For winter and midseason, the output is the same, for 2018 results show better reflective capacity.

[64] Ibid [63]



*Does intervention mitigation bring new effects to light
in UHI intensity?*

The answer to this question has been computed using the variations registered between 2001 and 2018 (see chapter 5, paragraph [5.1](#)) of the main variables contributing to the phenomenon, and the normalised intensities of UHI values for the same time considered. In fact, starting from the importance scores resulted from the Random Forest regression, we observed a high influence of Buildings, NDMI, NDVI, DTM, and Albedo.

Buildings, as previously mentioned, are the main elements that contribute to the phenomenon, underlining the fact that our analysis is not enough descriptive of all the phenomenon. DTM, for sure is a variable that influences the intensity of UHI, in fact, from the UHI maps we can observe the heating effect in the downtown, instead in the hills, and more elevated areas, a cooling effect can be displayed.

What is important for us to highlight is the effect/impact that planning practices can have on the phenomenon in question, by showing practical results in decreasing percentage of intensity, enacted by NDMI, NDVI and Albedo, that are indexes that can be measured and compared consistently.

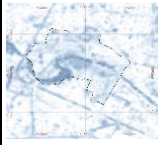
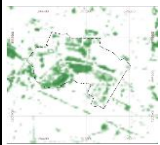
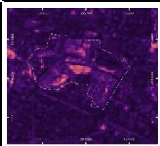
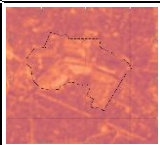
Variables		Incidence % UHI	Pearson coefficient (r)	RF Importance Score
NDMI		-16.39 %	-0.797	11733.990
NDVI		-15.52 %	-0.679	4767.600
ALBEDO		-7.48 %	-0.594	2151.600
EMISSION		0.07 %	-0.658	4509.740

Table 22 - Variables' incidence on UHI mitigation

Can we highlight differences in UHI intensity across different urban morphologies?

Yes, we can indeed highlight some differences in Urban Heat Island (UHI) intensity across **various urban morphologies**. UHI intensity varies based on the characteristics and layout of urban areas. For instance, densely built-up urban cores with tall buildings and limited green spaces tend to exhibit higher UHI intensities due to increased heat retention and reduced natural cooling mechanisms, an example can be the “Sacchi” neighbourhood in Turin. In contrast, open areas often have lower UHI intensities as they feature more open spaces and vegetation. Therefore, by studying these variations, we can gain valuable insights into how urban planning and design can mitigate the impacts of UHI.

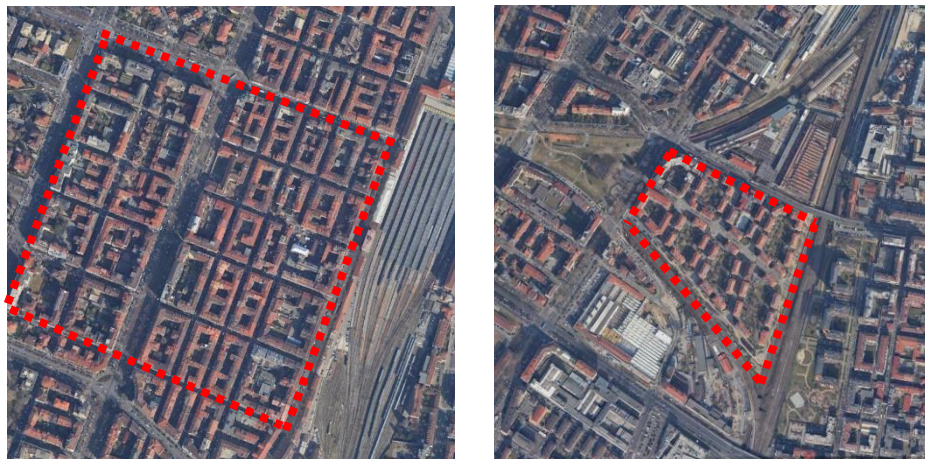


Figure 35 - Sacchi vs Arquata neighbourhood

TABLE III. HOMOGENOUS ZONES OF TURIN (IT)

Zone	Building variables and urban parameters					
	<i>BCR</i>	<i>MOS</i>	<i>BD</i>	<i>H/W</i>	<i>H/H_{avg}</i>	<i>A</i>
Arquata	0.18	0.59	3.56	0.27	1.01	0.19
Crocetta	0.28	0.45	5.86	0.52	0.97	0.15
Raffaello	0.38	0.45	8.10	0.57	0.95	0.15
Sacchi	0.40	0.54	7.72	0.60	0.95	0.14
Mediterraneo	0.24	0.58	6.96	0.62	1.01	0.15
Olympic Village	0.16	0.29	4.13	0.34	1.02	0.16

Table 23 - Table retrieved by Urban scale energy models, G. Mutani, et al (2020)

District	UHI _{av} 18/08/2018	UHI _{av} 11/02/2018	UHI _{av} 16/04/2018
Sacchi	1.3445	0.2819	0.6041
Arquata	1.2560	0.3824	0.4382

Table 24 - Sacchi vs Arquata neighbourhood UHI averages

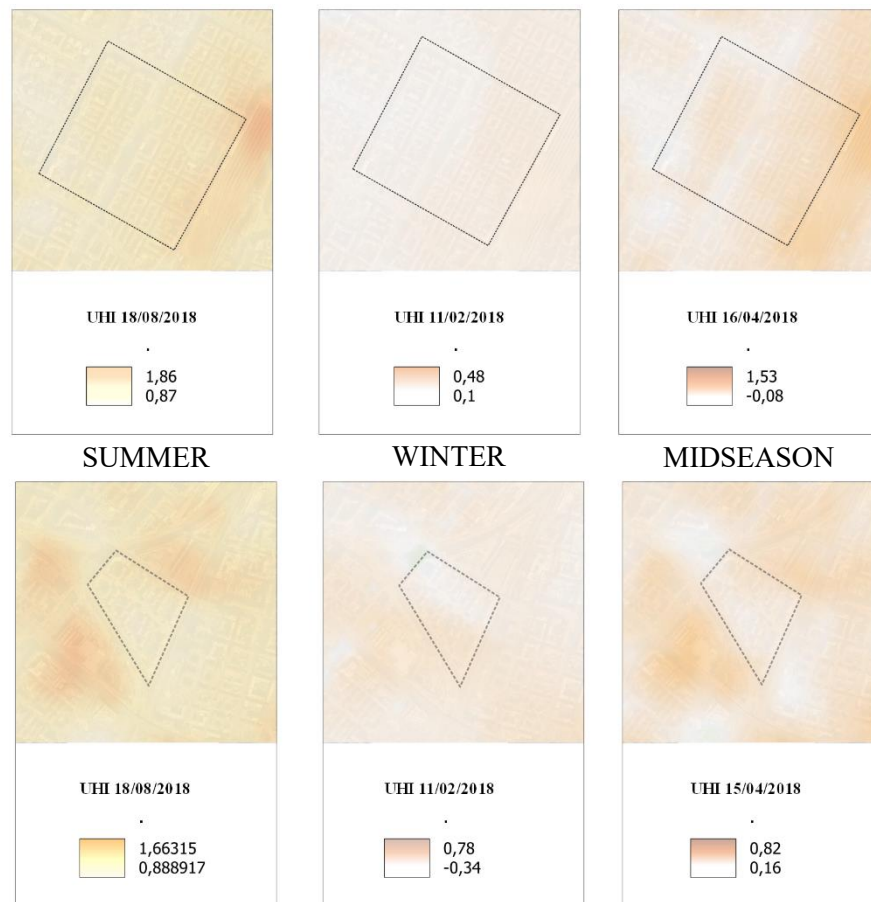


Figure 36 - UHI Sacchi vs Arquata neighborhoods UHI

The intensity of the Urban Heat Island (UHI) varies slightly between the Sacchi and Arquata neighbourhoods. Sacchi features a densely built urban environment, whereas Arquata is characterized by a more open layout of buildings. During the summer and midseason, UHI tends to be more pronounced in the densely populated Sacchi neighbourhood. Conversely, during the winter, UHI levels are higher in the Arquata neighbourhood, likely influenced by industrial activities near Arquata.

6.2 Planning documents and future directions for UHIs research

In the ongoing effort to combat the adverse effects of Urban Heat Islands (UHI), it becomes increasingly evident that key elements such as vegetation, humidity, and albedo play crucial roles in achieving effective mitigation. To gain insights into the **Turin's regulatory framework addressing UHI concerns**, we turned our attention to the planning documents of the city. Among the specific documents, the “*Allegato Energetico Ambientale di Torino*” emerged as a noteworthy resource. Within this document, a range of mitigation strategies are outlined, emphasizing normativism and voluntary elements.

An example of content of the building code attachment is the colour of external building surfaces that carries energy implications, as it correlates with the absorption and emission coefficients of these surfaces [65]. A low emission coefficient coupled with a high absorption coefficient can lead to passive behaviour, causing these surfaces to heat up in the sun and transmit a portion of this thermal energy. Additionally, the roughness of external surfaces also impacts thermal exchange, with heat transfer decreasing as surface roughness increases.

In fact, we could appreciate this by analysing the results of Albedo for Turin, where surfaces with high values are exceptionally reflective and tend to bounce back a significant portion of incoming solar radiation, helping in reducing the retaining of heat in the city.

Another question seems to appear of interest:

Why do we frequently encounter newly constructed buildings that feature dark colors? Is this a deliberate design choice or is there another reason behind it?

[65] Allegato Energetico - Ambientale al Regolamento Edilizio della Città di Torino, p.53, 2020

The answer to this question has been derived from the extensive analysis of numerous articles in the literature, as well as from the results of our analysis, which highlighted a reduction of UHI intensity by almost 7.5% through albedo modifications with the new intervention in the Ex-Teksid area.

Considering this, as one of the mitigation initiatives that can be implemented to “modify albedo”, in fact could be useful, at least in new developments, to observe the **use of light colours for outer surfaces**. Generally, seems more aesthetically to privilege the use of dark colours, as we can observe from these real estate listings in Turin:

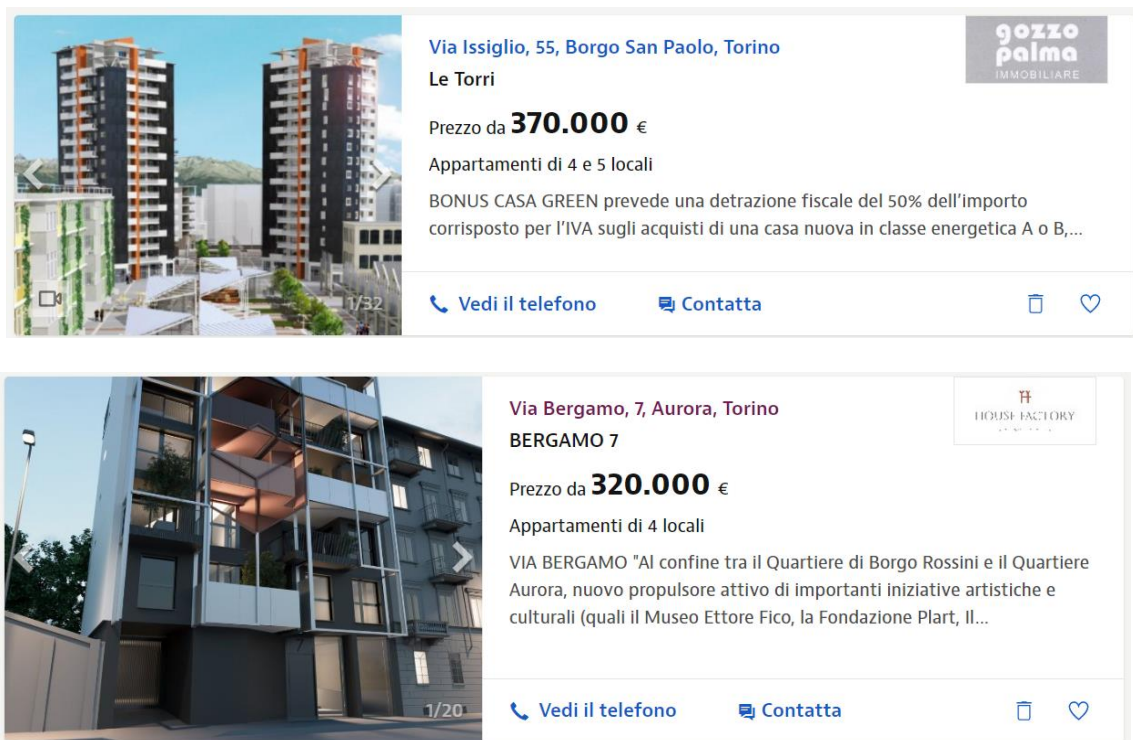


Figure 37 - New buildings in Borgo San Paolo and Aurora – Turin – Source: Idealista.it

Albedo, in a new development, for sure is not the only variable to consider, but is also one of the main variables to consider when we discuss about UHI phenomenon.

In this regard, we also inquired the “Piano del Colore” to understand if there were specific binding limitations to the use of colours in surfaces, but nothing noteworthy related to UHI topic.

In conclusion, this study represents a significant step forward in understanding the dynamics of the Urban Heat Island (UHI) phenomenon in the context of Turin, Italy, from 2001 to 2018. The application of machine learning techniques has allowed us to comprehensively analyse and quantify the variations in UHI intensity over this period.

Our findings have highlighted the critical importance of several key variables, in driving the fluctuations of UHI. These variables have emerged as influential factors in shaping the urban heat landscape and should serve as pivotal considerations for future urban planning.

Furthermore, this research offers practical implications for urban regeneration and mitigation strategies. By demonstrating the potential for tangible UHI reduction through urban regeneration interventions (as the one of the EX-Teksid area) we underscore the significance of incorporating UHI mitigation measures in urban planning and development projects. The successful case of mitigating UHI through thoughtful urban regeneration initiatives in Turin serves as a model for other cities striving to create more sustainable and climate-resilient urban environments.

As we move forward, it is imperative to translate these insights into actionable policies and practices that prioritize the reduction of UHI effects. The collaboration between data-driven analysis, machine learning, and urban planning holds great promise in addressing the challenges posed by UHI and fostering healthier, more sustainable, and resilient cities.

7 Bibliography

1. Equere, V., Mirzaei, P. A., Riffat, S., & Wang, Y. (2021). Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using GIS and ANN. *Sustainable Cities and Society*, 69, 102825. <https://doi.org/10.1016/j.scs.2021.102825>
2. Espino, D. J., Manchado, C., Valcarce, A. R., & Moscardò, V. (2022). ArcUHI: A GIS add-in for automated modelling of the Urban Heat Island effect through machine learning. *Environmental Modelling & Software*, 146, 105176. doi: 10.1016/j.envsoft.2021.105176
3. Howard, L. (1833). *The Climate of London*. London, UK: Harvey and Darton.
4. IPCC Report 2023, “Longer Report”, page 105: IPCC_AR6_SYR_LongerReport.pdf
5. IPCC Report 2023, “Longer Report”, page 48: IPCC_AR6_SYR_LongerReport.pdf
6. IPCC Report 2023, “Longer Report”, page 49: IPCC_AR6_SYR_LongerReport.pdf
7. IPCC Report 2023, “Longer Report”, page 50: IPCC_AR6_SYR_LongerReport.pdf
8. IPCC Report 2023, “Summary for Policymakers”, page 16-17: IPCC_AR6_SYR_SPM.pdf

9. Istituto Per L'innovazione E Trasparenza Degli Appalti E La Compatibilità Ambientale Itaca, Protocollo Itaca A Scala Urbana Sintetico, Metodologia E Strumento Di Verifica, Versione 2.02 (2020)
10. Lan, T., Peng, J., Liu, Y., Zhao, Y., Dong, J., Jiang, S., Cheng, X., & Corcoran, J. (2023). The future of China's urban heat island effects: A machine learning based scenario analysis on climatic-socioeconomic policies. *Science of The Total Environment*, 153114. <https://doi.org/10.1016/j.scitotenv.2022.153114>
11. Liao, K., Hong, Y., & Heo, J. (2018). The effect of spatial heterogeneity in urban morphology on surface urban heat islands. *Environmental Research Letters*, 13(6), 064014. doi: 10.1088/1748-9326/aac2d8
12. Lin, J., Qiu, S., Tan, X., & Zhuang, Y. (2022). Measuring the relationship between morphological spatial pattern of green space and urban heat island using machine learning methods. *Sustainable Cities and Society*, 95, 103877. doi: 10.1016/j.scs.2021.103877
13. Liu, S., Zhang, J., Li, J., Li, Y., Zhang, J., & Wu, X. (2021). Simulating and mitigating extreme urban heat island effects in a factory area based on machine learning. *Sustainable Cities and Society*, 70, 102977. doi: 10.1016/j.scs.2021.102977.
14. Liu, X., Huang, B., Li, R., Zhang, J., Gou, Q., Zhou, T., & Huang, Z. (2018). Wind environment assessment and planning of urban natural ventilation corridors using GIS: Shenzhen as a case study. *Building and Environment*, 129, 100-113.
15. Mohammad, P., Goswami, A., Chauhan, S., & Nayak, S. (2022). Machine learning algorithm-based prediction of land use land cover and land surface temperature changes to characterize the surface urban heat island phenomena over

Ahmedabad city, India. *Sustainable Cities and Society*, 96, 103952. doi:
10.1016/j.scs.2021.103952

16. Mutani, G., & Beltramino, S. (2022). Geospatial assessment and modeling of outdoor thermal comfort at urban scale. *Sustainable Cities and Society*, 77, 103245. <https://doi.org/10.1016/j.scs.2021.103245>
17. Mutani, G., & Todeschi, V. (2020). The Effects of Green Roofs on Outdoor Thermal Comfort, Urban Heat Island Mitigation and Energy Savings. *Sustainability*, 12(2), 616. doi: 10.3390/su12020616.
18. Mutani, G., Todeschi, V., & Matsuo, K. (2019). Urban Heat Island Mitigation: A GIS-based Model for Hiroshima. *Sustainability*, 11(8), 2369. <https://doi.org/10.3390/su11082369>
19. Nakata-Osaki, C. M., Souza, L. C. L., & Rodrigues, D. S. (2017). Tool for Heat Island Simulation: A GIS extension model to calculate urban heat island intensity based on urban geometry. *Computers, Environment and Urban Systems*, 65, 104-117. doi: 10.1016/j.compenvurbsys.2017.04.002
20. Oke T.R. (1976) The distinction between canopy and boundary-layer urban heat islands, *Atmosphere*, 14:4, 268-277, DOI: 10.1080/00046973.1976.9648422
21. Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108(455), 1-24.
22. Oke, T. R. (1987). *Boundary Layer Climates* (2nd ed.). Methuen.
23. Oliveira, A., Lopes, A., Niza, S., & Soares, A. (2021). An urban energy balance-guided machine learning approach for synthetic nocturnal surface Urban Heat Island prediction: A heatwave event in Naples. *Building and Environment*, 195, 107749. <https://doi.org/10.1016/j.buildenv.2021.107749>

24. Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>
25. Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>
26. Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T. (2022). Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools. *Climate*, 10, 3. <https://doi.org/10.3390/cli10010003>
27. Reis, C., Lopes, A., & Santos Nouri, A. (2021). Urban heat island data by local weather types in Lisbon metropolitan area based on Copernicus climate variables dataset for European cities. *Sustainable Cities and Society*, 65, 102618.
28. Scott, A. A., Waugh, D. W., & Zaitchik, B. F. (2019). Reduced urban heat island intensity under warmer conditions. *Environmental Research Letters*, 14(9), 094011.
29. Sun, T., Gao, J., Yu, Q., & Wang, X. (2021). Mapping the spatial distribution of nocturnal urban heat island based on Local Climate Zone framework. *Sustainable Cities and Society*, 72, 103072. <https://doi.org/10.1016/j.scs.2021.103072>
30. Thakur, Abhishek. *Approaching (Almost) Any Machine Learning Problem (English Edition)* (p.14). Abhishek Thakur. Kindle Edition.

31. Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.32). Abhishek Thakur. Edition Kindle.
32. Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.69). Abhishek Thakur. Edition Kindle.
33. Thakur, Abhishek. Approaching (Almost) Any Machine Learning Problem (p.69). Abhishek Thakur. Edition Kindle.
34. Todeschi, V., Mutani, G., Baima, L., Nigra, M., & Robiglio, M. (2021). Smart Solutions for Sustainable Cities—The Re-Coding Experience for Harnessing the Potential of Urban Rooftops. *Sustainability*, 13(16), 8954. <https://doi.org/10.3390/su13168954>
35. Vujovic, S., Haddad, B., Karaky, H., Sebaibi, N., & Boutouil, M. (2020). Urban Heat Island: Causes, Consequences, and Mitigation Measures with Emphasis on Reflective and Permeable Pavements. *Sustainability*, 12(23), 9836. doi: 10.3390/su12239836.
36. Wang S., W. Cai, Y. Tao, Q. Chayn Sun, P. Pui Yun Wong, X.Huang, Y. Liu, (2023). Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models. *Science of The Total Environment*. Volume 871. 162005. ISSN 0048-9697.
37. Wang, S., Cai, W., Tao, Y., Sun, Q. C., Wong, P. P. Y., Huang, X., & Liu, Y. (2023). Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models. *Environmental Research*, 204, 112012. <https://doi.org/10.1016/j.envres.2022.112012>

38. Wang, W., Wang, D., Chen, H., Wang, B., & Chen, X. (2022). Identifying urban ventilation corridors through quantitative analysis of ventilation potential and wind characteristics. *Building and Environment*, 208, 108759.
39. Xu, H.; Lin, D.; Tang, F. The impact of impervious surface development on land surface temperature in a subtropical city: Xiamen, China. *Int. J. Climatol.* 2013, 33, 1873–1883.
40. Yang, L., Qian, F., Song, D.-X., & Zheng, K.-J. (2016). Research on Urban Heat-Island Effect. *Procedia Engineering* 169 (2016) 11 – 18. Doi: 10.1016/j.proeng.2016.10.002

8 Sitography

1. Britannica, T. Editors of Encyclopaedia (2022) specific heat. Encyclopedia Britannica. <https://www.britannica.com/science/specific-heat>
2. Britannica, The Editors of Encyclopaedia. "albedo". Encyclopedia Britannica, 29 Dec. 2022, <https://www.britannica.com/science/albedo>.
3. Comune di Torino: Parco Dora: aperti al pubblico il lotto Valdocco Nord e Iron Valley – Verde Pubblico (comune.torino.it)
4. Esri. (2023). Random Forest guide in ArcGIS. Retrieved July 31, 2023, from <https://pro.arcgis.com/en/pro-app/latest/toolreference/spatialstatistics/forestbasedclassificationregression.htm>
5. Ex Stabilimento TEKSID, ex Ferriere Fiat Ingest: Ex Stabilimento Teksid, ex Ferriere Fiat Ingest – MuseoTorino
6. Ex Stabilimento TEKSID, ex Ferriere Fiat Valdocco: Ex stabilimento Teksid, ex Ferriere Fiat Valdocco – MuseoTorino
7. Ex Stabilimento TEKSID, ex Ferriere Fiat Vitali: Ex stabilimento Teksid, ex Ferriere Fiat Vitali – MuseoTorino
8. <https://geog.ubc.ca/profile/tim-oke/>
9. <https://www.scienceandsociety.co.uk/index.asp>
10. Il Piano Strategico dell’Infrastruttura Verde Torinese

11. Immagini del cambiamento: [PD04 \(polito.it\)](https://polito.it)
12. Intergovernmental Panel on Climate Change (IPCC). (2014). Climate Change 2014: Synthesis Report. Retrieved from <https://www.ipcc.ch/report/ar5/syr/>.
13. Oslo Opera House, Oslo, Norway, Daniel Ghio, Unsplash: <https://unsplash.com/it/foto/LHf1WVqdmCI>
14. [Piano di Resilienza Climatica](#)
15. Singapore National Parks Board (NParks): Parks & Nature Reserves - Gardens, Parks & Nature - National Parks Board (NParks)
16. Stewart, K. (2023). thermal conductivity. Encyclopedia Britannica. <https://www.britannica.com/science/thermal-conductivity>
17. Torino 1938|45 Luoghi memoria: Torino 1938|45 - La città delle fabbriche. Fiat Ferriere (istoreto.it)
18. Unsplash: <https://unsplash.com/it/foto/bIpKSEsaN6Q>

9 Table of figures

Figure 1 - Representation by Oke T.R. The energetic basis of the UHIs (Oke,1982).....	9
Figure 2 - Howard's examination of the urban effect in London (1833)	10
Figure 3 - Schematic representation of the urban atmosphere illustrating a two-layer classification of thermal modification (Illustration by T.R. Oke, 1976)	11
Figure 4 - Urban Canyon (Illustration by Y. Choi, S.Lee, H.Moon, 2018).....	11
Figure 5 - Photo by Fabio Fistarol on Unsplash	34
Figure 6 - Weather station localization.....	38
Figure 7 - IPCC Longer Report, 2023 (MED area means Mediterranean area, it also includes Italy)	51
Figure 8 - IPCC Longer Report, 2023	52
Figure 9 – Area EX TEKSID captured in the orthophoto of CGR IT2000 Provincia di Torino (year 2000).....	53
Figure 10 - Ex ferriere Fiat Valdocco (left), Ex ferriere Fiat Vitali (right)	55
Figure 11 - Teksid area and other factories, from Verolengo Street (view looking south), 1970	56
Figure 12 - Dora Park, condominiums, urban empty residue, 2018.....	56
Figure 13 - STAC API Browser (Qgis Plugin).....	57
Figure 14 - Selection of the images	58
Figure 15 - SCP Preprocessing for satellite images (Qgis)	59
Figure 16 - Python Console (QGis)	60
Figure 17 - Fractional soil cover [%] - Elaboration on data by Copernicus.....	68
Figure 18 – Scale factor by Landsat 8's metadata	76
Figure 19 - SRI Calculation [QGis].....	81
Figure 20 - S_V [Shape and heatmap QGis]	83
Figure 21 – SKY VIEW FACTOR.....	83
Figure 22 - LST retrieving process (Authors elaboration)	84
Figure 23 - Data Spell Python code example for hyper parametrization tuning	123
Figure 24 - Best parameters output from Python code	125

Figure 25 - Forest-based Classification and Regression (ArcGis Pro).....	130
Figure 26 - Calculated vs Predicted RF UHI [2001]	153
Figure 27 - Calculated vs Predicted RF UHI [2018]	154
Figure 28 - Parco Dora renewal, Ex Area TEKSID [].....	160
Figure 29 – Oslo Opera House [].....	161
Figure 30 - Yuan Ching Road, Jurong Lake Gardens, Singapore [].....	162
Figure 31 - Singapore National Parks Board (NParks) []	165
Figure 32 - Green system in the city of Turin.....	166
Figure 33 - Piano di Resilienza Climatica - Torino 2030	167
Figure 34 - Retrieved from page 105, chapter 4.3.	170
Figure 35 - Sacchi vs Arquata neighbourhood	179
Figure 36 - UHI Sacchi vs Arquata neighborhoods UHI	180
Figure 37 - New buildings in Borgo San Paolo and Aurora – Turin – Source: Idealista.it	182

APPENDIX

Automating processes in QGIS using Python can be a powerful way to streamline repetitive tasks, improve efficiency, and enhance geospatial workflows. The Python Console in QGIS provides an interactive environment where Python scripts can be executed to interact with QGIS' functionality and data.

A. Python scripts (Variables calculation)

NDVI *Change paths (blu part)

```
import rasterio
import numpy as np
from qgis.core import QgsRasterLayer, QgsProject, QgsMapLayerType
from qgis.utils import iface

# Define the paths to the red and NIR bands of Landsat 8
red_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O
- Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B4.TIF'
nir_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O
- Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B5.TIF'
output_ndvi_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino
Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI
TERMICHE/2018/Agosto 2018/NDVI/output_ndvi.tif'
```

```
# Open the red band and read the data
with rasterio.open(red_band_path) as red_ds:
    red = red_ds.read(1).astype(np.float32)

# Open the NIR band and read the data
with rasterio.open(nir_band_path) as nir_ds:
    nir = nir_ds.read(1).astype(np.float32)

# Calculate NDVI
ndvi = (nir - red) / (nir + red)

# Write NDVI to output file
with rasterio.open(red_band_path) as src:
    profile = src.profile
    profile.update(dtype=rasterio.float32, count=1)

    with rasterio.open(output_ndvi_path, 'w', **profile) as dst:
        dst.write(ndvi, 1)

# Add the NDVI layer to QGIS
layer = QgsRasterLayer(output_ndvi_path, 'NDVI')
QgsProject.instance().addMapLayer(layer)
layer.setRenderer(layer.renderer().clone())
QgsProject.instance().layerTreeRoot().insertLayer(0, layer)
```



```
# Refresh the map canvas
iface.mapCanvas().refreshAllLayers()

# Open the NDVI output file in QGIS
iface.addRasterLayer(output_ndvi_path, 'NDVI')
```

PVI *Change band paths (blu part)

```
import rasterio
import numpy as np
from math import sqrt
from qgis.core import QgsRasterLayer, QgsProject
from qgis.utils import iface

# Define the paths to the red and NIR bands of satellite image
red_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O
- Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B4.TIF'
nir_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O
- Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B5.TIF'
output_pvi_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino
Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI
TERMICHE/2018/Agosto 2018/PVI/output_pvi.tif'

# Open the red band and read the data
```

```
with rasterio.open(red_band_path) as red_ds:
    red = red_ds.read(1).astype(np.float32)

# Open the NIR band and read the data
with rasterio.open(nir_band_path) as nir_ds:
    nir = nir_ds.read(1).astype(np.float32)

# Calculate NDVI
ndvi = (nir - red) / (nir + red)

# Calculate NDVImin and NDVImax
ndvi_min = np.nanmin(ndvi)
ndvi_max = np.nanmax(ndvi)

# Calculate PVI
pvi = np.sqrt((ndvi - ndvi_min) / (ndvi_max - ndvi_min))

# Write PVI to output file
with rasterio.open(red_band_path) as src:
    profile = src.profile
    profile.update(dtype=rasterio.float32, count=1)

    with rasterio.open(output_pvi_path, 'w', **profile) as dst:
        dst.write(pvi, 1)

# Add the PVI layer to QGIS
```

```
layer = QgsRasterLayer(output_pvi_path, 'PVI')
QgsProject.instance().addMapLayer(layer)

# Refresh the map canvas
iface.mapCanvas().refreshAllLayers()

# Open the PVI output file in QGIS
iface.addRasterLayer(output_pvi_path, 'PVI')
```

NDMI *Change band paths (blu part)

```
import rasterio
import numpy as np
from qgis.core import QgsRasterLayer, QgsProject
from qgis.utils import iface

# Define the paths to the near-infrared (NIR) and shortwave infrared
(SWIR) bands of Landsat 8
nir_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B5.TIF'

swir_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B6.TIF'

output_ndmi_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/Agosto 2018/NDMI/output_ndmi.tif'

# Open the NIR band and read the data
with rasterio.open(nir_band_path) as nir_ds:
    nir = nir_ds.read(1).astype(np.float32)
```

```
# Open the SWIR band and read the data
with rasterio.open(swir_band_path) as swir_ds:
    swir = swir_ds.read(1).astype(np.float32)

# Calculate NDMI
ndmi = (nir - swir) / (nir + swir)

# Write NDMI to output file with the same metadata as the NIR band
with rasterio.open(nir_band_path) as src:
    profile = src.profile
    profile.update(dtype=rasterio.float32, count=1)

    with rasterio.open(output_ndmi_path, 'w', **profile) as dst:
        dst.write(ndmi, 1)

# Add the NDMI layer to QGIS
layer = QgsRasterLayer(output_ndmi_path, 'NDMI')
QgsProject.instance().addMapLayer(layer)

# Refresh the map canvas
iface.mapCanvas().refreshAllLayers()

# Open the NDMI output file in QGIS
iface.addRasterLayer(output_ndmi_path, 'NDMI')
```

ALBEDO *Change band paths (blu part)

```
import rasterio
import numpy as np
from qgis.core import QgsRasterLayer, QgsProject
from qgis.utils import iface

# Define the paths to the necessary bands of Landsat 8
band2_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O -
Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B2.TIF'

# Blue band
```

```
band3_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O -  
Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO  
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B3.TIF'  
  
# Green band  
band4_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O -  
Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO  
2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B4.TIF'  
  
# Red band  
output_albedo_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino  
Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI  
TERMICHE/2018/Agosto 2018/ALBEDO/output_albedo_2.tif'  
  
# Open the Landsat 8 bands and read the data  
with rasterio.open(band2_path) as band2_ds:  
    band2 = band2_ds.read(1).astype(np.float32)  
  
with rasterio.open(band3_path) as band3_ds:  
    band3 = band3_ds.read(1).astype(np.float32)  
  
with rasterio.open(band4_path) as band4_ds:  
    band4 = band4_ds.read(1).astype(np.float32)  
  
# Calculate the reflectance  
reflectance_band2 = band2 / np.max(band2)  
reflectance_band3 = band3 / np.max(band3)  
reflectance_band4 = band4 / np.max(band4)  
  
# Calculate the albedo  
albedo = (reflectance_band2 + reflectance_band3 + reflectance_band4) /  
3.0  
  
# Normalize the albedo values between 0 and 1  
albedo = np.clip(albedo, 0, 1)  
  
# Write albedo to output file with the same metadata as the Blue band  
with rasterio.open(band2_path) as src:  
    profile = src.profile  
    profile.update(dtype=rasterio.float32, count=1)
```



```
with rasterio.open(output_albedo_path, 'w', **profile) as dst:
    dst.write(albedo, 1)

# Add the albedo layer to QGIS
layer = QgsRasterLayer(output_albedo_path, 'Albedo')
QgsProject.instance().addMapLayer(layer)

# Refresh the map canvas
iface.mapCanvas().refreshAllLayers()

# Open the albedo output file in QGIS
iface.addRasterLayer(output_albedo_path, 'Albedo')
```

EMISSIONITY *Change band paths (blu part)

```
import rasterio
import numpy as np
from qgis.core import QgsRasterLayer, QgsProject
from qgis.utils import iface

# Define the paths to the necessary bands of Landsat 8
pvi_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/Agosto 2018/PVI/output_pvi.tif'
# Proportional Vegetation Index (PVI) band
output_emissivity_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/Agosto 2018/EMISSIONITY/output_emissivity_2.tif'

# Open the PVI band and read the data
with rasterio.open(pvi_path) as pvi_ds:
    pvi = pvi_ds.read(1).astype(np.float32)

# Calculate emissivity
emissivity = (0.004 * pvi) + 0.986

# Write emissivity to output file with the same metadata as the PVI band
with rasterio.open(pvi_path) as src:
```

```
profile = src.profile
profile.update(dtype=rasterio.float32, count=1)

with rasterio.open(output_emissivity_path, 'w', **profile) as dst:
    dst.write(emissivity, 1)

# Add the emissivity layer to QGIS
layer = QgsRasterLayer(output_emissivity_path, 'Emissivity')
QgsProject.instance().addMapLayer(layer)

# Refresh the map canvas
iface.mapCanvas().refreshAllLayers()

# Open the emissivity output file in QGIS
iface.addRasterLayer(output_emissivity_path, 'Emissivity')
```

SRI *Change band paths (blu part)

```
import rasterio
import numpy as np

# Define the paths to the Landsat 8 bands

red_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B4.TIF'

green_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B3.TIF'

blue_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B2.TIF'

nir_band_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino Studenti)/O - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI TERMICHE/2018/22 AGO 2018/LC08_L2SP_195029_20180822_20200831_02_T1_SR_B5.TIF'
```

```
# Define the output path for the SRI raster
output_sri_path = 'C:/Users/sufa/Dropbox (Politecnico Di Torino
Studenti)/0 - Tesi_HUI/1_PARTE_OPERATIVA/4_IMMAGINI
TERMICHE/2018/Agosto 2018/SRI/output_SRI_hd2.tif'

# Open the Landsat 8 bands using rasterio
with rasterio.open(red_band_path) as red_ds, \
    rasterio.open(green_band_path) as green_ds, \
    rasterio.open(blue_band_path) as blue_ds, \
    rasterio.open(nir_band_path) as nir_ds:

    # Read the band data
    red = red_ds.read(1).astype(np.float32)
    green = green_ds.read(1).astype(np.float32)
    blue = blue_ds.read(1).astype(np.float32)
    nir = nir_ds.read(1).astype(np.float32)

    # Calculate the spectral reflectance values
    reflectance = (nir / (red + green + blue),)

    # Calculate the Solar Reflectance Index (SRI)
    sri = 100 - (0.75 * (100 - reflectance[0])) - (0.15 * (1 - reflectance[0])) -
    (0.10 * (1 - reflectance[0]))

    # Write the SRI raster to output file
    profile = red_ds.profile # Use the profile of one of the bands
    profile.update(dtype=rasterio.float32, count=1, compress='lzw')
    with rasterio.open(output_sri_path, 'w', **profile) as sri_ds:
        sri_ds.write(sri, 1)
```

B. Hyper parameters tuning (codes and results)

Hyper parameter tuning (winter 2018)**Change parameters (blu part)*

```
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna
# List of raster file paths
raster_paths = [
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB
    2018/CROP/0 TIFF/NDVI.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB
    2018/CROP/0 TIFF/NDMI.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
    METEOREOLOGICA/2018_02_11/Turin TIFF/0 RASTER TIFF/humidity.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB
    2018/CROP/0 TIFF/EMISSIVITY.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
    METEOREOLOGICA/2018_02_11/Turin TIFF/0 RASTER TIFF/wind sector.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB
    2018/CROP/0 TIFF/ALBEDO.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/Buildings_30.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
    METEOREOLOGICA/2018_02_11/Turin TIFF/0 RASTER TIFF/temperature.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB 2018/CROP/0 TIFF/NDVI.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB 2018/CROP/0 TIFF/SRI.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
    METEOREOLOGICA/2018_02_11/Turin TIFF/0 RASTER TIFF/wind speed.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
    METEOREOLOGICA/2018_02_11/Turin TIFF/0 RASTER TIFF/solar rad.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/S V/S V To.tif',
    'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/5 DSM 1M/SKYVIEW_30M.tif'
]
reference_raster_path = 'C:/Users/ales/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/11 FEB
2018/LST/1 UHI Febbraio 18/UHI TO Febbraio 18.tif'
# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band
# Initialize an empty array to hold all the raster data
raster_data_list = []
from rasterio.warp import reproject, Resampling
# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
            src_transform=src.transform,
            src_crs=src.crs,
            dst_transform=ref_src.transform,
            dst_crs=ref_src.crs,
            resampling=Resampling.bilinear
        )
        raster_data_list.append(resampled_raster.flatten())
# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)
# Generate synthetic labels for demonstration purposes (replace with your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)
# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )
    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))
# Create an Optuna study
```

```
study = optuna.create_study(direction='maximize')
# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=55, n_jobs=-1)
# Get the best parameters from the study
best_params = study.best_params
# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)
# Fit the final model on the training data
best_rf.fit(X_train, y_train)
# Make predictions
y_pred = best_rf.predict(X_test)
# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
# Print best parameters, cross-validation results, and evaluation metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f'{param}: {value}')
print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)
print(f'Accuracy: {accuracy:.2f}')
print(f'F1-score: {f1:.2f}')
print(f'ROC-AUC: {roc_auc:.2f}')
[I 2023-08-17 10:32:37.727] A new study created in memory with name: no-name-9fd03530-6ba2-4354-99b5-bec4f08429ad
[I 2023-08-17 10:32:40.290] Trial 5 finished with value: 0.75 and parameters: {'n_estimators': 60, 'max_depth': 25, 'min_samples_split': 5, 'min_samples_leaf': 5}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:40.577] Trial 3 finished with value: 0.75 and parameters: {'n_estimators': 136, 'max_depth': 34, 'min_samples_split': 5, 'min_samples_leaf': 5}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:40.714] Trial 7 finished with value: 0.75 and parameters: {'n_estimators': 99, 'max_depth': 37, 'min_samples_split': 6, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.186] Trial 2 finished with value: 0.75 and parameters: {'n_estimators': 153, 'max_depth': 40, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.193] Trial 1 finished with value: 0.75 and parameters: {'n_estimators': 133, 'max_depth': 46, 'min_samples_split': 9, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.608] Trial 6 finished with value: 0.75 and parameters: {'n_estimators': 131, 'max_depth': 26, 'min_samples_split': 8, 'min_samples_leaf': 1}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.723] Trial 8 finished with value: 0.75 and parameters: {'n_estimators': 72, 'max_depth': 43, 'min_samples_split': 2, 'min_samples_leaf': 5}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.821] Trial 0 finished with value: 0.75 and parameters: {'n_estimators': 168, 'max_depth': 29, 'min_samples_split': 6, 'min_samples_leaf': 2}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:41.938] Trial 4 finished with value: 0.75 and parameters: {'n_estimators': 189, 'max_depth': 41, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:42.403] Trial 10 finished with value: 0.5 and parameters: {'n_estimators': 107, 'max_depth': 23, 'min_samples_split': 5, 'min_samples_leaf': 1}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:42.621] Trial 9 finished with value: 0.5833333333333334 and parameters: {'n_estimators': 178, 'max_depth': 30, 'min_samples_split': 3, 'min_samples_leaf': 2}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:42.670] Trial 11 finished with value: 0.75 and parameters: {'n_estimators': 120, 'max_depth': 26, 'min_samples_split': 3, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:42.872] Trial 12 finished with value: 0.75 and parameters: {'n_estimators': 125, 'max_depth': 34, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:43.198] Trial 15 finished with value: 0.5833333333333334 and parameters: {'n_estimators': 52, 'max_depth': 46, 'min_samples_split': 3, 'min_samples_leaf': 2}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:43.243] Trial 14 finished with value: 0.75 and parameters: {'n_estimators': 103, 'max_depth': 48, 'min_samples_split': 7, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:43.319] Trial 13 finished with value: 0.5 and parameters: {'n_estimators': 172, 'max_depth': 26, 'min_samples_split': 4, 'min_samples_leaf': 1}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:43.486] Trial 16 finished with value: 0.75 and parameters: {'n_estimators': 87, 'max_depth': 43, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:43.992] Trial 17 finished with value: 0.75 and parameters: {'n_estimators': 191, 'max_depth': 12, 'min_samples_split': 8, 'min_samples_leaf': 2}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:44.029] Trial 18 finished with value: 0.75 and parameters: {'n_estimators': 162, 'max_depth': 11, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:44.441] Trial 19 finished with value: 0.75 and parameters: {'n_estimators': 164, 'max_depth': 12, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:44.474] Trial 20 finished with value: 0.75 and parameters: {'n_estimators': 166, 'max_depth': 48, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:44.668] Trial 21 finished with value: 0.75 and parameters: {'n_estimators': 167, 'max_depth': 14, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:45.105] Trial 22 finished with value: 0.75 and parameters: {'n_estimators': 162, 'max_depth': 15, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:45.144] Trial 23 finished with value: 0.75 and parameters: {'n_estimators': 156, 'max_depth': 14, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:45.168] Trial 24 finished with value: 0.75 and parameters: {'n_estimators': 155, 'max_depth': 10, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:45.733] Trial 25 finished with value: 0.75 and parameters: {'n_estimators': 156, 'max_depth': 18, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:45.767] Trial 26 finished with value: 0.75 and parameters: {'n_estimators': 147, 'max_depth': 18, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:46.283] Trial 27 finished with value: 0.75 and parameters: {'n_estimators': 147, 'max_depth': 19, 'min_samples_split': 9, 'min_samples_leaf': 7}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:46.315] Trial 28 finished with value: 0.75 and parameters: {'n_estimators': 148, 'max_depth': 39, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.
[I 2023-08-17 10:32:46.331] Trial 29 finished with value: 0.75 and parameters: {'n_estimators': 148, 'max_depth': 39, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.
```


[I 2023-08-17 10:32:46,799] Trial 30 finished with value: 0.75 and parameters: {'n_estimators': 147, 'max_depth': 39, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:46,904] Trial 31 finished with value: 0.75 and parameters: {'n_estimators': 146, 'max_depth': 38, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:46,961] Trial 32 finished with value: 0.75 and parameters: {'n_estimators': 147, 'max_depth': 39, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:47,410] Trial 33 finished with value: 0.75 and parameters: {'n_estimators': 141, 'max_depth': 40, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:47,460] Trial 34 finished with value: 0.75 and parameters: {'n_estimators': 145, 'max_depth': 38, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:48,090] Trial 35 finished with value: 0.75 and parameters: {'n_estimators': 140, 'max_depth': 39, 'min_samples_split': 6, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:48,544] Trial 36 finished with value: 0.75 and parameters: {'n_estimators': 135, 'max_depth': 32, 'min_samples_split': 6, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:48,783] Trial 37 finished with value: 0.75 and parameters: {'n_estimators': 200, 'max_depth': 30, 'min_samples_split': 6, 'min_samples_leaf': 3}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:48,854] Trial 38 finished with value: 0.75 and parameters: {'n_estimators': 134, 'max_depth': 33, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:49,025] Trial 39 finished with value: 0.75 and parameters: {'n_estimators': 134, 'max_depth': 33, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:49,397] Trial 40 finished with value: 0.75 and parameters: {'n_estimators': 134, 'max_depth': 33, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:49,887] Trial 41 finished with value: 0.75 and parameters: {'n_estimators': 200, 'max_depth': 33, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:49,966] Trial 42 finished with value: 0.75 and parameters: {'n_estimators': 133, 'max_depth': 32, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:50,434] Trial 43 finished with value: 0.75 and parameters: {'n_estimators': 182, 'max_depth': 34, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:50,692] Trial 44 finished with value: 0.75 and parameters: {'n_estimators': 182, 'max_depth': 35, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:50,755] Trial 45 finished with value: 0.75 and parameters: {'n_estimators': 125, 'max_depth': 35, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:51,048] Trial 46 finished with value: 0.75 and parameters: {'n_estimators': 116, 'max_depth': 43, 'min_samples_split': 4, 'min_samples_leaf': 7}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:51,232] Trial 47 finished with value: 0.75 and parameters: {'n_estimators': 114, 'max_depth': 44, 'min_samples_split': 4, 'min_samples_leaf': 5}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:51,428] Trial 48 finished with value: 0.75 and parameters: {'n_estimators': 182, 'max_depth': 43, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:51,530] Trial 49 finished with value: 0.75 and parameters: {'n_estimators': 114, 'max_depth': 42, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:52,023] Trial 50 finished with value: 0.75 and parameters: {'n_estimators': 116, 'max_depth': 42, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:52,066] Trial 51 finished with value: 0.75 and parameters: {'n_estimators': 115, 'max_depth': 42, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:52,129] Trial 52 finished with value: 0.75 and parameters: {'n_estimators': 119, 'max_depth': 43, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:52,504] Trial 53 finished with value: 0.75 and parameters: {'n_estimators': 175, 'max_depth': 43, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

[I 2023-08-17 10:32:52,583] Trial 54 finished with value: 0.75 and parameters: {'n_estimators': 177, 'max_depth': 45, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 5 with value: 0.75.

Best Parameters:

n_estimators: 60

max_depth: 25

min_samples_split: 5

min_samples_leaf: 5

Cross-Validation Results:

	number	value	...	params_n_estimators	state
0	0	0.750000	...	168	COMPLETE
1	1	0.750000	...	133	COMPLETE
2	2	0.750000	...	153	COMPLETE
3	3	0.750000	...	136	COMPLETE
4	4	0.750000	...	189	COMPLETE
5	5	0.750000	...	60	COMPLETE
6	6	0.750000	...	131	COMPLETE
7	7	0.750000	...	99	COMPLETE
8	8	0.750000	...	72	COMPLETE
9	9	0.583333	...	178	COMPLETE
10	10	0.500000	...	107	COMPLETE
11	11	0.750000	...	120	COMPLETE
12	12	0.750000	...	125	COMPLETE
13	13	0.500000	...	172	COMPLETE
14	14	0.750000	...	103	COMPLETE
15	15	0.583333	...	52	COMPLETE
16	16	0.750000	...	87	COMPLETE
17	17	0.750000	...	191	COMPLETE
18	18	0.750000	...	162	COMPLETE
19	19	0.750000	...	164	COMPLETE
20	20	0.750000	...	166	COMPLETE
21	21	0.750000	...	167	COMPLETE
22	22	0.750000	...	162	COMPLETE
23	23	0.750000	...	156	COMPLETE
24	24	0.750000	...	155	COMPLETE
25	25	0.750000	...	156	COMPLETE
26	26	0.750000	...	147	COMPLETE
27	27	0.750000	...	147	COMPLETE
28	28	0.750000	...	148	COMPLETE
29	29	0.750000	...	148	COMPLETE
30	30	0.750000	...	147	COMPLETE
31	31	0.750000	...	146	COMPLETE
32	32	0.750000	...	147	COMPLETE
33	33	0.750000	...	141	COMPLETE

```

34 34 0.750000 ... 145 COMPLETE
35 35 0.750000 ... 140 COMPLETE
36 36 0.750000 ... 135 COMPLETE
37 37 0.750000 ... 200 COMPLETE
38 38 0.750000 ... 134 COMPLETE
39 39 0.750000 ... 134 COMPLETE
40 40 0.750000 ... 134 COMPLETE
41 41 0.750000 ... 200 COMPLETE
42 42 0.750000 ... 133 COMPLETE
43 43 0.750000 ... 182 COMPLETE
44 44 0.750000 ... 182 COMPLETE
45 45 0.750000 ... 125 COMPLETE
46 46 0.750000 ... 116 COMPLETE
47 47 0.750000 ... 114 COMPLETE
48 48 0.750000 ... 182 COMPLETE
49 49 0.750000 ... 114 COMPLETE
50 50 0.750000 ... 116 COMPLETE
51 51 0.750000 ... 115 COMPLETE
52 52 0.750000 ... 119 COMPLETE
53 53 0.750000 ... 175 COMPLETE
54 54 0.750000 ... 177 COMPLETE
[55 rows x 10 columns]

```

Hyper parameter tuning (Summer 2018)*Change parameters (blue part)

```

C:\Users\alless\AppData\Local\Programs\Python\Python311\python.exe "C:/Program Files/JetBrains/DataSpell 2023.2/plugins/python-ce/helpers/pydev/pydevconsole.py" -
-mode=client --host=127.0.0.1 --port=55595
import sys; print('Python %s on %s' % (sys.version, sys.platform))
sys.path.extend(['C:\Users\alless\Dropbox (Politecnico Di Torino Studenti)\Tesi_HUI\1 PARTE OPERATIVA\1 - MODELLO', 'C:\Users\alless\Dropbox
(Politecnico Di Torino Studenti)\Tesi_HUI\1 PARTE OPERATIVA\1 - MODELLO\Python all var', 'C:\Users\alless\Downloads'])
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna
# List of raster file paths
raster_paths = [
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO 2018/0 TIFF/NDWI.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO 2018/0 TIFF/NDMI.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/HUMIDITY.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO
2018/0 TIFF/EMISSIVITY.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/DIREZIONE VENTO.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO 2018/0 TIFF/ALBEDO.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/Buildings_30.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI METEOROLOGICA/2018_08_22/Air
temp/temperatura_to.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO 2018/0 TIFF/NDVI.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO 2018/0 TIFF/SRI.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/MODELLO_VELOCITA_VENTO.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/SOLAR RADIATION.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/S_V/S_V_To.tif',
    'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/5 DSM 1M/SKYVIEW_30M.tif'
]
reference_raster_path = 'C:/Users/alless/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/22 AGO
2018/LST_UHI/UHI/UHI_To_Ago18.tif'
# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band
# Initialize an empty array to hold all the raster data
raster_data_list = []
from rasterio.warp import reproject, Resampling
# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
            src_transform=src.transform,
            src_crs=src.crs,
            dst_transform=ref_src.transform,
            dst_crs=ref_src.crs,
            resampling=Resampling.bilinear
        )
        raster_data_list.append(resampled_raster.flatten())
# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)
# Generate synthetic labels for demonstration purposes (replace with your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)

```

```

# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )
    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))

# Create an Optuna study
study = optuna.create_study(direction='maximize')
# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=70, n_jobs=-1)
# Get the best parameters from the study
best_params = study.best_params
# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)
# Fit the final model on the training data
best_rf.fit(X_train, y_train)
# Make predictions
y_pred = best_rf.predict(X_test)
# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
# Print best parameters, cross-validation results, and evaluation metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f"{param}: {value}")
print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)
print(f"Accuracy: {accuracy:.2f}")
print(f"F1-score: {f1:.2f}")
print(f"ROC-AUC: {roc_auc:.2f}")

```

PyDev console: starting.

Python 3.11.4 (tags/v3.11.4:d2340ef, Jun 7 2023, 05:45:37) [MSC v.1934 64 bit (AMD64)] on win32

[I 2023-08-17 12:08:31.321] A new study created in memory with name: no-name-d5aaa270-2caf-4839-8465-5c19e6fb43da

[I 2023-08-17 12:08:34.941] Trial 5 finished with value: 0.5 and parameters: {'n_estimators': 131, 'max_depth': 18, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is trial 5 with value: 0.5.

[I 2023-08-17 12:08:35.832] Trial 0 finished with value: 0.5 and parameters: {'n_estimators': 179, 'max_depth': 31, 'min_samples_split': 10, 'min_samples_leaf': 2}. Best is trial 5 with value: 0.5.

[I 2023-08-17 12:08:35.978] Trial 6 finished with value: 1.0 and parameters: {'n_estimators': 200, 'max_depth': 31, 'min_samples_split': 2, 'min_samples_leaf': 1}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.038] Trial 2 finished with value: 0.5 and parameters: {'n_estimators': 97, 'max_depth': 42, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.066] Trial 1 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 75, 'max_depth': 38, 'min_samples_split': 5, 'min_samples_leaf': 1}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.422] Trial 3 finished with value: 0.5 and parameters: {'n_estimators': 105, 'max_depth': 50, 'min_samples_split': 3, 'min_samples_leaf': 8}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.512] Trial 4 finished with value: 0.5 and parameters: {'n_estimators': 195, 'max_depth': 25, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.745] Trial 7 finished with value: 0.5 and parameters: {'n_estimators': 162, 'max_depth': 38, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:36.990] Trial 8 finished with value: 0.5 and parameters: {'n_estimators': 146, 'max_depth': 48, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.088] Trial 10 finished with value: 0.5 and parameters: {'n_estimators': 59, 'max_depth': 49, 'min_samples_split': 2, 'min_samples_leaf': 4}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.173] Trial 9 finished with value: 0.5 and parameters: {'n_estimators': 141, 'max_depth': 50, 'min_samples_split': 10, 'min_samples_leaf': 1}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.475] Trial 12 finished with value: 0.5 and parameters: {'n_estimators': 74, 'max_depth': 14, 'min_samples_split': 8, 'min_samples_leaf': 6}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.765] Trial 11 finished with value: 0.5 and parameters: {'n_estimators': 144, 'max_depth': 32, 'min_samples_split': 5, 'min_samples_leaf': 6}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.921] Trial 13 finished with value: 0.5 and parameters: {'n_estimators': 109, 'max_depth': 12, 'min_samples_split': 5, 'min_samples_leaf': 4}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:37.999] Trial 15 finished with value: 0.5 and parameters: {'n_estimators': 88, 'max_depth': 35, 'min_samples_split': 7, 'min_samples_leaf': 7}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:38.085] Trial 14 finished with value: 0.5 and parameters: {'n_estimators': 131, 'max_depth': 22, 'min_samples_split': 3, 'min_samples_leaf': 4}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:38.767] Trial 18 finished with value: 0.8333333333333334 and parameters: {'n_estimators': 51, 'max_depth': 32, 'min_samples_split': 6, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:39.012] Trial 16 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 188, 'max_depth': 25, 'min_samples_split': 3, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:39.061] Trial 17 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 200, 'max_depth': 14, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:39.325] Trial 19 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 103, 'max_depth': 32, 'min_samples_split': 5, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:39.519] Trial 20 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 105, 'max_depth': 25, 'min_samples_split': 2, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.

[I 2023-08-17 12:08:39.603] Trial 21 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 86, 'max_depth': 38, 'min_samples_split': 2, 'min_samples_leaf': 2}. Best is trial 6 with value: 1.0.

209

```
[I 2023-08-17 12:08:55,553] Trial 68 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 197, 'max_depth': 27, 'min_samples_split': 3, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.
[I 2023-08-17 12:08:55,710] Trial 69 finished with value: 0.9166666666666666 and parameters: {'n_estimators': 196, 'max_depth': 27, 'min_samples_split': 3, 'min_samples_leaf': 3}. Best is trial 6 with value: 1.0.
```

```
Best Parameters:
n_estimators: 200
max_depth: 31
min_samples_split: 2
min_samples_leaf: 1
```

Cross-Validation Results:

	number	value	... params	n_estimators	state
0	0	0.500000	...	179	COMPLETE
1	1	0.916667	...	75	COMPLETE
2	2	0.500000	...	97	COMPLETE
3	3	0.500000	...	105	COMPLETE
4	4	0.500000	...	195	COMPLETE
...
65	65	0.500000	...	185	COMPLETE
66	66	0.500000	...	184	COMPLETE
67	67	0.500000	...	197	COMPLETE
68	68	0.916667	...	197	COMPLETE
69	69	0.916667	...	196	COMPLETE

[70 rows x 10 columns]

Hyper parameter tuning (Mid-season 2018)*Change parameters (blu part)

```
C:/Users/aless/AppData/Local/Programs/Python/Python311/python.exe "C:/Program Files/JetBrains/DataSpell 2023.2/plugins/python-ce/helpers/pydev/pydevconsole.py" -
-mode=client --host=127.0.0.1 --port=54424
import sys; print('Python %s on %s' % (sys.version, sys.platform))
sys.path.extend(['C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 - MODELLO', 'C:/Users/aless/Dropbox
(Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 - MODELLO/Python all var', 'C:/Users/aless/Downloads'])
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna
# List of raster file paths
raster_paths = [
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/CROP/0 TIFF/NDWI.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/CROP/0 TIFF/NDMI.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2018_04_16/Interpolation_2018_04_16/APRILE/Turin/0 RASTER TIFF/humidity.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/CROP/0 TIFF/EMISSION.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2018_04_16/Interpolation_2018_04_16/APRILE/Turin/0 RASTER TIFF/wind_degree.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/CROP/0 TIFF/ALBEDO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/Buildings_30.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2018_04_16/Interpolation_2018_04_16/APRILE/Turin/0 RASTER TIFF/temperature.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/CROP/0 TIFF/NDVI.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR 2018/CROP/0 TIFF/SRI.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2018_04_16/Interpolation_2018_04_16/APRILE/Turin/0 RASTER TIFF/wind_speed.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2018_04_16/Interpolation_2018_04_16/APRILE/Turin/0 RASTER TIFF/sol_rad.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/0 SHP REQUIREMENTS/Buildings/S_V/S_V_To.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/5 DSM 1M/SKYVIEW_30M.tif'
]
reference_raster_path = 'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2018/16 APR
2018/LST/UHI_TO_16_04_2018.tif'
# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band
# Initialize an empty array to hold all the raster data
raster_data_list = []
from rasterio.warp import reproject, Resampling
# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
```



```

src_transform=src.transform,
src_crs=src.crs,
dst_transform=ref_src.transform,
dst_crs=ref_src.crs,
resampling=Resampling.bilinear
)
raster_data_list.append(resampled_raster.flatten())
# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)
# Generate synthetic labels for demonstration purposes (replace with your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)
# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )
    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))
# Create an Optuna study
study = optuna.create_study(direction='maximize')
# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=55, n_jobs=-1)
# Get the best parameters from the study
best_params = study.best_params
# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)
# Fit the final model on the training data
best_rf.fit(X_train, y_train)
# Make predictions
y_pred = best_rf.predict(X_test)
# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
# Print best parameters, cross-validation results, and evaluation metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f"{param}: {value}")
print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)
print(f"Accuracy: {accuracy:.2f}")
print(f"F1-score: {f1:.2f}")
print(f"ROC-AUC: {roc_auc:.2f}")
PyDev console: starting.
Python 3.11.4 (tags/v3.11.4:d2340ef, Jun 7 2023, 05:45:37) [MSC v.1934 64 bit (AMD64)] on win32
[1 2023-08-17 10:50:50.948] A new study created in memory with name: no-name-c655009b-e4f9-4c27-bde9-a9fec4b767c
[1 2023-08-17 10:50:53.300] Trial 7 finished with value: 0.75 and parameters: {'n_estimators': 59, 'max_depth': 13, 'min_samples_split': 9, 'min_samples_leaf': 3}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:53.499] Trial 1 finished with value: 0.75 and parameters: {'n_estimators': 64, 'max_depth': 21, 'min_samples_split': 10, 'min_samples_leaf': 3}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:53.583] Trial 2 finished with value: 0.75 and parameters: {'n_estimators': 142, 'max_depth': 33, 'min_samples_split': 9, 'min_samples_leaf': 7}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:53.781] Trial 6 finished with value: 0.75 and parameters: {'n_estimators': 93, 'max_depth': 16, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:53.873] Trial 5 finished with value: 0.75 and parameters: {'n_estimators': 198, 'max_depth': 34, 'min_samples_split': 4, 'min_samples_leaf': 7}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:53.889] Trial 3 finished with value: 0.75 and parameters: {'n_estimators': 75, 'max_depth': 22, 'min_samples_split': 9, 'min_samples_leaf': 8}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.149] Trial 0 finished with value: 0.75 and parameters: {'n_estimators': 121, 'max_depth': 44, 'min_samples_split': 6, 'min_samples_leaf': 4}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.201] Trial 4 finished with value: 0.75 and parameters: {'n_estimators': 164, 'max_depth': 38, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.311] Trial 8 finished with value: 0.75 and parameters: {'n_estimators': 145, 'max_depth': 28, 'min_samples_split': 10, 'min_samples_leaf': 5}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.579] Trial 9 finished with value: 0.75 and parameters: {'n_estimators': 133, 'max_depth': 39, 'min_samples_split': 10, 'min_samples_leaf': 6}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.624] Trial 11 finished with value: 0.75 and parameters: {'n_estimators': 111, 'max_depth': 41, 'min_samples_split': 6, 'min_samples_leaf': 6}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:54.660] Trial 10 finished with value: 0.75 and parameters: {'n_estimators': 170, 'max_depth': 20, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:55.002] Trial 13 finished with value: 0.75 and parameters: {'n_estimators': 100, 'max_depth': 42, 'min_samples_split': 8, 'min_samples_leaf': 3}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:55.064] Trial 12 finished with value: 0.75 and parameters: {'n_estimators': 173, 'max_depth': 48, 'min_samples_split': 9, 'min_samples_leaf': 3}. Best is trial 7 with value: 0.75.
[1 2023-08-17 10:50:55.373] Trial 16 finished with value: 0.75 and parameters: {'n_estimators': 73, 'max_depth': 45, 'min_samples_split': 2, 'min_samples_leaf': 7}. Best is trial 7 with value: 0.75.

```

```

[I 2023-08-17 10:50:55.408] Trial 14 finished with value: 0.75 and parameters: {'n_estimators': 146, 'max_depth': 29, 'min_samples_split': 6, 'min_samples_leaf': 2}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:55.446] Trial 15 finished with value: 0.75 and parameters: {'n_estimators': 195, 'max_depth': 14, 'min_samples_split': 9, 'min_samples_leaf': 4}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:55.583] Trial 17 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 104, 'max_depth': 49, 'min_samples_split': 2,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:55.870] Trial 18 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 99, 'max_depth': 48, 'min_samples_split': 2,
'min_samples_leaf': 2}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:55.948] Trial 19 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 54, 'max_depth': 50, 'min_samples_split': 2,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:55.980] Trial 21 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 53, 'max_depth': 26, 'min_samples_split': 2,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.026] Trial 20 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 54, 'max_depth': 49, 'min_samples_split': 2,
'min_samples_leaf': 2}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.213] Trial 22 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 56, 'max_depth': 26, 'min_samples_split': 4,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.251] Trial 23 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 50, 'max_depth': 23, 'min_samples_split': 4,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.384] Trial 25 finished with value: 0.75 and parameters: {'n_estimators': 53, 'max_depth': 24, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.523] Trial 26 finished with value: 0.75 and parameters: {'n_estimators': 56, 'max_depth': 23, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.614] Trial 24 finished with value: 0.6666666666666666 and parameters: {'n_estimators': 118, 'max_depth': 23, 'min_samples_split': 4,
'min_samples_leaf': 1}. Best is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.706] Trial 27 finished with value: 0.75 and parameters: {'n_estimators': 118, 'max_depth': 23, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:56.934] Trial 28 finished with value: 0.75 and parameters: {'n_estimators': 126, 'max_depth': 33, 'min_samples_split': 4, 'min_samples_leaf': 5}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.091] Trial 29 finished with value: 0.75 and parameters: {'n_estimators': 123, 'max_depth': 33, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.171] Trial 30 finished with value: 0.75 and parameters: {'n_estimators': 133, 'max_depth': 34, 'min_samples_split': 8, 'min_samples_leaf': 4}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.347] Trial 31 finished with value: 0.75 and parameters: {'n_estimators': 126, 'max_depth': 10, 'min_samples_split': 8, 'min_samples_leaf': 4}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.498] Trial 32 finished with value: 0.75 and parameters: {'n_estimators': 120, 'max_depth': 35, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.648] Trial 33 finished with value: 0.75 and parameters: {'n_estimators': 123, 'max_depth': 32, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.806] Trial 34 finished with value: 0.75 and parameters: {'n_estimators': 132, 'max_depth': 33, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.841] Trial 35 finished with value: 0.75 and parameters: {'n_estimators': 135, 'max_depth': 31, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:57.857] Trial 36 finished with value: 0.75 and parameters: {'n_estimators': 84, 'max_depth': 33, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.074] Trial 37 finished with value: 0.75 and parameters: {'n_estimators': 83, 'max_depth': 19, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.179] Trial 38 finished with value: 0.75 and parameters: {'n_estimators': 77, 'max_depth': 18, 'min_samples_split': 8, 'min_samples_leaf': 8}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.265] Trial 39 finished with value: 0.75 and parameters: {'n_estimators': 74, 'max_depth': 19, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.299] Trial 40 finished with value: 0.75 and parameters: {'n_estimators': 75, 'max_depth': 19, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.496] Trial 41 finished with value: 0.75 and parameters: {'n_estimators': 80, 'max_depth': 18, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.590] Trial 42 finished with value: 0.75 and parameters: {'n_estimators': 82, 'max_depth': 17, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is
trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.670] Trial 44 finished with value: 0.75 and parameters: {'n_estimators': 75, 'max_depth': 19, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.712] Trial 45 finished with value: 0.75 and parameters: {'n_estimators': 65, 'max_depth': 17, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.784] Trial 43 finished with value: 0.75 and parameters: {'n_estimators': 82, 'max_depth': 20, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.812] Trial 46 finished with value: 0.75 and parameters: {'n_estimators': 69, 'max_depth': 38, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:58.897] Trial 47 finished with value: 0.75 and parameters: {'n_estimators': 66, 'max_depth': 37, 'min_samples_split': 10, 'min_samples_leaf': 6}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:59.356] Trial 48 finished with value: 0.75 and parameters: {'n_estimators': 162, 'max_depth': 37, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:59.441] Trial 49 finished with value: 0.75 and parameters: {'n_estimators': 159, 'max_depth': 37, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:59.768] Trial 50 finished with value: 0.75 and parameters: {'n_estimators': 163, 'max_depth': 37, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:59.843] Trial 51 finished with value: 0.75 and parameters: {'n_estimators': 162, 'max_depth': 37, 'min_samples_split': 9, 'min_samples_leaf': 7}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:50:59.913] Trial 52 finished with value: 0.75 and parameters: {'n_estimators': 160, 'max_depth': 38, 'min_samples_split': 9, 'min_samples_leaf': 3}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:51:00.193] Trial 53 finished with value: 0.75 and parameters: {'n_estimators': 156, 'max_depth': 38, 'min_samples_split': 9, 'min_samples_leaf': 3}. Best
is trial 7 with value: 0.75.
[I 2023-08-17 10:51:00.225] Trial 54 finished with value: 0.75 and parameters: {'n_estimators': 154, 'max_depth': 42, 'min_samples_split': 9, 'min_samples_leaf': 3}. Best
is trial 7 with value: 0.75.
Best Parameters:
n_estimators: 59
max_depth: 13
min_samples_split: 9
min_samples_leaf: 3
Cross-Validation Results:

```

	number	value	...	params	n_estimators	state
0	0	0.750000	...		121	COMPLETE
1	1	0.750000	...		64	COMPLETE
2	2	0.750000	...		142	COMPLETE
3	3	0.750000	...		75	COMPLETE
4	4	0.750000	...		164	COMPLETE
5	5	0.750000	...		198	COMPLETE

```
6 6 0.750000 ... 93 COMPLETE
7 7 0.750000 ... 59 COMPLETE
8 8 0.750000 ... 145 COMPLETE
9 9 0.750000 ... 133 COMPLETE
10 10 0.750000 ... 170 COMPLETE
11 11 0.750000 ... 111 COMPLETE
12 12 0.750000 ... 173 COMPLETE
13 13 0.750000 ... 100 COMPLETE
14 14 0.750000 ... 146 COMPLETE
15 15 0.750000 ... 195 COMPLETE
16 16 0.750000 ... 73 COMPLETE
17 17 0.666667 ... 104 COMPLETE
18 18 0.666667 ... 99 COMPLETE
19 19 0.666667 ... 54 COMPLETE
20 20 0.666667 ... 54 COMPLETE
21 21 0.666667 ... 53 COMPLETE
22 22 0.666667 ... 56 COMPLETE
23 23 0.666667 ... 50 COMPLETE
24 24 0.666667 ... 118 COMPLETE
25 25 0.750000 ... 53 COMPLETE
26 26 0.750000 ... 56 COMPLETE
27 27 0.750000 ... 118 COMPLETE
28 28 0.750000 ... 126 COMPLETE
29 29 0.750000 ... 123 COMPLETE
30 30 0.750000 ... 133 COMPLETE
31 31 0.750000 ... 126 COMPLETE
32 32 0.750000 ... 120 COMPLETE
33 33 0.750000 ... 123 COMPLETE
34 34 0.750000 ... 132 COMPLETE
35 35 0.750000 ... 135 COMPLETE
36 36 0.750000 ... 84 COMPLETE
37 37 0.750000 ... 83 COMPLETE
38 38 0.750000 ... 77 COMPLETE
39 39 0.750000 ... 74 COMPLETE
40 40 0.750000 ... 75 COMPLETE
41 41 0.750000 ... 80 COMPLETE
42 42 0.750000 ... 82 COMPLETE
43 43 0.750000 ... 82 COMPLETE
44 44 0.750000 ... 75 COMPLETE
45 45 0.750000 ... 65 COMPLETE
46 46 0.750000 ... 69 COMPLETE
47 47 0.750000 ... 66 COMPLETE
48 48 0.750000 ... 162 COMPLETE
49 49 0.750000 ... 159 COMPLETE
50 50 0.750000 ... 163 COMPLETE
51 51 0.750000 ... 162 COMPLETE
52 52 0.750000 ... 160 COMPLETE
53 53 0.750000 ... 156 COMPLETE
54 54 0.750000 ... 154 COMPLETE
[55 rows x 10 columns]
```

Hyper parameter tuning (Winter 2000)*Change parameters (blue part)

```
C:\Users\aless\AppData\Local\Programs\Python\Python311\python.exe "C:/Program Files/JetBrains/DataSpell 2023.2/plugins/python-ce/helpers/pydev/pydevconsole.py" -
-mode=client --host=127.0.0.1 --port=64124
import sys; print('Python %s on %s' % (sys.version, sys.platform))
sys.path.extend(['C:\Users\aless\Dropbox (Politecnico Di Torino Studenti)\Tesi HUI\1 PARTE OPERATIVA\1 - MODELLO', 'C:\Users\aless\Dropbox
(Politecnico Di Torino Studenti)\Tesi HUI\1 PARTE OPERATIVA\1 - MODELLO\Python all var', 'C:\Users\aless\Downloads'])
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna
# List of raster file paths
raster_paths = [
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/NDWI TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/NDMI TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2000 12 18/Turin Tiff/0 RASTER TIFF/humidity.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/EMISSIVITY TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2000 12 18/Turin Tiff/0 RASTER TIFF/wind_sector.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/ALBEDO TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2000 12 18/Turin Tiff/0 RASTER TIFF/temperature.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/NDVI TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/SRI TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2000 12 18/Turin Tiff/0 RASTER TIFF/wind_speed.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2000 12 18/Turin Tiff/0 RASTER TIFF/sol_rad.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Dicembre
2000/0 CROP/NDBI TO.tif'
]
reference_raster_path = 'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT
7/LE07 L1TP 195029 20001218 20211122 02 TI/CALCULATION NEW/UHI/UHI TO 2000 12 18.tif'
# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band
# Initialize an empty array to hold all the raster data
raster_data_list = []
from rasterio.warp import reproject, Resampling
# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
            src_transform=src.transform,
            src_crs=src.crs,
            dst_transform=ref_src.transform,
            dst_crs=ref_src.crs,
            resampling=Resampling.bilinear
        )
        raster_data_list.append(resampled_raster.flatten())
# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)
# Generate synthetic labels for demonstration purposes (replace with your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)
# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )
    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))
# Create an Optuna study
```

```

study = optuna.create_study(direction='maximize')
# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=33, n_jobs=-1)
# Get the best parameters from the study
best_params = study.best_params
# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)
# Fit the final model on the training data
best_rf.fit(X_train, y_train)
# Make predictions
y_pred = best_rf.predict(X_test)
# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
# Print best parameters, cross-validation results, and evaluation metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f'{param}: {value}')
print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)
print(f"Accuracy: {accuracy:.2f}")
print(f"F1-score: {f1:.2f}")
print(f"ROC-AUC: {roc_auc:.2f}")
PyDev console: starting.
Python 3.11.4 (tags/v3.11.4:d2340ef, Jun 7 2023, 05:45:37) [MSC v.1934 64 bit (AMD64)] on win32
[1 2023-08-17 17:19:00.516] A new study created in memory with name: no-name-ab1139e6-2f89-406b-960e-970a8c92271d
[1 2023-08-17 17:19:03.594] Trial 2 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 158, 'max_depth': 11, 'min_samples_split': 6, 'min_samples_leaf': 1}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:03.733] Trial 6 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 178, 'max_depth': 13, 'min_samples_split': 6, 'min_samples_leaf': 4}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:03.787] Trial 1 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 146, 'max_depth': 18, 'min_samples_split': 6, 'min_samples_leaf': 8}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:03.936] Trial 0 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 130, 'max_depth': 47, 'min_samples_split': 9, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:04.260] Trial 3 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 79, 'max_depth': 18, 'min_samples_split': 3, 'min_samples_leaf': 4}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:04.408] Trial 5 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 163, 'max_depth': 36, 'min_samples_split': 3, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:04.434] Trial 7 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 68, 'max_depth': 38, 'min_samples_split': 6, 'min_samples_leaf': 4}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:04.460] Trial 4 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 78, 'max_depth': 34, 'min_samples_split': 8, 'min_samples_leaf': 3}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:04.623] Trial 9 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 54, 'max_depth': 45, 'min_samples_split': 7, 'min_samples_leaf': 8}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:05.107] Trial 10 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 122, 'max_depth': 48, 'min_samples_split': 7, 'min_samples_leaf': 1}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:05.247] Trial 8 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 198, 'max_depth': 50, 'min_samples_split': 9, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:05.322] Trial 11 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 77, 'max_depth': 21, 'min_samples_split': 8, 'min_samples_leaf': 1}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:05.777] Trial 12 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 156, 'max_depth': 42, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:05.778] Trial 13 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 131, 'max_depth': 34, 'min_samples_split': 6, 'min_samples_leaf': 8}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:06.009] Trial 14 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 63, 'max_depth': 12, 'min_samples_split': 4, 'min_samples_leaf': 4}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:06.125] Trial 15 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 108, 'max_depth': 27, 'min_samples_split': 3, 'min_samples_leaf': 3}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:06.400] Trial 16 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 145, 'max_depth': 37, 'min_samples_split': 3, 'min_samples_leaf': 5}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:06.690] Trial 17 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 112, 'max_depth': 28, 'min_samples_split': 10, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:06.794] Trial 18 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 124, 'max_depth': 23, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:07.422] Trial 19 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 122, 'max_depth': 25, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:07.457] Trial 21 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 115, 'max_depth': 28, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:07.833] Trial 20 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 121, 'max_depth': 27, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:07.976] Trial 22 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 116, 'max_depth': 27, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:08.183] Trial 23 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 137, 'max_depth': 26, 'min_samples_split': 10, 'min_samples_leaf': 6}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:08.245] Trial 24 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 102, 'max_depth': 28, 'min_samples_split': 10, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:08.612] Trial 26 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 95, 'max_depth': 17, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:08.644] Trial 25 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 138, 'max_depth': 20, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:08.935] Trial 27 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 91, 'max_depth': 17, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[1 2023-08-17 17:19:09.430] Trial 28 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 141, 'max_depth': 16, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.

```



```
[I 2023-08-17 17:19:09.493] Trial 29 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 144, 'max_depth': 16, 'min_samples_split': 5, 'min_samples_leaf': 7}. Best is trial 2 with value: 0.6111111111111111.
[I 2023-08-17 17:19:09.669] Trial 30 finished with value: 0.6111111111111111 and parameters: {'n_estimators': 141, 'max_depth': 17, 'min_samples_split': 7, 'min_samples_leaf': 2}. Best is trial 2 with value: 0.6111111111111111.
[I 2023-08-17 17:19:09.783] Trial 31 finished with value: 0.7222222222222222 and parameters: {'n_estimators': 97, 'max_depth': 16, 'min_samples_split': 5, 'min_samples_leaf': 2}. Best is trial 31 with value: 0.7222222222222222.
[I 2023-08-17 17:19:10.229] Trial 32 finished with value: 0.7222222222222222 and parameters: {'n_estimators': 176, 'max_depth': 16, 'min_samples_split': 5, 'min_samples_leaf': 2}. Best is trial 31 with value: 0.7222222222222222.
Best Parameters:
n_estimators: 97
max_depth: 16
min_samples_split: 5
min_samples_leaf: 2
Cross-Validation Results:
  number  value  ...  params  n_estimators  state
0      0  0.611111  ...           130  COMPLETE
1      1  0.611111  ...           146  COMPLETE
2      2  0.611111  ...           158  COMPLETE
3      3  0.611111  ...            79  COMPLETE
4      4  0.611111  ...            78  COMPLETE
5      5  0.611111  ...           163  COMPLETE
6      6  0.611111  ...           178  COMPLETE
7      7  0.611111  ...            68  COMPLETE
8      8  0.611111  ...           198  COMPLETE
9      9  0.611111  ...            54  COMPLETE
10     10  0.611111  ...           122  COMPLETE
11     11  0.611111  ...            77  COMPLETE
12     12  0.611111  ...           156  COMPLETE
13     13  0.611111  ...           131  COMPLETE
14     14  0.611111  ...            63  COMPLETE
15     15  0.611111  ...           108  COMPLETE
16     16  0.611111  ...           145  COMPLETE
17     17  0.611111  ...           112  COMPLETE
18     18  0.611111  ...           124  COMPLETE
19     19  0.611111  ...           122  COMPLETE
20     20  0.611111  ...           121  COMPLETE
21     21  0.611111  ...           115  COMPLETE
22     22  0.611111  ...           116  COMPLETE
23     23  0.611111  ...           137  COMPLETE
24     24  0.611111  ...           102  COMPLETE
25     25  0.611111  ...           138  COMPLETE
26     26  0.611111  ...            95  COMPLETE
27     27  0.611111  ...            91  COMPLETE
28     28  0.611111  ...           141  COMPLETE
29     29  0.611111  ...           144  COMPLETE
30     30  0.611111  ...           141  COMPLETE
31     31  0.722222  ...            97  COMPLETE
32     32  0.722222  ...           176  COMPLETE
[33 rows x 10 columns]
```

Hyper parameter tuning (Summer 2001)*Change parameters (blue part)

```
C:\Users\alless\AppData\Local\Programs\Python\Python311\python.exe "C:/Program Files/JetBrains/DataSpell 2023.2/plugins/python-ce/helpers/pydev/pydevconsole.py" -
-mode=client --host=127.0.0.1 --port=63390
import sys; print('Python %s on %s' % (sys.version, sys.platform))
sys.path.extend(['C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO', 'C:/Users/aleess/Dropbox
(Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/Python all var', 'C:/Users/aleess/Downloads'])
>>> import numpy as np
... from sklearn.model_selection import train_test_split, cross_val_score
... from sklearn.ensemble import RandomForestClassifier
... from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
... import rasterio
... import optuna
...
... # List of raster file paths
... raster_paths = [
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/NDWI.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/NDMI.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2001_08_24/Turin_Tiff/0_RASTER_TIFF/humidity.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/EMISSIONITY.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2001_08_24/Turin_Tiff/0_RASTER_TIFF/wind_sector.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/ALBEDO.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2001_08_24/Turin_Tiff/0_RASTER_TIFF/temperature.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/NDVI.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/SRI.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2001_08_24/Turin_Tiff/0_RASTER_TIFF/wind_speed.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEOREOLOGICA/2001_08_24/Turin_Tiff/0_RASTER_TIFF/solar_rad.tif',
...     'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Agosto
2001/0_CROP/NDBI.tif'
... ]
...
... reference_raster_path = 'C:/Users/aleess/Dropbox (Politecnico Di Torino Studenti)/Tesi_HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT
7/LE07_L1TP_194029_20010824_20200917_02_T1/CALCULATIONS - NEW/UHI/UHI_TO_24_08_2001.tif'
...
... # Load reference raster using rasterio
... with rasterio.open(reference_raster_path) as ref_src:
...     reference_raster = ref_src.read(1) # Read the first band
...
... # Initialize an empty array to hold all the raster data
... raster_data_list = []
...
... from rasterio.warp import reproject, Resampling
...
... # Loop through raster paths and load data into the list
... for path in raster_paths:
...     with rasterio.open(path) as src:
...         raster_data = src.read(1) # Read the first band
...         # Reproject the raster data to match the reference raster
...         resampled_raster = np.empty_like(reference_raster)
...         reproject(
...             source=raster_data,
...             destination=resampled_raster,
...             src_transform=src.transform,
...             src_crs=src.crs,
...             dst_transform=ref_src.transform,
...             dst_crs=ref_src.crs,
...             resampling=Resampling.bilinear
...         )
...         raster_data_list.append(resampled_raster.flatten())
...
... # Combine all raster data into a single array
... raster_data_array = np.array(raster_data_list)
...
... # Generate synthetic labels for demonstration purposes (replace with your actual labels)
... labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
...
... # Split the data into training and testing sets
... X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)
...
... # Define the objective function to optimize
... def objective(trial, X, y):
...     n_estimators = trial.suggest_int('n_estimators', 50, 200)
...     max_depth = trial.suggest_int('max_depth', 10, 50)
...     min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
...     min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
...
...     rf = RandomForestClassifier(
...         n_estimators=n_estimators,
```

```

...     max_depth=max_depth,
...     min_samples_split=min_samples_split,
...     min_samples_leaf=min_samples_leaf,
...     random_state=42
... )
...
... return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))
...
... # Create an Optuna study
... study = optuna.create_study(direction='maximize')
...
... # Optimize the study
... study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=30, n_jobs=-1)
...
... # Get the best parameters from the study
... best_params = study.best_params
...
... # Create the final random forest classifier with the best parameters
... best_rf = RandomForestClassifier(
...     n_estimators=best_params['n_estimators'],
...     max_depth=best_params['max_depth'],
...     min_samples_split=best_params['min_samples_split'],
...     min_samples_leaf=best_params['min_samples_leaf'],
...     random_state=42
... )
...
... # Fit the final model on the training data
... best_rf.fit(X_train, y_train)
...
... # Make predictions
... y_pred = best_rf.predict(X_test)
...
... # Calculate accuracy, F1-score, and ROC-AUC
... accuracy = accuracy_score(y_test, y_pred)
... f1 = f1_score(y_test, y_pred)
... roc_auc = roc_auc_score(y_test, y_pred)
...
... # Print best parameters, cross-validation results, and evaluation metrics
... print("Best Parameters:")
... for param, value in best_params.items():
...     print(f'{param}: {value}')
...
... print("Cross-Validation Results:")
... cv_results = study.trials_dataframe()
... print(cv_results)
...
... print(f"Accuracy: {accuracy:.2f}")
... print(f"F1-score: {f1:.2f}")
... print(f"ROC-AUC: {roc_auc:.2f}")
...
... >>>
... >>>
... >>>
PyDev console: starting.
Python 3.11.4 (tags/v3.11.4:d2340ef, Jun 7 2023, 05:45:37) [MSC v.1934 64 bit (AMD64)] on win32
[1 2023-08-18 09:01:37.130] A new study created in memory with name: no-name-a88d15ea-ef09-46ff-826b-c304150a433a
[1 2023-08-18 09:01:39.769] Trial 6 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 59, 'max_depth': 50, 'min_samples_split': 4, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:39.907] Trial 0 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 64, 'max_depth': 33, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:39.935] Trial 5 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 165, 'max_depth': 39, 'min_samples_split': 2, 'min_samples_leaf': 3}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:40.041] Trial 4 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 93, 'max_depth': 50, 'min_samples_split': 4, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:40.541] Trial 1 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 145, 'max_depth': 21, 'min_samples_split': 5, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:40.645] Trial 2 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 183, 'max_depth': 46, 'min_samples_split': 7, 'min_samples_leaf': 5}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.110] Trial 7 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 100, 'max_depth': 18, 'min_samples_split': 10, 'min_samples_leaf': 8}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.157] Trial 3 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 164, 'max_depth': 19, 'min_samples_split': 7, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.184] Trial 8 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 111, 'max_depth': 20, 'min_samples_split': 8, 'min_samples_leaf': 3}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.412] Trial 10 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 67, 'max_depth': 21, 'min_samples_split': 6, 'min_samples_leaf': 3}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.678] Trial 12 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 72, 'max_depth': 16, 'min_samples_split': 9, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:41.780] Trial 9 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 191, 'max_depth': 38, 'min_samples_split': 4, 'min_samples_leaf': 8}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.106] Trial 11 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 191, 'max_depth': 12, 'min_samples_split': 7, 'min_samples_leaf': 5}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.299] Trial 13 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 132, 'max_depth': 12, 'min_samples_split': 8, 'min_samples_leaf': 7}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.365] Trial 14 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 115, 'max_depth': 36, 'min_samples_split': 7, 'min_samples_leaf': 5}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.824] Trial 15 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 135, 'max_depth': 23, 'min_samples_split': 5, 'min_samples_leaf': 4}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.885] Trial 18 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 51, 'max_depth': 32, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[1 2023-08-18 09:01:42.939] Trial 17 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 56, 'max_depth': 40, 'min_samples_split': 4, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.

```

```
[I 2023-08-18 09:01:42.976] Trial 16 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 166, 'max_depth': 29, 'min_samples_split': 7, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:43.241] Trial 19 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 53, 'max_depth': 30, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:43.311] Trial 20 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 50, 'max_depth': 30, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:43.350] Trial 21 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 52, 'max_depth': 29, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:43.545] Trial 22 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 50, 'max_depth': 32, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:43.702] Trial 23 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 50, 'max_depth': 29, 'min_samples_split': 2, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.011] Trial 24 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 78, 'max_depth': 27, 'min_samples_split': 3, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.121] Trial 25 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 88, 'max_depth': 44, 'min_samples_split': 3, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.251] Trial 26 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 83, 'max_depth': 44, 'min_samples_split': 3, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.345] Trial 29 finished with value: 0.6944444444444443 and parameters: {'n_estimators': 78, 'max_depth': 43, 'min_samples_split': 3, 'min_samples_leaf': 4}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.396] Trial 27 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 83, 'max_depth': 43, 'min_samples_split': 4, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
[I 2023-08-18 09:01:44.559] Trial 28 finished with value: 0.8055555555555555 and parameters: {'n_estimators': 80, 'max_depth': 45, 'min_samples_split': 3, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8055555555555555.
Best Parameters:
n_estimators: 50
max_depth: 50
min_samples_split: 4
min_samples_leaf: 1
Cross-Validation Results:
  number  value  ...  params  n_estimators  state
0    0  0.694444  ...      64  COMPLETE
1    1  0.694444  ...     145  COMPLETE
2    2  0.694444  ...     183  COMPLETE
3    3  0.694444  ...     164  COMPLETE
4    4  0.694444  ...     93  COMPLETE
5    5  0.694444  ...     165  COMPLETE
6    6  0.805556  ...     59  COMPLETE
7    7  0.694444  ...     100  COMPLETE
8    8  0.694444  ...     111  COMPLETE
9    9  0.694444  ...     191  COMPLETE
10   10  0.694444  ...     67  COMPLETE
11   11  0.694444  ...     191  COMPLETE
12   12  0.694444  ...     72  COMPLETE
13   13  0.694444  ...     132  COMPLETE
14   14  0.694444  ...     115  COMPLETE
15   15  0.694444  ...     135  COMPLETE
16   16  0.694444  ...     166  COMPLETE
17   17  0.805556  ...     56  COMPLETE
18   18  0.694444  ...     51  COMPLETE
19   19  0.694444  ...     53  COMPLETE
20   20  0.694444  ...     50  COMPLETE
21   21  0.694444  ...     52  COMPLETE
22   22  0.694444  ...     50  COMPLETE
23   23  0.694444  ...     50  COMPLETE
24   24  0.805556  ...     78  COMPLETE
25   25  0.805556  ...     88  COMPLETE
26   26  0.805556  ...     83  COMPLETE
27   27  0.805556  ...     83  COMPLETE
28   28  0.805556  ...     80  COMPLETE
29   29  0.694444  ...     78  COMPLETE
[30 rows x 10 columns]
```

Hyper parameter tuning (Mid-season 2001)*Change parameters (blu part)

```
C:\Users\aless\AppData\Local\Programs\Python\Python311\python.exe "C:/Program Files/JetBrains/DataSpell 2023.2/plugins/python-ce/helpers/pydev/pydevconsole.py" -
-mode=client --host=127.0.0.1 --port=62412
import sys; print('Python %s on %s' % (sys.version, sys.platform))
sys.path.extend(['C:\Users\aless\Dropbox (Politecnico Di Torino Studenti)\Tesi HUI\1 PARTE OPERATIVA\1 - MODELLO', 'C:\Users\aless\Dropbox
(Politecnico Di Torino Studenti)\Tesi HUI\1 PARTE OPERATIVA\1 - MODELLO\Python all var', 'C:\Users\aless\Downloads'])
import numpy as np
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, f1_score, roc_auc_score
import rasterio
import optuna
# List of raster file paths
raster_paths = [
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/NDWI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/NDMI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/humidity.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/EMISSIVITY_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 - MODELLO/DTM.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/wind_sector.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/ALBEDO_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/temperature.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/NDVI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/SRI_TO.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/wind_speed.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/1 BANCA DATI
METEREOLOGICA/2001_05_27/Turin_Tiff/0_RASTER_TIFF/solar_rad.tif',
    'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT 7/CALCOLI Maggio
2001/0_CROP/NDBI_TO.tif',
]
reference_raster_path = 'C:/Users/aless/Dropbox (Politecnico Di Torino Studenti)/Tesi HUI/1 PARTE OPERATIVA/4 IMMAGINI TERMICHE/2001/LANDSAT
7/LE07_L1TP_195029_20010527_20200917_02_T1/CALCULATION_NEW/UHI/UHI_TO_27_05_2001.tif'
# Load reference raster using rasterio
with rasterio.open(reference_raster_path) as ref_src:
    reference_raster = ref_src.read(1) # Read the first band
# Initialize an empty array to hold all the raster data
raster_data_list = []
from rasterio.warp import reproject, Resampling
# Loop through raster paths and load data into the list
for path in raster_paths:
    with rasterio.open(path) as src:
        raster_data = src.read(1) # Read the first band
        # Reproject the raster data to match the reference raster
        resampled_raster = np.empty_like(reference_raster)
        reproject(
            source=raster_data,
            destination=resampled_raster,
            src_transform=src.transform,
            src_crs=src.crs,
            dst_transform=ref_src.transform,
            dst_crs=ref_src.crs,
            resampling=Resampling.bilinear
        )
        raster_data_list.append(resampled_raster.flatten())
# Combine all raster data into a single array
raster_data_array = np.array(raster_data_list)
# Generate synthetic labels for demonstration purposes (replace with your actual labels)
labels = np.random.randint(0, 2, size=raster_data_array.shape[0])
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(raster_data_array, labels, test_size=0.2, random_state=42)
# Define the objective function to optimize
def objective(trial, X, y):
    n_estimators = trial.suggest_int('n_estimators', 50, 200)
    max_depth = trial.suggest_int('max_depth', 10, 50)
    min_samples_split = trial.suggest_int('min_samples_split', 2, 10)
    min_samples_leaf = trial.suggest_int('min_samples_leaf', 1, 8)
    rf = RandomForestClassifier(
        n_estimators=n_estimators,
        max_depth=max_depth,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )
    return np.mean(cross_val_score(rf, X, y, n_jobs=-1, cv=3, scoring='accuracy'))
# Create an Optuna study
study = optuna.create_study(direction='maximize')
# Optimize the study
study.optimize(lambda trial: objective(trial, X_train, y_train), n_trials=15, n_jobs=-1)
# Get the best parameters from the study
```



```

best_params = study.best_params
# Create the final random forest classifier with the best parameters
best_rf = RandomForestClassifier(
    n_estimators=best_params['n_estimators'],
    max_depth=best_params['max_depth'],
    min_samples_split=best_params['min_samples_split'],
    min_samples_leaf=best_params['min_samples_leaf'],
    random_state=42
)
# Fit the final model on the training data
best_rf.fit(X_train, y_train)
# Make predictions
y_pred = best_rf.predict(X_test)
# Calculate accuracy, F1-score, and ROC-AUC
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
roc_auc = roc_auc_score(y_test, y_pred)
# Print best parameters, cross-validation results, and evaluation metrics
print("Best Parameters:")
for param, value in best_params.items():
    print(f"{param}: {value}")
print("Cross-Validation Results:")
cv_results = study.trials_dataframe()
print(cv_results)
print(f"Accuracy: {accuracy:.2f}")
print(f"F1-score: {f1:.2f}")
print(f"ROC-AUC: {roc_auc:.2f}")
PyDev console: starting.
Python 3.11.4 (tags/v3.11.4:d2340ef, Jun 7 2023, 05:45:37) [MSC v.1934 64 bit (AMD64)] on win32
[1 2023-08-17 16:14:39.241] A new study created in memory with name: no-name-478c75d7-eb45-4959-be51-794700609c01
[1 2023-08-17 16:14:42.370] Trial 1 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 98, 'max_depth': 23, 'min_samples_split': 3, 'min_samples_leaf': 5}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:42.793] Trial 0 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 109, 'max_depth': 45, 'min_samples_split': 5, 'min_samples_leaf': 5}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:43.587] Trial 2 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 120, 'max_depth': 14, 'min_samples_split': 7, 'min_samples_leaf': 3}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:43.611] Trial 3 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 121, 'max_depth': 48, 'min_samples_split': 8, 'min_samples_leaf': 5}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:43.615] Trial 5 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 177, 'max_depth': 45, 'min_samples_split': 5, 'min_samples_leaf': 4}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:43.851] Trial 4 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 117, 'max_depth': 30, 'min_samples_split': 2, 'min_samples_leaf': 7}. Best is trial 1 with value: 0.38888888888888884.
[1 2023-08-17 16:14:44.050] Trial 6 finished with value: 0.8888888888888888 and parameters: {'n_estimators': 164, 'max_depth': 44, 'min_samples_split': 4, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:44.233] Trial 7 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 189, 'max_depth': 35, 'min_samples_split': 8, 'min_samples_leaf': 3}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:44.351] Trial 8 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 189, 'max_depth': 32, 'min_samples_split': 4, 'min_samples_leaf': 6}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:44.616] Trial 10 finished with value: 0.7777777777777777 and parameters: {'n_estimators': 70, 'max_depth': 11, 'min_samples_split': 4, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:44.711] Trial 9 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 184, 'max_depth': 10, 'min_samples_split': 7, 'min_samples_leaf': 5}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:45.312] Trial 11 finished with value: 0.7777777777777777 and parameters: {'n_estimators': 161, 'max_depth': 43, 'min_samples_split': 2, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:45.426] Trial 12 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 193, 'max_depth': 47, 'min_samples_split': 9, 'min_samples_leaf': 2}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:45.447] Trial 14 finished with value: 0.38888888888888884 and parameters: {'n_estimators': 99, 'max_depth': 30, 'min_samples_split': 6, 'min_samples_leaf': 4}. Best is trial 6 with value: 0.8888888888888888.
[1 2023-08-17 16:14:45.517] Trial 13 finished with value: 0.8888888888888888 and parameters: {'n_estimators': 78, 'max_depth': 41, 'min_samples_split': 2, 'min_samples_leaf': 1}. Best is trial 6 with value: 0.8888888888888888.
Best Parameters:
n_estimators: 164
max_depth: 44
min_samples_split: 4
min_samples_leaf: 1
Cross-Validation Results:
  number  value  ...  params n_estimators  state
0      0  0.388889  ...      109  COMPLETE
1      1  0.388889  ...      98  COMPLETE
2      2  0.388889  ...     120  COMPLETE
3      3  0.388889  ...     121  COMPLETE
4      4  0.388889  ...     117  COMPLETE
5      5  0.388889  ...     177  COMPLETE
6      6  0.888889  ...     164  COMPLETE
7      7  0.388889  ...     189  COMPLETE
8      8  0.388889  ...     189  COMPLETE
9      9  0.388889  ...     184  COMPLETE
10     10  0.777778  ...      70  COMPLETE
11     11  0.777778  ...     161  COMPLETE
12     12  0.388889  ...     193  COMPLETE
13     13  0.888889  ...      78  COMPLETE
14     14  0.388889  ...      99  COMPLETE
[15 rows x 10 columns]

```

C. Literature Table

N°	Title	Author	Year	Place/Locali zation	Scale	Object	Factors	Tools/Analysis	Results	Others
1	Approaching (Almost) Any Machine Learning Problem	Thakur Abhishek, Aakash Nain, Aditya Soni, Andreas Müller, Andrey Lukyanenko, Ayon Roy, Bojan Tunguz, Gilberto Titericz Jr., Konrad Banachewicz, Luca Massaron, Nabajeet Barman, Parul Pandey, Ram Ramrakhya, Sanyam Bhutani, Sudalai Rajkumar, Tanishq Abraham, Walter Reade, Yuval Reina	2020	India	-	Book about machine learning coding with explanation and resolution about machine learning problems	-	-	-	-
2	The Climate of London	L. Howard	1833	London	Urban	Provide a scientific analysis of the climate of London, including its meteorological and environmental conditions.	Temperature, Humidity, Barometric pressure, Rainfall, wind direction and speed, Cloud cover, Atmospheric phenomena	Thermometers, Hygrometer, Barometer, Rain gauge, Wind vane, Anemometer, Cloud observation instruments	Temperature: the mean temperature of London varied from 8°C (46°F) in January to 18°C (64°F) in July, with the warmest and coldest months being July and January, respectively. The relative humidity of the air in London was highest in the early morning and lowest in the mid-afternoon. The barometric pressure in London tended to be higher in the winter and lower in the summer. the amount of rainfall in London varied throughout the year, with the heaviest rainfall occurring in the months of October and November. the prevailing winds in London were from the southwest, and that the wind speed tended to be highest in the winter months. he amount of cloud cover in London varied throughout the year, with the cloudiest months being December and January.	First scientific studies on climate

3	The energetic basis of the urban heat island	T. R. Oke	1980	Vancouver	Urban	The article aims to provide a detailed understanding of the physical processes responsible for the UHI effect, with a focus on the energy balance in urban environments.	Surface cover, Building materials, Anthropogenic heat, Weather conditions, Urban form	Metereological measurements, Remote sensing (Satellite imagery, aerial photography), Energy balanced modeling	The study showed that urban areas have higher temperatures than surrounding rural areas, particularly at night, due to the absorption and storage of heat energy by buildings and other impervious surfaces, and the release of anthropogenic heat. The research also highlighted the importance of the energy balance in urban environments and the need for urban planners and policymakers to consider the impact of urban design and development on the UHI effect.	Urban energy balance
4	The distinction between canopy and boundary-layer urban heat islands	T.R. Oke	2010	Vancouver	Urban	Understand the underlying causes of the urban heat island phenomenon and how it is influenced by various factors such as land use, vegetation cover, and building density.	Land use, vegetation cover, building density, wind speed, city size	Oke is not a research paper reporting on new findings or experiments, but rather a review article that synthesizes and discusses previous research on urban heat islands.	Oke argues that understanding the difference between canopy and boundary-layer urban heat islands is important for urban planners and designers, as different strategies may be needed to mitigate their effects.	Urban heat island
5	Research on Urban Heat-island Effect	L. Yanga, F. Qiana, D.X. Songa, K.J. Zhenga	2016	Bozhou	Urban	The study seeks to understand the UHI effect by examining the factors that contribute to it, such as surface cover, urban form, and anthropogenic heat, as well as the influence of weather conditions.	Surface cover, Urban form, Anthropogenic heat, Weather conditions	The study relied on a comprehensive review of the literature to examine the factors that contribute to the UHI effect, the impacts of the UHI on urban environments and human health, and various mitigation strategies.	<p>The UHI effect is a complex phenomenon influenced by various factors, such as surface cover, urban form, anthropogenic heat, and weather conditions.</p> <p>The UHI effect can have significant impacts on urban environments, including increased energy consumption, reduced air and water quality, and compromised human health.</p> <p>Various mitigation strategies, such as green infrastructure, building design, and urban planning, can help reduce the UHI effect and its negative impacts.</p> <p>The effectiveness of these mitigation strategies depends on the specific urban context, and a combination of approaches may be required to achieve optimal results.</p>	Urban heat island analysis

6	Smart Solutions for Sustainable Cities—The Re-Coding Experience for Harnessing the Potential of Urban Rooftops	V.Todeschi, G. Mutani, L.Baima, M.Nigra, M.Robiglio	2020	Turin	Urban	Smart Rooftops	Roof suitability (architectural characteristics), morphological context, building codes, and regulations.	GIS; NDVI; LST	A methodology was presented to evaluate rooftop renovation opportunities, assessing energy savings, thermal comfort, and environmental and economic benefits.	Green/High reflectance
7	Mapping the spatial distribution of nocturnal urban heat island based on Local Climate Zone framework	Y.Zheng, C.Ren, Y. Shi, S. H.L.Yim, D. Y.F. Lai, Y.Xu, C.Fang, W.Li	2023	Hong Kong	Local	Nocturnal street scale Urban Heat Island [UHI] spatial distribution	Climate data & Urban morphology data (total street lenght, pervious surface fraction, sky view factor)	GIS, GPS, Landsat 8 images; Multiple Linear Regression (MLR) and Partial Least Square Regression (PLSR)	Identification of UHI hotspots through the statistical models, with the statment that urban forms have significant influences on UHI developement.	Urban morphology
8	The effect of spatial heterogeneity in urban morphology on surface urban heat islands	W. Liao, T. Hong, Y. Heo	2021	Greater London, Seoul	Urban	Urban Heat Island [UHI]: the role of the urban morphology	Spatial heterogeneity	GIS, MODIS products; LST; Correlation analysis (Spearman, Genizi, partial correlation method)	Heterogeneity measures of the green space ratio, building coverage ratio, and canyon H/W ratio had a significant impact on the LST.	Urban morphology
9	Urban heat island data by local weather types in Lisbon metropolitan area based on Copernicus climate variables dataset for European cities	C.Reis, A.Lopes, A. Santos Nouri	2021	Lisbon Metropolitan Area	Regional	Local weather types data based on copernicus Land Monitoring Service climate variables dataset	Thermal patterns (seasons); hourly air temperature, specific humidity, relative humidity and wind speed data	GIS; GeoTIFF	UHI Intensity estimated in a R script using Local Weather patterns	Copernicus Datasets, Remote Sensing

10	Wind environment assessment and planning of urban natural ventilation corridors using GIS: Shenzhen as a case study	X.Liu, B.Huang, R.Li, J.Zhang, Q.Gou, T.Zhou, Z.Huang	2018	Shenzen	Urban	ventilation corridor planning to mitigate UHI	1km2 grid cells climate data; land use (vegetation and water body), building characteristics (height, density), road data (density, connectivity, open area)	GIS, DEM, Landsat 8 images	Wind environment assessment scheme and ventilation corridor construction plan in GIS	Urban morphology/Wind
11	Identifying urban ventilation corridors through quantitative analysis of ventilation potential and wind characteristics	W. Wang, D. Wang, H. Chen, B. Wang, X. Chen	2022	Hangzhou (Shangcheng District, Gongshu District, Xihu District, Binjiang District, Xiaoshan District, Yuhang District, Linping District, Qiantang District.	Urban	ventilation corridor planning to mitigate UHI	Land surface temperature, weather data	GIS, Landsat 8 images	Seven primary and four secondary ventilation corridors were identified. Primary ventilation corridors are primarily water bodies and green land areas; secondary ventilation corridors are mainly roads.	Urban morphology/Ventilation corridors/Wind
12	Integration of topological aspect of city terrains to predict the spatial distribution of urban heat island using GIS and ANN	V. Equere, P.A. Mirzaei, S. Riffat, Y. Wang	2021	Illinois, Chicago	Urban	UHI prediction by integrating the impact of land surface elevation using a novel parameter of terrain factor (TF), which is integrated with other morphological parameters to develop an artificial neural network (ANN) model to predict the spatial distribution of UHI indicated by land surface temperature (LST).	Terrain Factor (TF), Artificial neural network (ANN), LST	ArcGis 10.4, Lidar data, Landsat images, Google Earth, MatLab	Terrain Factor has significant impact on the surface UHI in areas with a significant surface relief variation	Terrain factor/Topography variations
13	Tool for Heat Island Simulation: A GIS extension model to calculate urban heat island intensity based on urban geometry	C. Mayumi Nakata-Osaki, L.C. Lucas Souza, D. Souto Rodrigues	2017	São José do Rio Preto and Bauru	Urban	Max intensity of UHI based on geometry data (H/W)	Street axes, perimeters of buildings, heights of buildings, distance radius of building-axis, roughness	GIS, MatLab	The final result is a computational tool called THIS (Tool for Heat Island simulation), which can be used to simulate hypothetical urban scenarios and verify the potential of the influence on the urban geometry in the UHI _{max} .	Theoretical-numerical basis (Oke's model); Algorithm; GIS extension

14	An urban energy balance-guided machine learning approach for synthetic nocturnal surface Urban Heat Island prediction: A heatwave event in Naples	A. Oliveira, A. Lopes, S. Niza, A. Soares	2021	Naples functional urban area (FUA)	Urban	Predicting nocturnal (SUHI) during night, use of random forest approach to predict LST.	Local climate zones, Land cover, Air temperature, LST, short wave and long wave radiation, albedo, DEM	R (with random forest package), QGIS, Landsat-8, MSG-SEVIRI, Random Forest Algorithm, NASA Atmospheric calculator, GRASS-GIS	Develop and test an energy balance-based machine learning approach, to describe the daily cycle of the heat flux components and predict the nocturnal LST and SUHI, during an HW event.	Nocturnal LST and SUHI
15	Geospatial Assessment and Modeling of Outdoor Thermal Comfort at Urban Scale	G. Mutani, S. Beltramino	2022	Turin	Urban	Outdoor thermal condition at urban scale	Urban Morphology, vegetation, outdoor thermal comfort, sky view factor, mean radiant temperature, urban surfaces	QGIS [UMEP-SOLWEIG (QGIS)], DSM	SOLWEIG is a more suitable tool for assessment and analyses at the urban scale, while ENVI-met is more useful for feasibility studies with high spatial and temporal resolution or for the pre-design phase of little neighborhoods	Outdoor spaces after COVID-19
16	Urban Heat Island Mitigation: A GIS-based Model for Hiroshima	G. Mutani, V. Todeschi, K. Matsuo	2019	Hiroshima	Urban	GIS tool, that can be used to analyze the microclimate of outdoor spaces, considering the relationship between the air temperature and the characteristics of an urban environment.	Urban Morphology, vegetation, density population, surfaces types	GIS, Satellite images (Landsat 7,8)	UHI effect decreases proportionally with the presence of vegetation and with higher values of the albedo of urban surfaces, as well as of the altitude and the distance from the sea. The UHI effect instead increases proportionally for higher values of the canyon height-to-width ratio, the building density and the Land Surface Temperature	Air temperature
17	ArcUHI: A GIS add-in for automated modelling of the Urban Heat Island effect through machine learning	D. J. Espino, C. Manchado, A. R. Valcarce, V. Moscardò	2022	Metropolitan area of Madrid	Urban	Modelling SHUI through ArcGis plug-in, predict observed LST with high accuracy, using Random Forest Regression (RFR).	Building height, albedo, DTM, Land Cover map, LST	ArcGis PRO, R, Landsat-8, Arcgisbinding package	Developing of an add-in (ArcUHI) for automated modelling the UHI effect.	Strategic roof and wall greening was suggested as a measure to mitigate the street canyon effect entailed by buildings and offset the heat retention capacity of built-up surfaces.
18	Measuring the relationship between morphological spatial pattern of green space and urban heat island using machine learning methods	J. Lin, S. Qiu, X. Tan, Y. Zhuang	2022	Shenzhen	Urban	Investigate the relationship between the morphological spatial pattern of green space and UHI intensity using machine learning methods.	Urban morphology, Land cover, morphological characteristics of green spaces, Population, Buildings	Landsat-8, GIS software	The UHI intensity was negatively correlated with the cores, perforations, and loops of green space, but positively correlated with islets.	Suggestions on UHI mitigation and land use planning, especially when the size of green space cannot be unlimitedly increased.

19	Machine learning algorithm based prediction of land use land cover and land surface temperature changes to characterize the surface urban heat island phenomena over Ahmedabad city, India	P. Mohammad, A. Goswami, S. Chauhan, S. Nayak	2022	Ahmedabad (India)	Urban	the prediction of LULC, seasonal LST, and urban thermal field variance (UTFVI) over Ahmedabad city, India using multi date Landsat data.	LULC maps, Bareness Index (BI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), Normalized Difference Moisture Index (NDMI), Modified Normalized Difference Water Index (MNDWI), Latitude and Longitude	Landsat-8 OLI/TIRS, Landsat-5 TM, Google Earth Engine, CART, QGIS	The result demonstrated that more than 70% area in summer and 40% in winter would likely face higher temperature zones than their respective present situation	MOLUSCE plugin in QGIS
20	Simulating and mitigating extreme urban heat island effects in a factory area based on machine learning	S. Liu, J. Zhang, J. Li, Y. Li, J. Zhang, X. Wu	2021	Wu'an	Urban	To mitigate the E-UHI, machine learning is used for simulating and quantifying the marginal utility of the scale, shape, type, stage, and structure of the factory on the land surface temperature (LST), factory LST (LSTf), surrounding LST (LSTs) and increase value (Δ LST) level.	Industrial shape, type, production stage, scale, and internal structure.	Landsat-8, WorldView-3, ALOS, FLUKE F572-2 laser, HI-TARGET Qpad X3 (GPS), ENVI, MATLAB	The results show that the scale of all types of factories affects LSTf and LSTs, and the shape of steel factories affects LSTs and Δ LST.	Industrial E-UHI
21	The Effects of Green Roofs on Outdoor Thermal Comfort, Urban Heat Island Mitigation and Energy Savings	G. Mutani, V. Todeschi	2020	Turin	Urban	The main aim of this work has been to analyze how the presence of green surfaces integrated with the building envelopes, together with the urban morphology, can mitigate the UHI, improve indoor and outdoor thermal comfort, and help save energy	Land surface temperature distribution, geometrical and typological characteristics of the city, empty spaces, characteristics of buildings, critical areas, Land cover, NDVI, orientation of the street, H/W ratio of urban canyons, sky view factor, DSM, albedo, vegetation	Local weather stations, Landsat-8, ArcGIS, Feature Analyst 5.2 , ortophoto, technical city's map of Turin	Land-surface temperature, and therefore the air temperature, tend to decrease as the green areas increase.	Feature Analyst 5.2

22	Unpacking the inter- and intra-urban differences of the association between health and exposure to heat and air quality in Australia using global and local machine learning models	S. Wang, W. Cai, Y. Tao, Q. C. Sun, P. P. Y. W., X. Huang, Y. Liu	2023	Australia	National	Quantify and identify the potential confounders that are important to the heat-air-health relationship, unveil the cumulative heat and air pollution on self-reported physical and mental health, reveal the spatial heterogeneity of such effects across urban-rural and intra urban spaces	Census, Health, Land use, Point of interest, Google earth engine, LST	GIS	The results show that social and built environmental factors are more influential to physical and mental health outcomes than heat and air pollution, especially in rural areas.	Health and LST
23	The future of China’s urban heat island effects: A machine learning based scenario analysis on climatic-socioeconomic policies	T. Lan, J. Peng, Y. Liu, Y. Zhao, J. Dong, S. Jiang, X. Cheng, J. Corcoran	2023	Changsha-Zhuzhou-Xiangtan	Urban	Spatially estimate UHI effects from 2040 to 2100.	LST, Land Cover map, population, DEM	GIS	<p>The risk of UHI effects was assessed through UHI hazard and population exposure. Results showed that compared with 2020, heat islands would decrease by 30% - 34% in area but increase by 29% - 37% in mean intensity by 2100. The population impacted by UHIs would increase by >10 million under the third scenario.</p> <p>Areas at high risk of UHI effects are mainly located in the urban core, with the total value of risk across the study area by 2100 being about 2 - 4 times higher than that in 2020.</p>	UHI vs Health
24	Satellite and Ground-Based Sensors for the Urban Heat Island Analysis in the City of Rome	R. Fabrizi, S. Bonafoni, R. Biondi	2010	Rome	Urban	The research aims to analyze the Urban Heat Island (UHI) phenomenon in Rome using both satellite and ground-based sensors, focusing on identifying the areas that are most affected by high temperatures.	Urban morphology, temperature	Satellite data and ground-based sensors	The study found that the phenomenon is more pronounced in densely urbanized areas, with a temperature difference of up to 8°C between the urban center and the surrounding rural areas.	Remote sensing

25	Racial and Socioeconomic Disparities in Heat-Related Health Effects and Their Mechanisms: a Review	C. J. Gronlund	2014	Michigan	-	Examine the ways in which race and socioeconomic status contribute to disparities in heat-related health outcomes. The study focuses on the mechanisms that underlie these disparities, including differences in exposure to heat, pre-existing health conditions, and access to health care. The objective is to identify strategies for reducing these disparities and improving overall health outcomes in vulnerable populations.	access to air conditioning, outdoor occupation, social isolation, and pre-existing medical conditions, urban design, environmental factors	It was a literature review, the study did not involve any tools or experiments, but rather analyzed and synthesized data from various sources.	The study found that people of color and those from low-income backgrounds were more likely to experience negative health impacts due to heat waves. This was due to factors such as lack of access to air conditioning, outdoor occupations, social isolation, and pre-existing medical conditions.	Health disparities
26	Role of local climate zones and urban ventilation in canopy urban heat island–heatwave interaction in Nanjing megacity, China	W. Tian, Y. Yang, L. Wang, L. Zong, Y. Zhang, D. Liu	2023	Nanjing	Urban	Spatio-temporal variations of the CUHI intensity (canopy urban heat island) and its association with HWs	Metereological stations	GIS	Significant diurnal variation in CUHII, showing a valley at daytime and a peak at nighttime. The CUHII is significantly stronger in the HW periods than in the non-heatwave	Canopy
27	Urban Heat Island: Causes, Consequences, and Mitigation Measures with Emphasis on Reflective and Permeable Pavements	Vujovic, S., Haddad, B., Karaky, H., Sebaibi, N., & Boutouil, M.	2020	More Places has been taken into consideration	Urban	The objective of this research is to investigate the economic and social development in urban and rural areas of developing countries in the context of increasing impervious surfaces, primarily paved surfaces, and its consequences on urban heat island formation, while also exploring measures like reflective and permeable pavements for mitigation and urban microclimate improvement.	Green Areas, Building materials, Morphology, weather condition	-	The results indicate that the literature extensively documents the UHI phenomenon, its causes, and consequences, including its effects on human health, thermal comfort, increased cooling energy consumption, air pollution, and surface water quality deterioration; furthermore, the study highlights the importance of landscape design in urban planning and the potential of cool pavements, specifically highly reflective and permeable pavements, for improving the urban microclimate by reducing pavement surface temperatures and their impact on air temperatures, albeit with considerations such as site-dependent environmental impacts and the need for periodic watering and pavement structure optimization; finally, the study suggests the need for further research through experimental models or numerical simulations to assess the efficiency of combined strategies involving reflective and permeable pavements as potential UHI mitigation measures.	Building materials

28	The impact of impervious surface development on land surface temperature in a subtropical city: Xiamen, China.	Xu, H.; Lin, D.; Tang, F.	2013	Xiamen	Regional	The research focuses on estimating impervious surface area in different regions of China using Landsat 8 imagery, considering factors like climate, soil, vegetation, and image acquisition timing, with the aim of improving accuracy and providing recommendations for optimal seasonal image use and the integration of various data sources and algorithms to reduce spectral confusion in impervious surface estimation.	Water, Impervious surface, Permeable surface, Climate, Soil, Vegetation	Landsat 5,7,8 ; ENVI 5.1	This study used Landsat 8 OLI imagery to classify land cover types in China, focusing on impervious surfaces, and found that spectral resolution was more important than spatial resolution for accurate classification. They recommended specific image acquisition times for different regions to achieve the best accuracy, with potential for further improvements by integrating additional data and algorithms in future research.	-
29	Reduced urban heat island intensity under warmer conditions.	Scott, A. A., Waugh, D. W., & Zaitchik, B. F.	2019	For this study has been taken into consideration 54 US cities	National	The objective of this research is to analyze the Urban Heat Island (UHI) phenomenon in 54 US cities over a 15-year period, investigating how UHI intensity relates to temperature variations, weather conditions, and urban warming trend.	UHI, Temperature variation, Climate zones, Weather conditions, Urban Warming Trends	Global Historical Climatology Network (climate stations), Defense 86 Meteorological Satellite Program's Operational Linescan System	The results of the research indicate that in 38 out of 54 US cities studied, the intensity of the Urban Heat Island (UHI) tends to decrease as temperatures increase, and this pattern holds across different climate zones; the decrease in UHI intensity is primarily driven by changes in rural areas and is most pronounced during moist weather conditions; furthermore, the study finds that warming cities have not experienced an increasing UHI effect.	Meteorological stations
30	Impact of Urbanization on Urban Heat Island Intensity in Major Districts of Bangladesh Using Remote Sensing and Geo-Spatial Tools.	Rahman, M. N., Rony, M. R. H., Jannat, F. A., Chandra Pal, S., Islam, M. S., Alam, E., & Islam, A. R. M. T.	2022	Bangladesh	Urban	The research aims to analyze how rapid urbanization has affected the intensity of Urban Heat Islands (UHIs) during the winter dry period in major districts of Bangladesh from 2000 to 2019.	LULC maps, Hotspots and LST, Urbanization, UHI, NDVI	Landsat 5,7,8 ; GIS	The research in seven urbanized districts of Bangladesh over 20 years revealed that urban growth resulted in less vegetation, higher land temperatures, and increasing Urban Heat Island (UHI) intensities, with Mymensingh having the highest UHI intensity (10°C) and Dhaka the lowest (1.46°C), highlighting significant thermal changes and underscoring the importance of sustainable urban planning to address UHI effects and potential regional climate impacts.	Rapid urbanization

31	Local climate zones for urban temperature studies.	Stewart, I. D., & Oke, T. R.	2012	More Places has been taken into consideration	Urban	The research aims to introduce the "local climate zone" (LCZ) classification system, which categorizes urban and rural areas into 17 unique zone types, providing a standardized framework for temperature observations and benefiting urban heat island researchers and related fields.	LCZ classification, UHI, Urban Development, Physical and Climatological Characteristics	GIS	In summary, the local climate zones classification system offers a cost-effective and versatile approach for characterizing urban and rural areas' climate characteristics, aiding in standardized reporting, UHI comparisons, and improving our understanding of urban climates and site selection processes.	Local Climate Zone (LCZ) Classification
32	Vegetation as a climatic component in the design of an urban street: An empirical model for predicting the cooling effect of urban green areas with trees.	Shashua-Bar, L., & Hoffman, M. E.	2009	Not provided	Urban	The research aims to develop an empirical model to predict the cooling effect of urban green areas with trees, specifically focusing on the role of vegetation as a climatic component in the design of urban streets.	The main factors considered in this research are likely related to tree characteristics, urban design, and local climate conditions that influence the cooling effect of trees in urban settings.	Data collection on tree characteristics, urban design elements, and climate conditions.	The research would present an empirical model for estimating the cooling effect of trees in urban green areas, providing valuable insights for urban planners and designers regarding the use of vegetation to mitigate urban heat and improve the local climate on streets.	Urban Tree Cooling Model