

Research article

Estimating the importance of environmental factors influencing the urban heat island for urban areas in Greece. A machine learning approach



Ilias Petrou ^{*}, Pavlos Kassomenos

Laboratory of Meteorology, Department of Physics, University of Ioannina, University Campus, GR-45110, Ioannina, Greece

1. Introduction

Without a doubt, urbanization offers many advantages and as a result, humankind has prospered across the planet. However, such advantages come with unanticipated consequences. Considerable changes in land cover result from urban growth. It is well-established that, the replacement of natural (vegetated) land cover with artificial built-up surfaces (impervious), such as roads, pavements, buildings, and other constructions required to support urban expanding, has a great impact on turbulent heat and moisture exchange between land and atmosphere (Chen et al., 2018), on aerodynamic properties (Grimmond and Oke, 1999), on the absorption of solar radiation and the presence of shade within cities (Dudorova and Belan, 2022), on the process of evapotranspiration (Chen et al., 2023) and on the emission of greenhouse gases and anthropogenic heat (Narumi et al., 2021). These mechanistic pathways, directly or indirectly contribute to the urban environment's excessive warmth in comparison to the rural environment. This effect, is referred to as the urban heat island (UHI), and regarding its dynamics, operating space, and time scales, it is quite complicated. It emerges as a result of the variations in the surface energy balance between rural and urban environments, therefore is interrelated to global warming, rising global threats related to climate change (Akbari et al., 2015). Heat islands can be classified into three basic categories: atmospheric, surface and subsurface, depending on the medium used for sensing. The atmospheric type, is the most commonly studied UHI type exists in the atmosphere. Canopy-layer UHI (CUHI), which is also the subject of the present study, is a type of atmospheric UHI associated with shallower atmospheric layer extending from the ground up to the mean roof level of the city (Oke, 1976). The effects of CUHI are influenced by a variety of factors, including the complicated urban infrastructure, paved surfaces, reduced vegetation, the release of air pollutants in urban areas, the local morphological and surface geometry, the size of the city, the population density, and the local climate/weather phenomena (Li et al., 2020; Lee et al., 2020).

One of the major variables determining the onset and intensity of

CUHI is also the local weather conditions. Many studies have been conducted in order to find associations between large-scale prevalent meteorological conditions in a specific location and the magnitude of the CUHI (Reis et al., 2022; Khan et al., 2021). Additionally, the interactions of certain climatic factors (such as wind speed, temperature, relative humidity, cloud cover, etc.) and CUHI have been studied in the past and are well understood (Zheng et al., 2023). The significance of atmospheric humidity for the urban energy balance is widely acknowledged (Huang et al., 2021). Several previous studies have demonstrated that cities are generally considered to be drier than their surrounding non-urbanized areas, an effect that is known as Urban Dry Island (UDI) (Milelli et al., 2023). For instance, Zhang et al. (2023) examined the urban heat effect and its correlation with humidity and concluded that UDI offsets the CUHI. Similarly, Chakraborty et al. (2022) determined that lower relative humidity due to urbanization partly offsets the effect of CUHI and thus moderates the potential for urban heat stress. On the other hand, previous studies have also reported that cities could be moister than the surrounding rural areas (Kuttler et al., 2017). The terms of "urban moisture island" (UMI) and "urban moisture excess" (UME) refers to the phenomenon wherein an urban region has more moisture than its rural surrounds, especially during nighttime (Wang et al., 2015). In addition, wind conditions have been also shown to have an effect on the development and the intensity of the CUHI (Ngarambe et al., 2021). The high-density urban built environments generally experience weak wind conditions, preventing convective heat exchange and trapping hot air within cities, resulting in strengthened CUHI conditions. In a recent study, Al-Obaidi et al. (2021) revealed that wind speeds less than 2 m/s provide the highest CUHI intensities, whereas wind speeds more than 6 m/s produce the lowest magnitudes, in five large Australian cities. Moreover, previous research demonstrated the interactions between local winds (sea and land breezes) and CUHI (Yang et al., 2022; Bauer, 2020). In general terms, sea breezes could mitigate the adverse UHI effects and cool coastal cities, whereas the land breeze could moderately increase the strength of CUHI effect.

Apart from climatic component and land surface characteristics, the

* Corresponding author.

E-mail address: petroui@hotmail.com (I. Petrou).

formation and growth of the CUHI effect are also significantly influenced by urban morphology. Numerous studies have demonstrated that the spatial structure and character of a city (including their land uses, roadway networks and building layouts) significantly impact urban microclimate (Liu et al., 2023; Zhou and Chen, 2018). For example, Yang et al. (2019) concluded that elevated surface temperatures may be caused by high-density high-rise buildings. On the contrary, in case of large Chinese megacities, was found that land surface temperatures (LST) are greater in areas with low-rise buildings and lower in areas with high-rise ones, due to the effects of building shadows, nearby wind disturbances and building layout (Wang and Xu, 2021). On the other hand, an extensive body of research has been conducted on the associations between impervious surfaces (IPS) and LST, which are the primary sources of the surface UHI (SUHI) (Zhang et al., 2012, 2022; Morabito et al., 2018). IPS refers to artificial structures (such as pavements, roads, parking lots, airports, ports and paved areas) that can prevent the infiltration of surface water into the ground, rising the surface runoff and lowering of groundwater levels. Through changes in surface albedo and specific emissivity, IPS directly impact the vertical radiation balance, increasing also the intensity of the CUHI and surface sensible heat flow. On the whole, the vast majority of studies demonstrate strong correlations between the IPS and the urban thermal environment. This suggests that an increase in the IPS area will result in an increase in the regional LST.

Most of the published literature that examined the contribution of different environmental factors to the CUHI effect, is focused on the trends in different metrics (means, maximums, minimums, etc.) of individual or multiple variables, applying basics and commonly used predictive statistical methods, such as linear regression analysis. However, estimating complicated nonlinear and realistic synergistic roles of multiple environmental factors in CUHI intensity remains largely unexplored. This is particularly important when the dataset of this factors is quite large, with fine temporal and spatial resolution. Over the last few years, machine learning (ML) techniques have proven very useful tool in applied climatological research, highlighting the methodological advances and their applications to different research areas. For instance, Kyros et al. (2023) applied ML techniques in order to determine the optimal correlations between meteorological parameters and cloud formation that result in precipitation and improve short-range forecasting of rainfall. In addition, applying ML techniques for statistical modeling aerosol changes (Li et al., 2022), for explaining heatwaves (Buschow et al., 2024), and for predicting sea surface temperature and marine heatwave occurrence (Bonino et al., 2024), have provided encouraging outcomes. Concerning the UHI study, several researchers have adopted ML techniques in order to investigate the impacts of CUHI in urban areas (Furuya et al., 2023; Addas, 2023; Rao et al., 2023). Moreover, ML-based analysis has adopted in order to study the relationship between urban morphology and CUHI. In this context, Yoo (2018) found that the most significant urban features in the creation of UHIs in the Marion County (Indiana region) were the normalized difference vegetation index (NDVI) and the percentage of built imperviousness. Similarly, Hou et al. (2023) employed ML algorithms and concluded that the most significant environmental driver of CUHI was determined to be the covering of urban vegetation, with surface albedo coming in second.

Up to now, it is still not clear which complex urban metrics associated with UHI in large Greek urban areas. The application of big data in this field is crucial, nevertheless has received rather little attention from the scientific community. In addition, to the authors' knowledge, although machine learning (ML) techniques are widely used to comprehend the degree of dependency of various environmental factors in the UHI effect, they have not been utilized in the investigation of UHI impacts in Greece. To fill this gap, this study aims to explore and quantify the relative significance of multiple environmental factors that may contribute to the UHI intensity (UHII), by applying a commonly-used machine learning algorithm, the Random Forest (RF). This

particular method, utilized meteorological factors (relative humidity, specific humidity and wind speed) as well as urban morphology factors (building heights and impervious surfaces) for Athens and Thessaloniki, the two largest and highest urbanized cities in Greece, during the summer period. Urban planning methods for more efficient improvement of the urban thermal environment can be prioritized by assessing their respective contributions to the intensity of the UHI.

2. Materials and methods

2.1. Study areas

In order to estimate the importance of environmental factors influencing the summer UHI, the present research is focused in two major cities in Greece, namely Athens and Thessaloniki (Fig. 1). Both cities are characterized by high levels of population density, large built-up areas and widespread suburban areas (Fig. 2).

More specifically, the first subject of this study is Athens Urban Area (AUA - also known as the Athens–Piraeus Urban Complex), which includes the whole city of Athens, the nation's capital and largest city of Greece. The total size of AUA is approximately 412 km², consists of 40 municipalities within Attica region and is a highly urbanized and built-up region with few open spaces and natural regions (Fig. 2). With a very high population density and with a population over three million, almost half of the population of the country lives there. The whole Athens–Piraeus complex has been built inside a sizable basin that is located in the peninsula's center. The area's local climate conditions, which allow for year-round outdoor activities, has influenced the architectural style as well as the city's way of life. The CUHI effect is well documented for the AUA, revealing that higher intensities of UHI affected among others by building forms (light-colored flat-roofed and multi-story apartments), building material (cement), insufficient green infrastructure, narrows streets, anthropogenic heat emissions (vehicle exhaust, energy consumption etc.) and atmospheric pollution (Santamouris et al., 2015; Stathopoulou and Cartalis, 2007; van der Schriek et al., 2020).

The second study area, includes Thessaloniki Urban Area (TUA), a built-up area lining the Thermaikos gulf, the second-largest urban center in the country and the capital of the region of Central Macedonia in Northern Greece (Fig. 2). TUA comprise seven municipalities and according to the 2021 census counts a total population of about one million (<https://www.statistics.gr/en/home>). The urban morphology of TUA is characterized by cemented and asphalted structures, without parks with natural terrain or materials with high reflexivity. The traffic in the city is huge and the pollution caused by cars is noticeable. In addition, the building density is high resulting in significant obstacles for the wind passes. Hence, these features affecting the local microclimate as temperature and relative humidity. Although the study of CUHI effect in TUA has received rather little attention, it has been reported to be affected by local weather conditions, urban density, building height and urban land uses (Keppas et al., 2021; Giannaros and Melas, 2012).

2.2. Copernicus Urban Climate dataset

In order to achieve a detailed quantification of the CUHI intensity, as well as to evaluate the importance of meteorological factors to UHII in both urban areas of Athens and Thessaloniki, the Copernicus Urban Climate dataset was used presented by the Copernicus Climate Change Service (Hooyberghs et al., 2019). This dataset contains hourly gridded values (air temperature, specific humidity, relative humidity, and wind speed) for 100 European cities from 2008 to 2017, with a spatial resolution of 100 m at the scale of a city neighborhood). The data were produced using the UrbClim urban climate model, which has undergone extensive verification by comparing hourly temperature measurements from many European cities with the model's output (De Ridder et al., 2015; Kourtidis et al., 2015; Lauwaet et al., 2016; Verdonck et al., 2018;

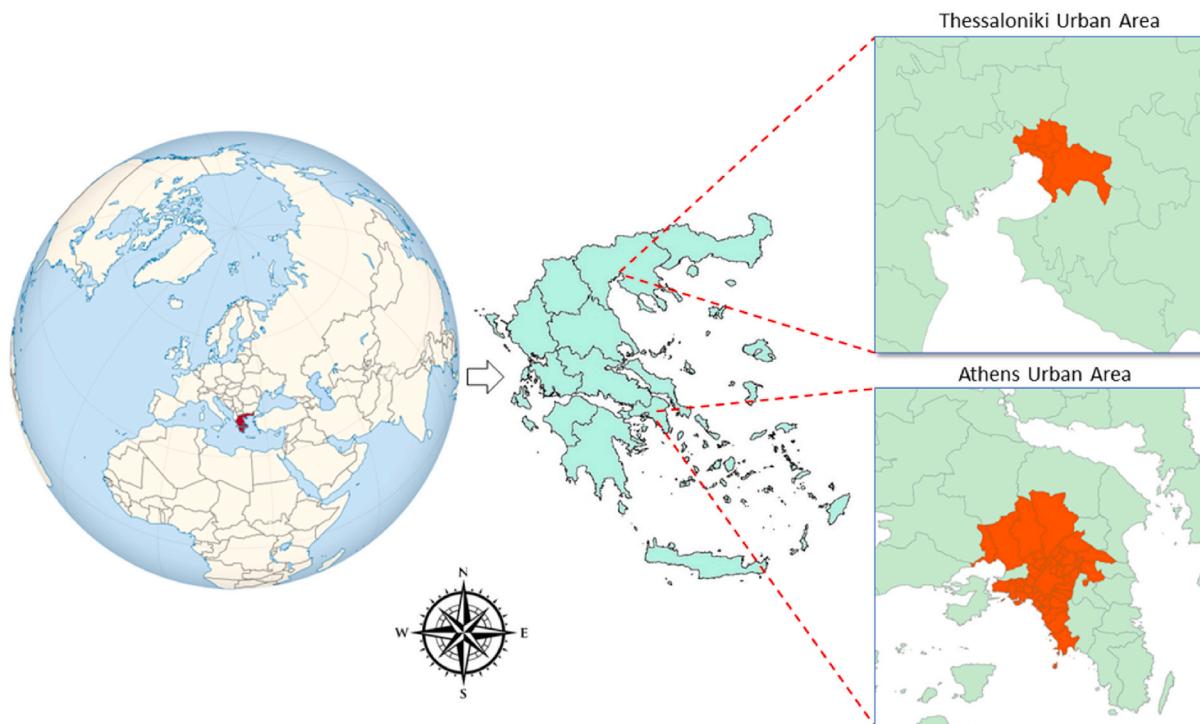


Fig. 1. Map of study areas including Thessaloniki Urban Area and Athens Urban Area.

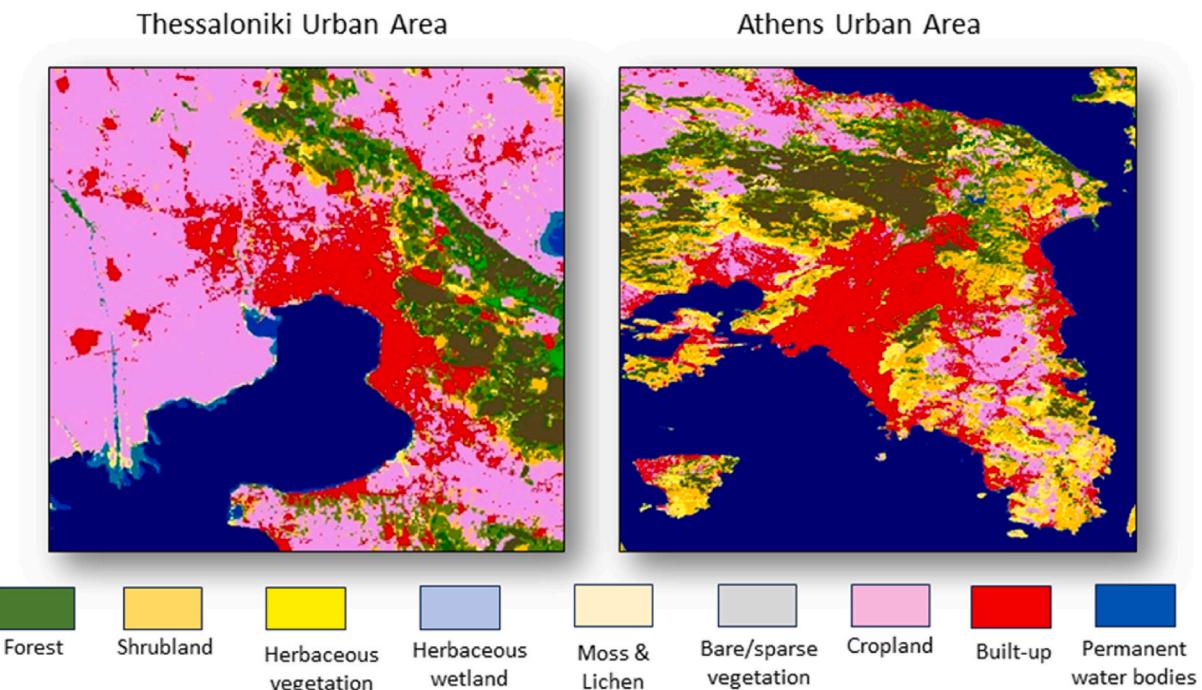


Fig. 2. The Land Cover composition for each study area, according to the discrete classes of the CGLS Dynamic Land Cover Map based on Copernicus Global Land Operations (Buchhorn et al., 2020).

Zhou et al., 2016). For this study, air temperature (K), relative humidity (%), specific humidity (kg kg^{-1}) and wind speed (m/s) hourly data at the height of 2m above the surface, in NetCDF-4 format were collected for Athens and Thessaloniki, between 2012 and 2017. Only the months of June, July, August, and September were chosen to depict the summer season.

It is well established that the intensity of the UHI can vary

significantly at hourly time scales. Much of the published literature on urban climate has highlighted that during nighttime the UHI effects become more evident in both urban areas. Due to the higher thermal inertia of the materials employed in the urban fabric, nocturnal UHI is more prominent. For instance, in TUA, high UHII values were recorded in the hours after sunset, continued through the night, and then sharply decreased the next morning (Giannaros and Melas, 2012). Kourtidis

et al. (2015), also reported that the Athen's center and coastline are the hottest areas in the nighttime under different meteorological conditions. On the other hand, during the daytime a reversed UHI (lower temperatures in the urban areas than in the rural ones) has been reported from some UHI studies (Kantzoura et al., 2012; Kassomenos and Katsoulis, 2006). In this context, understanding the dynamics of both daytime and nighttime UHIs is essential for developing effective policies and interventions to enhance urban resilience to climate change. Therefore, the above meteorological data were then separated into daytime (06:00–18:00 LT) and nighttime (23:00–05:00 LT) time segments.

Last but not least, urban and rural areas defined based on two additional variables that are provided to the dataset: (i) a land-sea mask and (ii) a rural-urban mask, derived from the CORINE land cover of 2012 (<https://land.copernicus.eu/pan-european/corine-land-cover>). The rural-urban mask utilized in order to calculate the UHI by presenting a value of 1 for rural surfaces and a NaN for urban surfaces, while land-sea mask was used to divide the land from sea areas, in each urban area.

2.3. Urban morphology data

In addition to the meteorological variables described above, urban morphology data were utilized, to explore the importance of urban parameters influencing the summer UHI in each urban area. Specifically, building height (BH) information for both urban areas, in a 10 m raster layer for the 2012 reference year were obtained from Copernicus Urban Atlas - European environment et al., 2012 dataset (<https://land.copernicus.eu/local/urban-atlas/building-height-2012>) (accessed on December 22, 2023)). The urban atlas from the Copernicus project specifies building heights for some major European cities. Aside from building heights datasets, also high-resolution layer imperviousness density (IMD) 2012 for both urban areas were downloaded from the Copernicus land monitoring service (CLMS) web site: <https://land.copernicus.eu/en/products/high-resolution-layer-imperviousness-imperviousness-density-2012> (accessed on December 23, 2023). The IMD is a 10 m raster dataset representing the degree of imperviousness within each pixel, where each pixel is coded in the range 0–100. It's worth mentioning that the coordinate reference system for both datasets is ETRS89-LAEA89 (EPSG:3035).

2.4. Calculation of the urban heat island intensity

As mentioned above, the present study focuses on heat island that occur in the urban canopy layer (UCL) in which processes of airflow and energy exchanges are primarily controlled by micro-scale and site-specific characteristics. Complementary observations of near-surface air temperature at urban and rural locations are typically used to research the CUHI impact. The difference in air temperature between an urban and the surrounding suburban/rural area(s), is considered to be the most widely known measure of UHII or UHI magnitude. The majority of heat island studies use mainly three methods to measure the UHII, including calculation of the difference in temperature between the warmest and coolest sites of a network of weather stations (Chandler, 1961), the calculation of the difference in spatial averages of temperature between urban and rural sites (Kim and Baik, 2005) and last, the calculating of the difference in temperature between an outlying rural site and a site close to the geographic center of a city (Runnalls and Oke, 2000). For this study, in both urban areas the summer UHII was calculated for every hour of the daytime and nighttime from 2012 to 2017 according to the following equation:

$$\Delta T_i = T_{\text{urban}} - T_{\text{rural(mean)}} \quad (1)$$

where ΔT_i represents the urban heat island intensity, T_{urban} is the 2 m temperature for each urban grid point, and $T_{\text{rural(mean)}}$ is the 2 m temperature averaged over the rural grid points. Additionally, in both cases

the rural classifications of CORINE (coordination of information on the environment)—which include woodland, grassland, farmland, shrubland, broadleaf forest, and needleleaf forest—represent the rural region of each area (<https://land.copernicus.eu/en/products/corine-land-cover> (accessed on March 20, 2024)). To be more specific, the land area is given a value of 1 by the land-sea mask, whereas the sea mask gives the sea area a missing value of NaN. On the other hand, the rural–urban mask assigns a value of 1 to the rural classes of CORINE and a missing value of NaN to the urban classes of CORINE (Fig. S1 and Fig. S2).

2.5. Random forest algorithm

The relative importance and impact of several environmental factors on UHII was quantified, using the Random Forest (RF) algorithm. The RF algorithm, is a widely-used machine learning approach that is very flexible, ease of use and is capable to provide high prediction accuracy and model high dimensional complex data. This kind of supervised learning algorithm is also applicable to tasks involving both regression and classification (Svetnik et al., 2003). In order to produce a prediction that is more reliable and accurate, it combines the predictions from several decision trees, usually trained with the bagging method. Each decision tree has a significant variance, but when combining them all at once, the variance that results is reduced since each decision tree is precisely trained on a specific sample data. As a result, the outcome is dependent on numerous decision trees rather than just one. When dealing with a classification issue, the majority voting classifier is used to determine the final result, while the mean of all the outputs is the final result in a regression problem (Aggregation). Moreover, along with the prediction or classification of a target variable (also known as the label) based on explanatory variables (features), RF algorithm is also capable to score the relative importance of each feature (Menze et al., 2009).

Therefore, in this study building height (BH), imperviousness density (IMD), relative humidity (RH), specific humidity (SH) and wind speed (WS) were adopted, as local environmental features, in order to estimate their importance to summer UHII (target) in each urban area. Both meteorological and urban morphology features represent the basic determinants of the canopy-layer UHII, that play a key role in shaping the absolute temperature difference between urban and rural areas. At first, for both urban areas, the mean daytime and nighttime values of meteorological data as well as the mean daytime and nighttime values of UHII (ΔT_i) for each urban grid point (coordinate) and for each month (June, July, August and September) were calculated between 2012 and 2017. After that, the urban morphology grid points match with those of metrological grid points above and subsequently a matrix is constructed for each month and for both day and night. Such a matrix contains either the mean daytime or the mean nighttime variations of all environmental factors at the corresponding grid points. The data of the matrixes separated into features and targets, in such a way that the six columns of each matrix correspond to five features (BH, IMD, RH, SH and WS) and the one target (UHII), while the rows correspond to space (coordinates). To put our data into machine-understandable terms, data manipulation relying on the structure known as a dataframe, using Python library Pandas and then converted the Pandas dataframes to Numpy arrays.

With the aim of achieving a detailed quantification of the UHI intensity not only during the summer season, but also conduct a monthly scale analysis of the importance of different factors on UHI, considering that throughout the summer months the climatic and other characteristics (such as population, tourism) of the two urban areas exhibit notable variations, two experiments were conducted based on RF algorithm. In the first one, the average of all summer months was calculated for each urban area, while the second one was established based on each month of the summer season, for both day and night. In all experiments, the data were split into training (75% of the data) and testing (25% of the data) sets, to train the RF model and evaluate the model's performance respectively. Additionally, applying the most common cross validation (CV) technique, the K-Fold CV, the training set further split

into five number of subsets (folds) of equal sizes, to fine-tune the hyperparameters of RF model (Pal and Patel, 2020). According to this method, the four folds are used as the training dataset (based on a set of hyperparameters), while the remaining set is used for testing. The algorithm is trained and tested five times, using the leftover sets for training and a new set for testing each time. Tables S1-S4 show the list of the hyperparameters used for training the algorithm in both experiments, for both urban areas. Then, to automatically find the best combination of hyper-parameters, for the model and test dataset, the grid search method (Gridsearch CV) was used, an approach that assesses every combination that specified rather than selecting a random sample from a distribution. The model is retrained using the optimal set of hyper-parameters to the complete training set, and its performance is assessed by calculating the accuracy (comparing the predicted values to the actual values) of the model in each set of experiment. Finally, the importance of each feature was calculated, by examining the degree to which all of the forest's trees' nodes that make use of that feature minimize impurity. After training, it automatically calculates each feature's score and adjusts the output so that the total relevance of all the features equals one. This efficient method called Mean Decrease Impurity (MDI), and has been widely used in a variety of applications (Han et al., 2016).

Fig. 3 depicts the schematic overview of the procedure in this study, describing the input training dataset to the RF model for both urban areas, as well as the diagram of the RF algorithm.

3. Results and discussion

For this study, the importance of various environmental factors to the UHII effect was investigated during the summer period from 2012 to 2017, for both day and night and for the largest urban areas in Greece, including Thessaloniki and Athens urban area. In order to determine the most important variables influencing UHII, RF model was utilized, training and testing two experiments, including the monthly mean of all data and mean daytime and nighttime values of each summer month respectively, for both urban areas. In each experiment, the performance of the model was evaluated, based on the optimal combination of hyper-

parameters.

3.1. Spatial and temporal patterns of urban heat island during summer

For the period between 2008 and 2017, Figs. 4 and 5 demonstrate the spatial characteristics of mean UHII values during day and night per summer month, in AUA and TUA respectively. It is evident that for both urban areas, the intensity of UHI (mean values of UHII) is higher during the night than during the day, in agreement with previous studies (such as Giannaros et al., 2013; Giannaros and Melas, 2012), reporting that the effects of the UHI are mostly nocturnal in Athens and Thessaloniki.

In case of AUA, a closer examination in Fig. 4 reveals that during the daytime, the UHII magnitudes varied between 0 and 2 °C on average, especially in early summer. The daytime UHII effects are mainly located in Athens metropolitan area, including southern urban areas such as Piraeus and Glyfada, and northwestern suburban areas such as Aspropyrgos. This agree well with earlier studies, in which the specific urban areas have been characterized as UHI hotspots (Petrou et al., 2023; Keramitsoglou et al., 2011). During the night, at the same areas the intensity of nocturnal UHI is greater than 4 °C and lower than 6 °C, reaching their maximum values in early summer. The nighttime UHI hotspots are also more spatially extensive in June and July, while they are partially limited towards the end of the summer. Looking at Fig. 4, one can also notice that in northeastern suburban areas surrounding Mount Penteli, the UHII values are considerably lower (from -1 °C to 1 °C). Remarkably, these values are also calculated during the night, highlighting the importance of keeping the green areas around the AUA, as an effective mitigation strategy of the UHI effect. At this point it is also worth clarifying, that negative values indicate that these specific urban areas are cooler than the surrounding non-urban areas.

Spatial variations for TUA per month during the warm period is illustrated in Fig. 5. It can be seen that the maximum UHII values are observed during the night, corroborating the results of previous studies that show that UHI is primarily a nocturnal phenomenon in Thessaloniki (Skoufali and Battisti, 2019; Giannaros and Melas, 2012). During the day, the UHI magnitudes varies between 0 °C and 2 °C, while the UHI phenomenon is more clearly reflected in June and September. As it

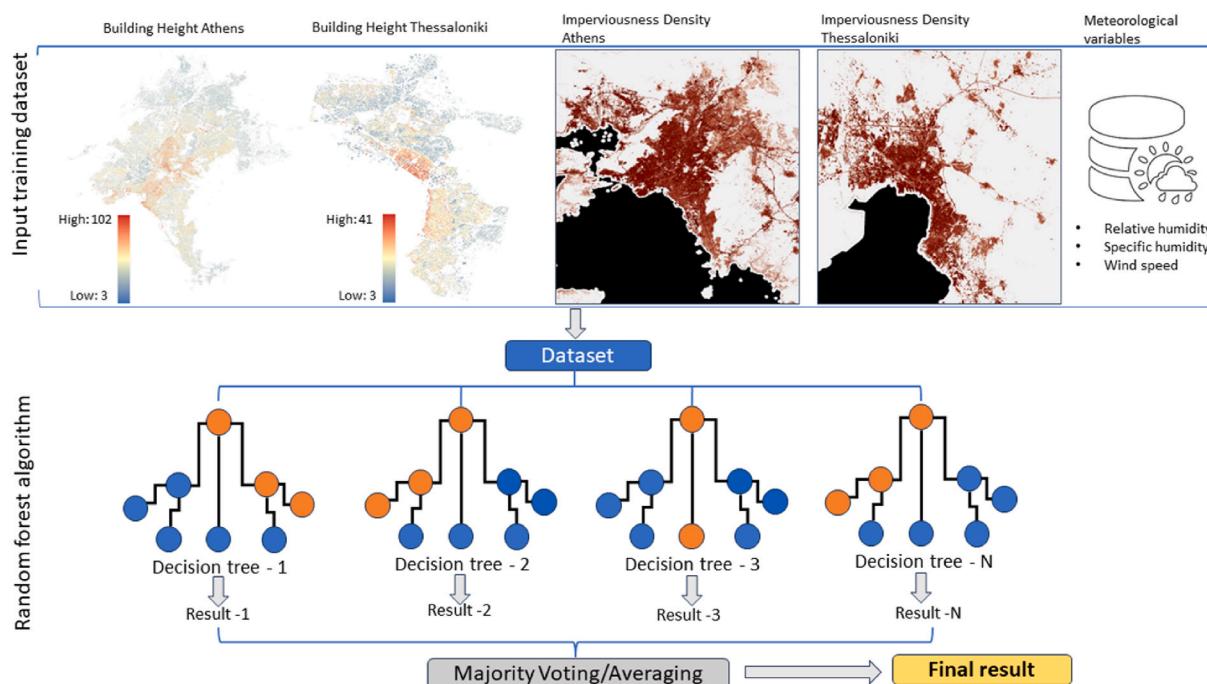


Fig. 3. The procedure of the proposed approach for quantifying the variable importance, including the input training dataset (building height, imperviousness density, relative humidity, specific humidity, and wind speed) for both urban areas and a schematic diagram of the random forest algorithm.

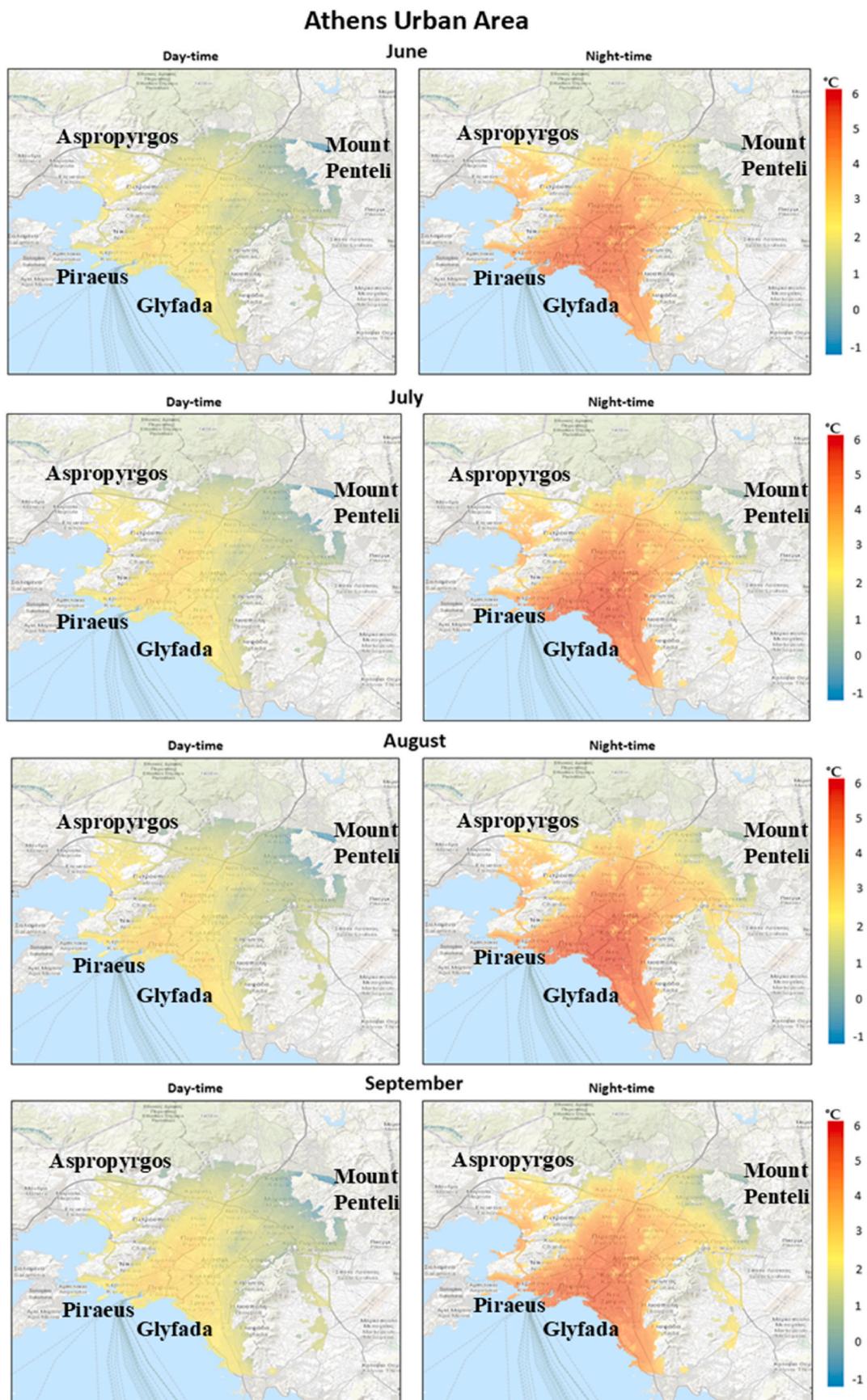


Fig. 4. Spatial variation of the urban heat island intensity per summer month during day and night in Athens urban area.

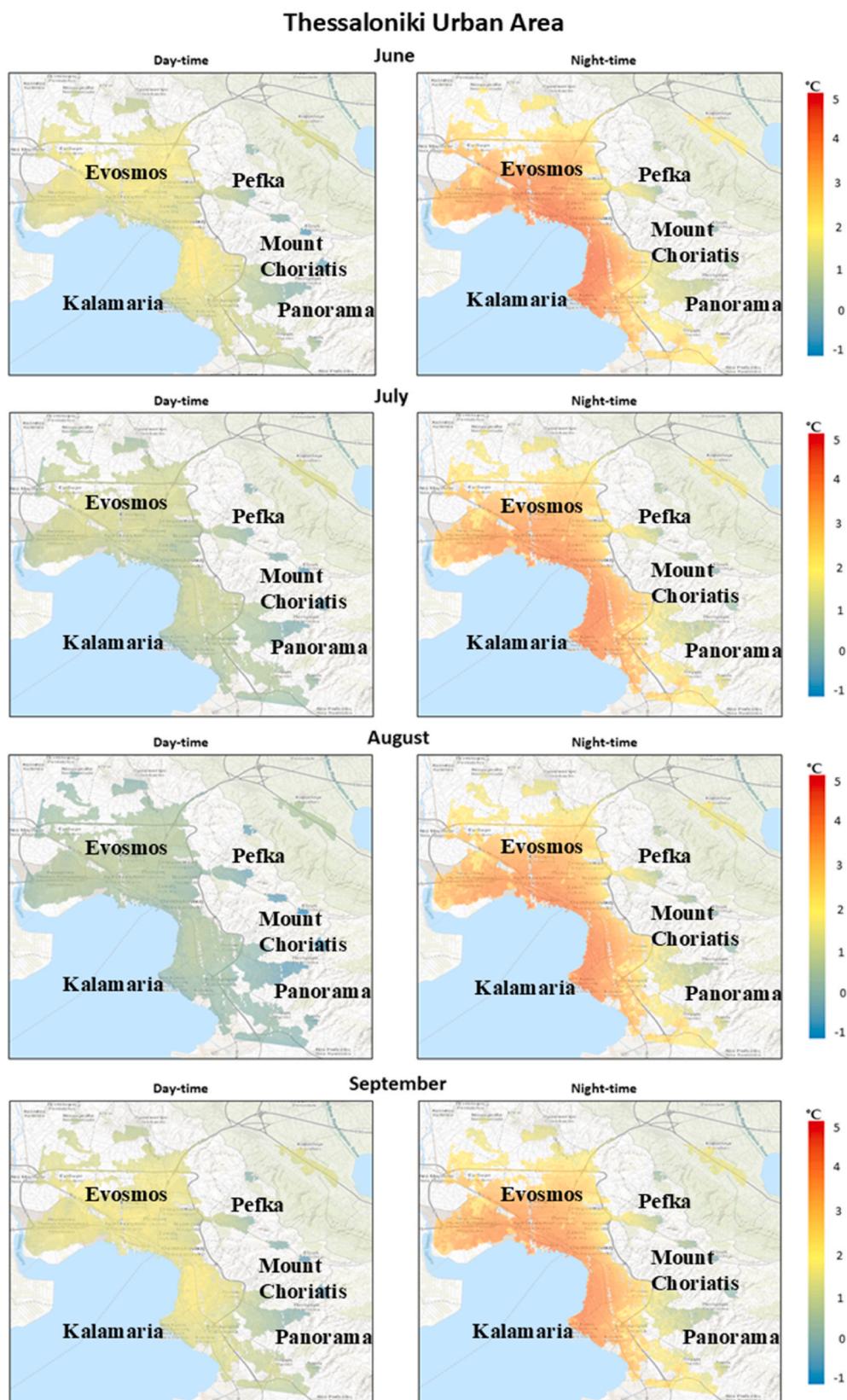


Fig. 5. Spatial variation of the urban heat island intensity per summer month during day and night in Thessaloniki urban area.3.2 Variable Importances during Summer.

would be expected, the nighttime UHI values are larger than in the daytime and range from 2 °C to 5 °C, especially in June and July. Additionally, during the summer months it is noted that the center of TUA, as well as the western districts (such as Evosmos) and the south-eastern districts (Kalamaria) experience more intense UHI, in both day and night. Not surprisingly, the common feature of these areas is intense urbanization with high urban densities and lack of green spaces. On the other hand, suburbs and semi-mountainous suburbs located at the foot of Mount Chortiatis (such as Pefka and Panorama), are related with the lowest values of UHII. This implies that Thessaloniki's UHI hot spots are determined by the combined effects of surface cover characteristics and topography, especially throughout the night.

The quantified attributions of five environmental factors to summer UHII during day and night for both urban areas, are illustrated to Fig. 6. It is apparent that the dominant factors that attribute to the summer UHII are RH, IMD and WS. As it mentioned above, it is well established that air humidity significantly influences the development of UHI since the existence of humidity variations between urban and rural have been highlighted from previous studies. Continuous evaporation during the night in urban areas and the removing vapor from atmosphere to the surface due to early morning dewfall in rural areas, have been reported as mechanisms of UMI development (Zipper et al., 2017; Hage, 1975). Furthermore, estimates of interactions between UMI and the development of the UHI, were found strong positive correlations between the vapor pressure difference and UHII (Lee, 1991), as well as that the synergistic effects of UHI and UMI aggravates the heat stress during hot summers (Huang et al., 2021). IMD, is also a variable with high importance. These results do not come as surprise, as it is widely acknowledged that build environment has a significant correlation with the urban thermal environment. Higher UHI intensities can be result from a density increase in impervious surfaces, such as paved areas and buildings. Closer examination in Fig. 6 reveals that, another significant determinant to UHII is WS during the summer in both urban areas. In general terms wind speed is one of the most significant meteorological factors that influence the development and the intensity of the UHI effect. For example, several studies agree that clear skies and calm winds favor the maximization of UHII, while UHII decreases with increasing wind speed (Kim and Baik, 2004). It is worth mentioning that, BH and SH are the less important variables in all cases.

To be more specific, in case of TUA, during the night the RH is calculated as the most important factor (0.68), followed by WS (0.23)

(Fig. 6). Higher surface temperature due to the nocturnal UHII effect may result in less condensation and more evaporation, as previous research suggested (Dudorova and Belan, 2019). Simultaneously, because of the increased radiative loss, it is anticipated that evaporation at the rural areas will be significantly impeded by the decrease in air temperature. These findings support those of Giannaros and Melas (2012), who found that Thessaloniki urban area was moister than the surroundings rural areas, while UHI plays a key role in shaping the urban moisture excess (UME), especially during the night. On the contrary, during the day the IMD and WS are the predominant contributors to UHII (0.41 and 0.28 for IMD and WS respectively), while the other environmental factors are less important for TUA. These results could be attributed to the direct effect of the sea breeze, which permits the transportation of humid marine air over Thessaloniki's urban regions along the coastline, especially during the warm period (Melas et al., 1994). On the other hand, the predominance of IMD to UHII, highlighting strong correlations between urban morphology and the microclimate during the day. Urban elements such as street layout, the urban canyon, and open and green spaces in TUA, contribute decisively to the creation of the microclimatic conditions, especially during hot days as noted by Kouklis and Yiannakou (2021). Although the building height is ranging from 3 to 41m (Fig. 2), TUA characterized by lack of open and green spaces and density building – blocks. Urban land uses in the Thessaloniki city are mixed, mainly residential, retail and services. These factors combined with impervious surfaces, directly affect the energy balance, the ventilation and insolation of the urban fabric, since the features and the geometry of Thessaloniki urban space determine both the degree of absorption and emission of solar radiation from building materials. It is worth noticing that the WS reported as the second more significant variable influencing the UHII. According to Giannaros and Melas (2012), 4 m/s is the threshold wind speed value, over which the UHI intensity is considerably reduced.

On the contrary to the case of Thessaloniki, in AUA the dominant environmental factor during the night is IMD (0.51), while the factor of RH is more important during the day (0.41) (Fig. 6). Comparing the two studies areas, AUA is more complex and quite larger, characterized by densely built areas, a network of high aspect ratio (height/width ratio) streets, tall building (3–102m), limited green and open space areas, lack of water evaporation and high heat storage capacities of building and surface materials. These high-density areas, provide shading to the surfaces of urban canyons and reduce solar radiation absorption, leading

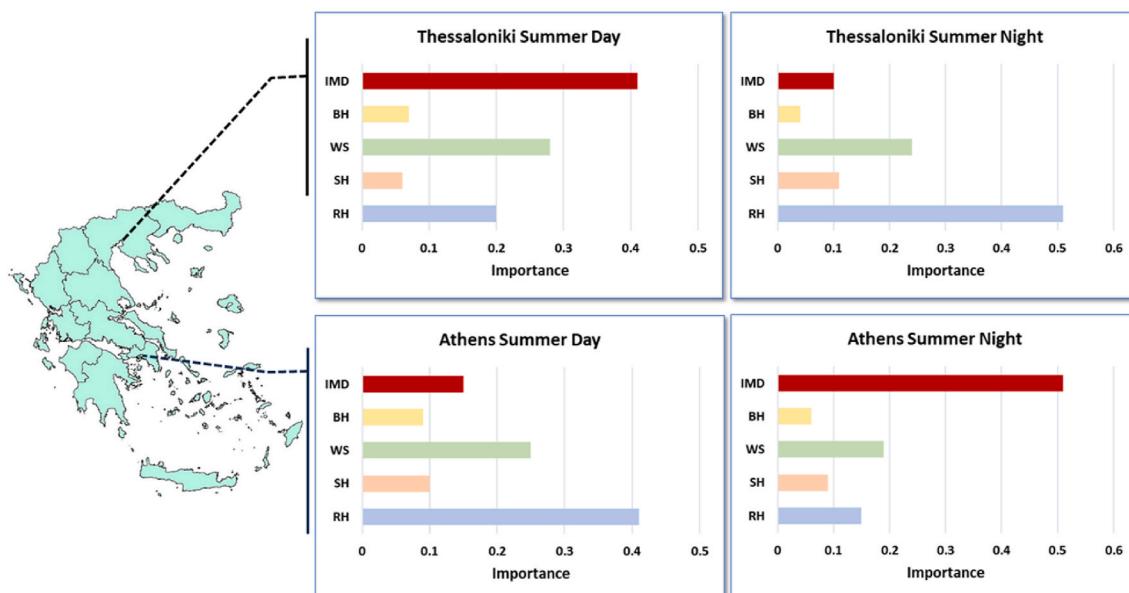


Fig. 6. The quantified importance of five environmental factors (IMD, BH, WS, SH and RH) to summer UHII during day and night for both urban areas.

to trapping the long-wave radiation emitted from the impervious surfaces and building materials at night. In this context, for AUA Agathangelidis et al. (2020) found that the impervious urban fraction had the most consistent and statistically significant relation to surface temperatures during the night. Similarly, Giannaros et al. (2023) in a recent study for AUA noted that nocturnal thermal phenomena were mainly affected by urban morphology, such as buildings characteristics, in contrast to atmospheric conditions which affect more rural areas surrounding AUA. Furthermore, Stathopoulou and Cartalis (2007) examined the UHI for Athens based on satellite thermal images of the city, concluded that thermal environment of Athens during nighttime influenced mainly by surface cover characteristics related with the residential urban zones (such as concrete and asphalt). As mentioned above, Fig. 6 illustrates that during the day the impact of RH to UHII becomes more significant. Such a finding is not uncommon, as during the day weather conditions associated with low wind speeds mainly from northerly directions and weak pressure gradient over the Aegean associated with high values of UHII for AUA (Petrou et al., 2023), favoring the formation of sea breezes and the transportation of humid air masses from the sea (Kassomenos and Katsoulis, 2006). On a similar note, Kassomenos et al. (2022) found that warm and humid conditions are strongly correlated with high UHII magnitudes during the day and with dry warm conditions at night. Additionally, Huang et al. (2021) reported that during the day in a compact high-rise city, such as AUA, anthropogenic heat emission from building cooling systems can aggravate UHI and UMI effects, by increasing air temperature and air specific humidity, respectively. Regarding the WS, it is notable that is the second most significant contributor to UHII in AUA, especially during the day. These findings support those of Founda and Santamouris (2017), who found that synoptic winds strongly affect the daytime UHII, especially during heat waves.

The underlying mechanism that contributes to the differential importance of the prominent factors between the two urban areas during the day and night, is possibly the built form of each city. AUA spreads out over the centre plain of a basin (Attica basin), several kilometers away from the sea, that is bound by four mountains (Fig. 2). This urban zone characterized by high density, extensive use of heat-retaining materials, limited green spaces, and congested transportation network. These factors collectively contribute to higher urban nighttime temperatures, justifying the predominance of IMD during the night. During the day, several researchers agree that AUA acts as an urban heat sink (Keramitsoglou et al., 2011; Giannaros et al., 2013), possibly reinforcing the importance of RH on UHII. On the other hand, in TUA the urban fabric stretches from the coast to higher elevations inland, where there is more plant cover and less urban congestion (Fig. 2). As a result, the thermal pattern of this coastal urban zone is in line with the city's architectural plan (Fig. 5). Such urban planning combined with the strong influence of the sea in this urban area, may explain the high and low attributions of RH on UHII during the night and day, respectively. Finally, this first experiment based on RF model was evaluated. Specifically, in order to determine if Grid search yielded a better model, we compare the base model with the best Grid search model, by calculating the accuracy using the mean average percentage error. It is worth mentioning that the accuracy is the proportion of correct predictions that model could make, while the error rate is the proportion of incorrect predictions. Table 1 summarizes the model's performance for Athens and Thessaloniki urban areas during daytime and nighttime and for

summer period. It is evident that using Grid search yielded an accuracy score that range between 93.45% and 96.85%. Furthermore, for Athens urban area the accuracies for both day and night are higher than Thessaloniki urban area. This may be driven by the amount of training data points, which is related to the size of urban areas and can contribute to the accuracy of the model. More specific, the amount of data points in case where in case of Athens (39,925 data points) is approximately three times higher than those of Thessaloniki urban area (12,190 data points).

3.2. Variable Importances during summer months

To achieve a detailed temporal distribution of the most important environmental variables influencing summer UHII, the relative importance of these environmental factors described above is also calculated for each month during the summer season. The results are displayed in Figs. 7 and 8, for Thessaloniki and Athens urban area for both daytime and nighttime.

In case of Thessaloniki, it is evident that during daytime the importance of IMD in influencing the UHII is more significant in July and August, the hottest months for TUA. The opposite pattern was observed for WS, the second most important factor, although the importance values are much lower compared to IMD. Additionally, the highest importance values for RH and SH are found in June and September during the day. Contrary to daytime, during the night the most important attribute to UHII is RH. The highest importance values for this factor, recorded in early/later summer (June and September). Closer examination reveals that the WH factor is the second important factor, especially during July and August, while IMD is the third most important factor for the hottest months of the summer for TUA during the night. Last but not least, in all cases for TUA, the BH and SH were found to be the most weakened factors during day and night (Fig. 7).

Moreover, Fig. 8 present the quantified attribution of all environmental factors to the UHII for each summer month, in AUA. It can be seen that, during the day RH is the dominant factor contributing to the UHII, with WS coming in second. The highest importance values for RH factor were found during August. The WS factor follows the same pattern as RH, reaching its maximum also during August. Interestingly, IMD factor during the day shows the highest values in June and September. On the other hand, during the night in AUA, it is clear that IMD is the most important factor for all months, captures the highest importance value in August. Considering the remaining environmental variables, WS and RH follow the same pattern, reaching their maximum importance values in early/later summer. As for the SH and BH factors, their significance is extremely weak (<0.1) in all cases for AUA (Fig. 8).

The performance of this set of experiments was also evaluated, by calculating the accuracy scores of the model on the data in each case. Table 2 summarizes the calculated scores. It general, it can be seen that the performance of model pretty impressive in all cases. More interestingly, the results revealed that during the second set of experiments, the model accuracy is slightly higher for the AUA, indicating that the amount of training data may also improve the model performance in this case of experiment.

At this point it is worth mentioning that ideally, this research would have a few more environmental variables to determine more broadly the importance of other urban metrics to the UHI phenomenon. However, this limitation cannot affect the accuracy of the model, since it is directly related to the spatial and temporal resolutions of the data. It is believed that in this study, the model performance was significantly improved and provided a better understanding of the underlying mechanisms, by employing fine resolution data. Therefore, in order to reveal more intricate relationships and patterns between the UHI and environmental factors as well as increase the precision and accuracy of UHI modelling, future research should consider combining more environmental data at higher resolutions.

Table 1
RF performance for Athens and Thessaloniki urban areas during summer period.

Urban Areas	Accuracy	
	Day	Night
Athens	96.85%	96.84%
Thessaloniki	93.45%	94.76%

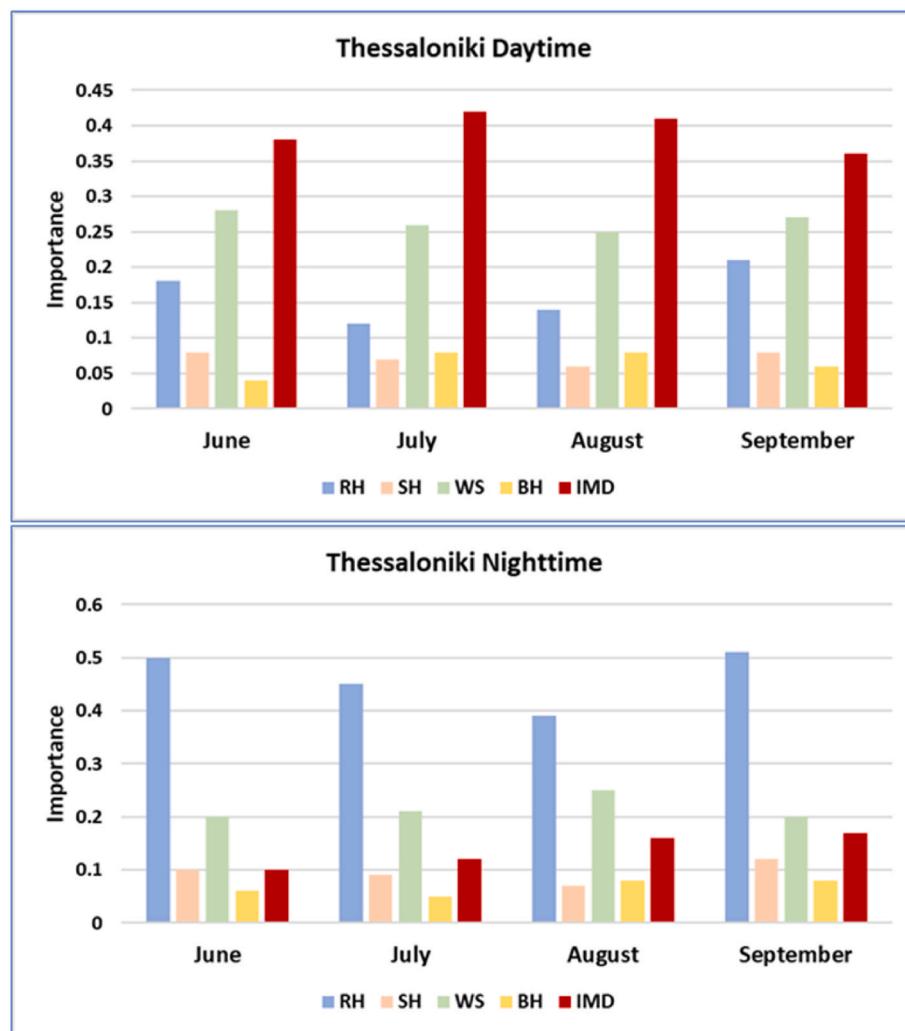


Fig. 7. The quantified importance of five environmental factors (IMD, BH, WS, SH and RH) to summer UHII for each month during day (up) and night (down) for TUA.

4. Conclusions

The quantification of various environmental attributes to summer UHII during the day and night in largest urban areas in Greece, is presented in the current study. For this purpose, a powerful machine learning algorithm was used in two sets of experiments, including calculating the variable importance of all environmental factors to UHII, both in summer and in each month of the warm period during day and night. The study's conclusions may be summed up as follows:

- For both urban areas, the dominant factors that attribute to the summer UHII are RH and IMD.
- In Thessaloniki urban area, during the summer night the RH is calculated as the most important factor, reaching its highest importance values in early summer. On the other hand, during the day the IMD is the predominant contributor to UHII, especially in July and August.
- On the contrary, in case of Athens urban area, it was found that the dominant environmental factor influencing the UHII during the summer night is IMD, especially in August. During the day in summer, RH is the most important factor on UHII, captures the highest importance value in August.
- WS was reported as a significant factor in all cases, reported as the second important factor, while BH and SH were found to be the most weakened factors in all cases.

5. The performance of the model was quite satisfactory in both experiments, suggesting potential associations between the amount of training data points and the accuracies scores of the model.

These findings highlight that studying the summer UHI effect under ML algorithms could provide essential information on the environmental contributors affecting the formation and the intensity of the UHI, in large urban areas. The considerable spatiotemporal variability of UHI attributions, highlights also that the importance of environmental factors to the UHI are relative to the location, to the time of the year, and to the climate variability of each area. Based on this, it is important to consider local environmental and socioeconomic variables when designing and implementing realistic sustainable heat mitigation strategies. For instance, in AUA the IMD factor during night reported as the dominant environmental factor influencing the UHI, yet not in the case of TUA. Furthermore, IMD and RH were the most significant environmental features in the formation of UHII for the summer season, in both urban areas. The importance of these features depends on the size of the urban area, as well as on the month and time of the day. This demonstrates how crucial is to increase green spaces and vegetation in strategic location of these urban areas in order to minimize the impacts of UHI. Additionally, effective heat island mitigation plans may focus on improving the urban planning, by installing green or cool roofs, white and pervious pavements and planning ventilation corridors within urban areas. Measurement and collection of accurate humidity and wind

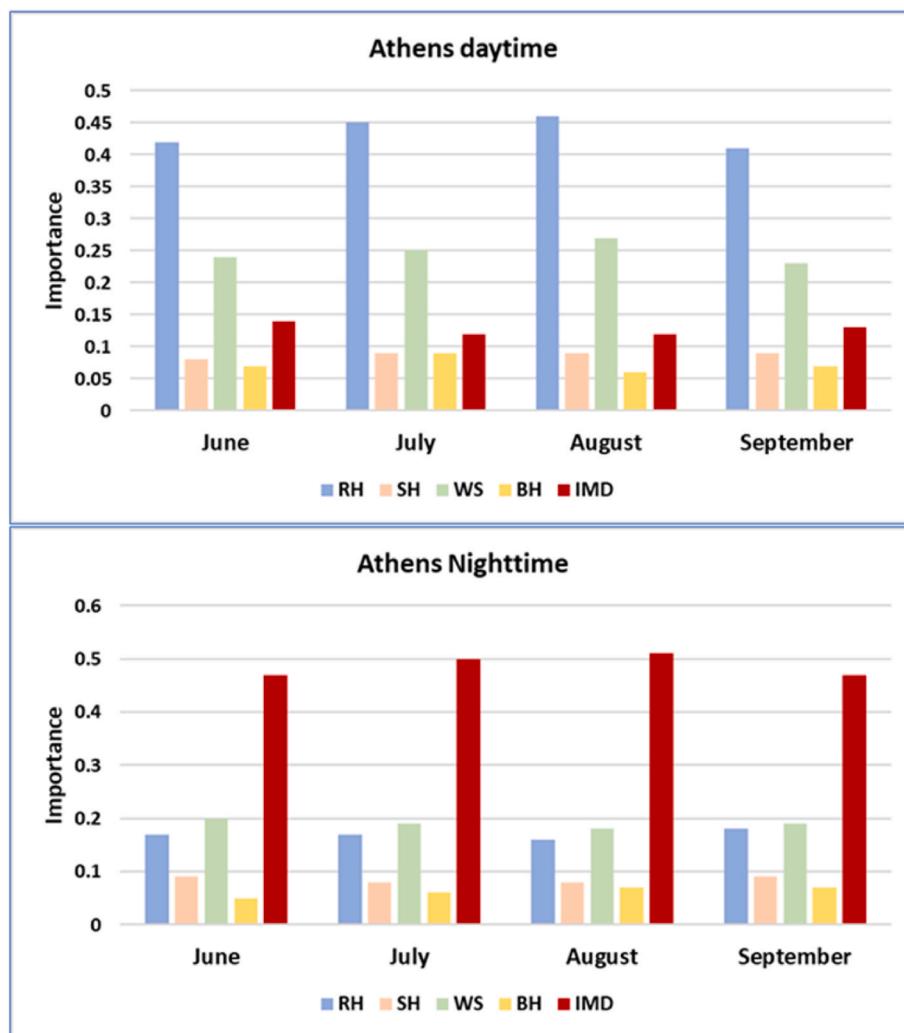


Fig. 8. The quantified importance of five environmental factors (IMD, BH, WS, SH and RH) to summer UHII for each month during day (up) and night (down) for AUA.

Table 2
RF performance for Athens and Thessaloniki urban areas during each month of the summer period.

	Accuracy			
	Athens		Thessaloniki	
	Day	Night	Day	Night
June	93.35%	96.84%	92.89%	95.45%
July	94.58%	95.85%	93.78%	94.37%
August	94.03%	94.87%	93.32%	93.54%
September	95.58%	96.32%	94.02%	93.88%

data, as well as the accurately predicting of these meteorological factors variables in Athens and Thessaloniki should be increased, since RH and WS are significant determinants to UHII in both areas. Taking into account the presented results, it becomes clear that having a catalogue or an archive of diverse and well-coordinated solutions (such as increase green spaces or planning ventilation corridors within urban areas) rather than relying just on one approach is more realistic for effectively mitigating the UHI impact in various urban areas.

In general terms, studies have demonstrated that BH impacts the UHI effect through increased surface area, heat emissions, urban canyon effects, altered wind patterns, and reduced green spaces (Li et al., 2020; Bakarman and Chang, 2015; Loughner et al., 2012). However, the

relationship between building height and urban heat island (UHI) effects is complex and multifaceted. For instance, Wang and Xu (2021) reported that when a building's height is between 0 and 66 m, the land surface temperatures will drastically drop as the building becomes higher, suggesting a notable cooling impact in this height range in megacities. On the contrary, compared to low (~7m) and middle (~24m) rise buildings, the UHII of high (~71m) rise buildings is 0–0.7 °C greater, as noted by Xi et al. (2021). Nevertheless, it is noteworthy that in both urban areas in this study, the urban morphology factor BH is found of less importance on UHII than IMD factor. This could be partly attributed to the fact that BH factor possess no significant spatial variability and is comparatively constant, comparing to IMD. The BH grid points that matched the meteorological data, refer to urban areas where the height of the building does not show significant variations in both urban areas. Another explanation could be that in both urban areas, the main urban characteristic is the building density and not so much the height, which constitutes constraint in this study. Concerning the heat mitigation strategies, these finding highlight that it is crucial to accurately quantify the contribution of anthropogenic activities and urban morphology to the UHI impact than modify the patterns of human activity or urban shape in already-existing cities.

Although the performance of the RF model was quite impressive, future research would benefit from extending the number of critical factors that have been missed from this research, such as surface albedo, vegetation coverage, road distribution data and socioeconomic

characteristics for each urban area. In this context, there is belief that more attention should be paid in ML techniques for UHI studies, in the hope that the more application of different ML algorithms, the more interesting and comparable results will emerge. Lastly, the municipal authority in urban areas might also use these findings to implement urban heat mitigation, helping decision-makers decide what has to be done first and where to concentrate their efforts.

CRediT authorship contribution statement

Ilias Petrou: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pavlos Kassomenos:** Writing – review & editing, Supervision, Data curation, Conceptualization.

Declaration of competing interest

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

Data availability

No data was used for the research described in the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jenvman.2024.122255>.

References

- Addas, A., 2023. Machine learning techniques to Map the impact of urban heat island: investigating the city of jeddah. *Land* 12 (6), 1159. <https://doi.org/10.3390/land12061159>.
- Agathangelidis, I., Cartalis, C., Santamouris, M., 2020. Urban morphological controls on surface thermal dynamics: a comparative assessment of major European cities with a focus on Athens, Greece. *Climate* 8 (11), 131. <https://doi.org/10.3390/cli8110131>.
- Akbari, H., Cartalis, C., Kolokotsa, D., Muscio, A., Pisello, A.L., Rossi, F., Santamouris, M., Symnet, A., Wong, N.H., Zinzi, M., 2015. Local climate change and urban heat island mitigation techniques – the state of the art. *J. Civ. Eng. Manag.* 22 (1), 1–16. <https://doi.org/10.3846/13923730.2015.1111934>.
- Al-Obaidi, I., Rayburg, S., Pórolnica, M., Neave, M., 2021. Assessing the impact of wind conditions on urban heat islands in large Australian cities. *Journal of Ecological Engineering* 22 (11), 1–15. <https://doi.org/10.12911/22998993/142967>.
- Bakarman, M.A., Chang, J.D., 2015. The influence of height/width ratio on urban heat island in hot-arid climates. *Procedia Eng.* 118, 101–108. <https://doi.org/10.1016/j.proeng.2015.08.408>.
- Bauer, T.J., 2020. Interaction of urban heat island effects and land-sea breezes during a New York city heat event. *J. Appl. Meteorol. Climatol.* 59, 477–495. <https://doi.org/10.1175/JAMC-D-19-0061.1>.
- Bonino, G., Galimberti, G., Masina, S., McAdam, R., Clementi, E., 2024. Machine learning methods to predict sea surface temperature and marine heatwave occurrence: a case study of the Mediterranean Sea. *Ocean Sci.* 20, 417–432. <https://doi.org/10.5194/os-20-417-2024>.
- Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsedbazar, N.E., Linlin, L., Tarko, A., 2020. Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015–2019: Product User Manual. Zenodo, Geneve, Switzerland. September 2020.
- Buschow, S., Keller, J., Wahl, S., 2024. Explaining heatwaves with machine learning. *Q. J. R. Meteorol. Soc.* 1–15. <https://doi.org/10.1002/qj.4642>.
- Chakraborty, T., Venter, Z.S., Qian, Y., Lee, X., 2022. Lower urban humidity moderates outdoor heat stress. *AGU Advances* 3, e2022AV000729. <https://doi.org/10.1029/2022AV000729>.
- Chandler, T.J., 1961. The changing climate of London's heat island. *Geography* 46, 295–307.
- Chen, H., Huang, J.J., Li, H., Wei, Y., Zhu, X., 2023. Revealing the response of urban heat island effect to water body evaporation from main urban and suburb areas. *J. Hydrol.* 623, 129687. <https://doi.org/10.1016/j.jhydrol.2023.129687>.
- Chen, Y., Zhang, N., 2018. Urban heat island mitigation effectiveness under extreme heat conditions in the suzhou-wuxi-changzhou metropolitan area, China. *J. Appl. Meteorol. Climatol.* 57, 235–253. <https://doi.org/10.1175/JAMC-D-17-0101.1>.
- De Ridder, K., Lauwaet, D., Maiheu, B., 2015. UrbClim – a fast urban boundary layer climate model. *Urban Clim.* 12, 21–48. <https://doi.org/10.1016/j.ulc.2015.01.001>.
- Dudorova, N.V., Belan, B.D., 2022. The energy model of urban heat island. *Atmosphere* 13, 457. <https://doi.org/10.3390/atmos13030457>.
- Dudorova, V.N., Belan, B.D., 2019. The role of evaporation and condensation of water in the formation of the urban heat island. *Proc. SPIE 11208, 25th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 112088J, 18 December 2019.*
- European environment agency (EEA), 2012. Building Height. <https://land.copernicus.eu/local/urban-atlas/building-height-2012>. (Accessed 22 December 2023).
- Founda, D., Santamouris, M., 2017. Synergies between urban heat island and heat waves in Athens (Greece), during an extremely hot summer (2012). *Sci. Rep.* 7, 10973 <https://doi.org/10.1038/s41598-017-11407-6>.
- Furuya, M.T.G., Furuya, D.E.G., de Oliveira, L.Y.D., da Silva, P.A., Cicerelli, R.E., Gonçalves, W.N., Junior, J.M., Osco, L.P., Ramos, A.P.M., 2023. A machine learning approach for mapping surface urban heat island using environmental and socioeconomic variables: a case study in a medium-sized Brazilian city. *Environ. Earth Sci.* 82, 325. <https://doi.org/10.1007/s12665-023-11017-8>.
- Giannaros, C., Agathangelidis, I., Papavasileiou, G., Galanaki, E., Kotrotsi, V., Lagouvardos, K., Giannaros, T.M., Cartalis, C., Matzarakis, A., 2023. The extreme heat wave of July–August 2021 in the Athens urban area (Greece): atmospheric and human-biometeorological analysis exploiting ultra-high resolution numerical modeling and the local climate zone framework. *Sci. Total Environ.* 857, 159300. <https://doi.org/10.1016/j.scitotenv.2022.159300>.
- Giannaros, T.M., Melas, D., 2012. Study of the urban heat island in a coastal Mediterranean City: the case study of Thessaloniki, Greece. *Atmos. Res.* 118, 103–120. <https://doi.org/10.1016/j.atmosres.2012.06.006>.
- Giannaros, T.M., Melas, D., Daglis, I.A., Keramitsoglou, I., Kourtidis, K., 2013. Numerical study of the urban heat island over Athens (Greece) with the WRF model. *Atmos. Environ.* 73, 103–111. <https://doi.org/10.1016/j.atmosenv.2013.02.055>.
- Grimmond, C.S.B., Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Meteorol. Climatol.* 38, 1262–1292. [https://doi.org/10.1175/1520-0450\(1999\)038<1262:APOUAD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1262:APOUAD>2.0.CO;2).
- Hage, K.D., 1975. Urban–rural humidity differences. *J. Appl. Meteorol.* 14, 1277–1283.
- Han, H., Guo, X., Yu, H., 2016. Variable selection using mean decrease accuracy and mean decrease gini based on random forest, 2016 7th. In: IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, pp. 219–224.
- Hooyberghs, H., Berckmans, J., Lefebre, F., De Ridder, K., 2019. C3S_422_Lot2 SIS European health. UrbClim Extra Documentation 1–18. Available online: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-urban-climate-cities?tab=form>. (Accessed 15 December 2023).
- Hou, H., Longyang, Q., Su, H., Zeng, R., Xu, T., Wang, Z.H., 2023. Prioritizing environmental determinants of urban heat islands: a machine learning study for major cities in China. *Int. J. Appl. Earth Obs. Geoinf.* 122, 103411. <https://doi.org/10.1016/j.jag.2023.103411>.
- Huang, X., Song, J., Wang, C., Chui, T.F.M., Chan, P.W., 2021. The synergistic effect of urban heat and moisture islands in a compact high-rise city. *Build. Environ.* 205, 108274. <https://doi.org/10.1016/j.buildenv.2021.108274>.
- Kantzoura, A., Kosmopoulos, P., Zoras, S., 2012. Urban surface temperature and microclimate measurements in Thessaloniki. *Energy Build.* 44, 63–72.
- Kassomenos, P., Katsoulis, B., 2006. Mesoscale and macroscale aspects of the morning urban heat island around Athens, Greece. *Meteorol. Atmos. Phys.* 94, 209–218. <https://doi.org/10.1007/s00703-006-0191-x>.
- Kassomenos, P., Kissas, G., Petrou, I., Begou, P., Khan, H.S., Santamouris, M., 2022. The influence of daily weather types on the development and intensity of the urban heat island in two Mediterranean coastal metropolises. *Sci. Total Environ.* 819, 153071. <https://doi.org/10.1016/j.scitotenv.2022.153071>.
- Keppas, S., Parliari, D., Kontos, S., Poupkou, A., Papadogiannaki, S., Tzoumaka, P., Kelessis, A., Melas, D., 2021. Urban heat island and future projections: a study in Thessaloniki, Greece. In: Dobrinkova, N., Gadzhiev, G. (Eds.), Environmental Protection and Disaster Risks. EnviroRISK 2020. Studies in Systems, Decision and Control, vol. 361. Springer, Cham. https://doi.org/10.1007/978-3-030-70190-1_14.
- Keramitsoglou, I., Kiranoudis, C.T., Ceriola, G., Weng, Q., Rajasekar, U., 2011. Identification and analysis of urban surface temperature patterns in Greater Athens, Greece, using MODIS imagery. *Rem. Sens. Environ.* 115, 3080–3090.
- Khan, H.S., Santamouris, M., Kassomenos, P., Paolini, R., Caccetta, P., Petrou, I., 2021. Spatiotemporal variation in urban overheating magnitude and its association with synoptic airmasses in a coastal city. *Sci. Rep.* 11, 6762. <https://doi.org/10.1038/s41598-021-86089-2>.
- Kim, Y.H., Baik, J.J., 2004. Daily maximum urban heat island intensity in large cities of Korea. *Theor. Appl. Climatol.* 79, 151–164. <https://doi.org/10.1007/s00704-004-0070-7>.
- Kim, Y.H., Baik, J.J., 2005. Spatial and temporal structure of the urban heat island in Seoul. *J. Appl. Meteorol.* 44, 651–659.
- Kouklis, G.R., Yiannakou, A., 2021. The contribution of urban morphology to the formation of the microclimate in compact urban cores: a study in the city center of Thessaloniki. *Urban Science* 5 (2), 37. <https://doi.org/10.3390/urbansci5020037>.
- Kourtidis, K., Georgoulias, A.K., Rapsomanikis, S., Amiridis, V., Keramitsoglou, I., Hooyberghs, H., Maiheu, B., Melas, D., 2015. A study of the hourly variability of the urban heat island effect in the Greater Athens Area during summer. *Sci. Total Environ.* 517, 162–177.
- Kuttler, W., Weber, S., Schonnefeld, J., Hesselschwerdt, A., 2017. Urban/rural atmospheric water vapour differences and urban moisture excess in Krefeld, Germany. *Int. J. Climatol.* 25, 2005–2015. <https://doi.org/10.1002/joc.1558>.
- Kyros, G., Manolas, I., Diamantaras, K., Dafis, S., Lagouvardos, K., 2023. A machine learning approach for rainfall nowcasting using numerical model and observational

- data. In: Environmental Sciences Proceedings, vol. 26, p. 11. <https://doi.org/10.3390/environsciproc2023026011>, 1.
- Lauwaet, D., De Ridder, K., Saeed, S., Brisson, E., Chatterjee, F., van Lipzig, N.P.M., Maiheu, B., Hooyberghs, H., 2016. Assessing the current and future urban heat island of Brussels. *Urban Clim.* 15, 1–15. <https://doi.org/10.1016/j.ulclim.2015.11.008>.
- Lee, D.O., 1991. Urban-rural humidity differences in London. *Int. J. Climatol.* 11, 577–582. <https://doi.org/10.1002/joc.3370110509>.
- Lee, K., Kim, Y., Sung, H.C., Ryu, J., Jeon, S.W., 2020. Trend analysis of urban heat island intensity according to urban area change in asian mega cities. *Sustainability* 12, 112. <https://doi.org/10.3390/su12010112>.
- Li, H., Yang, Y., Wang, H., Wang, P., Yue, X., Liao, H., 2022. Projected aerosol changes driven by emissions and climate change using a machine learning method. *Environ. Sci. Technol.* 56 (7), 3884–3893.
- Li, Y., Schubert, S., Kropp, J.P., Rybski, D., 2020. On the influence of density and morphology on the Urban Heat Island intensity. *Nat. Commun.* 11, 2647. <https://doi.org/10.1038/s41467-020-16461-9>.
- Liu, B., Guo, X., Jiang, J., 2023. How urban morphology relates to the urban heat island effect: a multi-indicator study. *Sustainability* 15 (14), 10787. <https://doi.org/10.3390/su151410787>.
- Loughner, C., Allen, D., Zhang, D., Pickering, K., Dickerson, R., Landry, L., 2012. Roles of urban tree canopy and buildings in urban heat island effects: parameterization and preliminary results. *J. Appl. Meteorol. Climatol.* 51 (10), 1775–1793. <https://doi.org/10.1175/JAMC-D-11-0228.1>.
- Melas, D., Ziomas, I.C., Zerefos, C.S., 1994. A numerical study of dispersion in a coastal urban area. Part A. Air flow. *Fresenius Environ. Bull.* 3, 306–311.
- Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinf.* 10, 213. <https://doi.org/10.1186/1471-2105-10-213>.
- Milelli, M., Bassani, F., Garbero, V., Poggi, D., Hardenberg, J., Ridolfi, L., 2023. Characterization of the urban heat and dry island effects in the turin metropolitan area. *Urban Clim.* 47, 101397 <https://doi.org/10.1016/j.ulclim.2022.101397>.
- Morabito, M., Crisci, A., Georgiadis, T., Orlandini, S., Munafò, M., Congedo, L., Rota, P., Zazzi, M., 2018. Urban imperviousness effects on summer surface temperatures nearby residential buildings in different urban zones of parma. *Rem. Sens.* 10 (1), 26. <https://doi.org/10.3390/rs10010026>.
- Narumi, D., Levinson, R., Shimoda, Y., 2021. Effect of urban heat island and global warming countermeasures on heat release and carbon dioxide emissions from a detached house. *Atmosphere* 12 (5), 572. <https://doi.org/10.3390/atmos12050572>.
- Ngarambe, J., Oh, J.W., Su, M.A., Santamouris, M., Yun, G.Y., 2021. Influences of wind speed, sky conditions, land use and land cover characteristics on the magnitude of the urban heat island in Seoul: an exploratory analysis. *Sustain. Cities Soc.* 71, 102953 <https://doi.org/10.1016/j.scs.2021.102953>.
- Oke, T.R., 1976. The distinction between canopy and boundary-layer heat islands. *Atmosphere* 14, 268–277.
- Pal, K., Patel, B.V., 2020. Data classification with K-Fold cross validation and holdout accuracy estimation methods with 5 different machine learning techniques. In: 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). Erode, India, pp. 83–87.
- Petrou, I., Kyriazis, N., Kassomenos, P., 2023. Evaluating the spatial and temporal characteristics of summer urban overheating through weather types in the Attica region, Greece. *Sustainability* 15, 10633. <https://doi.org/10.3390/su151310633>.
- Rao, P., Tassinari, P., Torreggiani, D., 2023. Exploring the land-use urban heat island nexus under climate change conditions using machine learning approach: a spatio-temporal analysis of remotely sensed data. *Heliyon* 9 (8), e18423. <https://doi.org/10.1016/j.heliyon.2023.e18423>.
- Reis, C., Lopes, A., Nouri, S.A., 2022. Assessing urban heat island effects through local weather types in Lisbon's Metropolitan Area using big data from the Copernicus service. *Urban Clim.* 43, 101168 <https://doi.org/10.1016/j.ulclim.2022.101168>.
- Runnalls, K.E., Oke, T.R., 2000. Dynamics and controls of the near-surface heat island of Vancouver, British Columbia. *Phys. Geogr.* 21 (4), 283–304. <https://doi.org/10.1080/02723646.2000.10642711>.
- Santamouris, M., Cartalis, C., Synnefa, A., 2015. Local urban warming, possible impacts and a resilience plan to climate change for the historical center of Athens, Greece. *Sustain. Cities Soc.* 19, 281–291. <https://doi.org/10.1016/j.scs.2015.02.001>.
- Skoufali, I., Battisti, A., 2019. Microclimate of urban canopy layer and outdoor thermal comfort: a case study in pavlou mela, Thessaloniki. *Urban Science* 3, 84. <https://doi.org/10.3390/urbansci3030084>.
- Stathopoulou, M., Cartalis, C., 2007. Daytime urban heat islands from Landsat ETM+ and Corine land cover data: an application to major cities in Greece. *Sol. Energy* 81 (3), 358–368. <https://doi.org/10.1016/j.solener.2006.06.014>.
- Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and qsar modeling. *J. Chem. Inf. Comput. Sci.* 43 (6), 1947–1958. <https://doi.org/10.1021/ci034160g>.
- van der Schrieck, T., Varotsos, K.V., Giannakopoulos, C., Founda, D., 2020. Projected future temporal trends of two different urban heat islands in Athens (Greece) under three climate change scenarios: a statistical approach. *Atmosphere* 11 (6), 637. <https://doi.org/10.3390/atmos11060637>.
- Verdonck, M.-L., Demuzere, M., Hooyberghs, H., Beck, C., Cyrys, J., Schneider, A., Dewulf, R., Van Coillie, F., 2018. The potential of local climate zones maps as a heat stress assessment tool, supported by simulated air temperature data. *Landsl. Urban Plann.* 178, 183–197. <https://doi.org/10.1016/J.LANDURBPLAN.2018.06.004>.
- Wang, Y., Li, Y., Chan, P.W., 2015. The impact of urbanization on moisture excess in Hong Kong. *Energy Proc.* 78, 3061–3065.
- Wang, M., Xu, H., 2021. The impact of building height on urban thermal environment in summer: a case study of Chinese megacities. *PLoS One* 16 (4), e0247786. <https://doi.org/10.1371/journal.pone.0247786>.
- Xi, C., Ren, C., Wang, J., Feng, Z., Cao, S., 2021. Impacts of urban-scale building height diversity on urban climates: a case study of Nanjing, China. *Energy Build.* 251, 111350 <https://doi.org/10.1016/j.enbuild.2021.111350>.
- Yang, J., Jin, S., Xiao, X., Jin, C., Xia, J.C., Li, X., Wang, S., 2019. Local climate zone ventilation and urban land surface temperatures: towards a performance-based and wind-sensitive planning proposal in megacities. *Sustain. Cities Soc.* 47, 101487 <https://doi.org/10.1016/j.scs.2019.101487>.
- Yang, J., Xin, J., Zhang, Y., Xiao, X., Xia, J.C., 2022. Contributions of sea–land breeze and local climate zones to daytime and nighttime heat island intensity. *npj Urban Sustainability* 2, 12. <https://doi.org/10.1038/s42949-022-00055-z>.
- Yoo, S., 2018. Investigating important urban characteristics in the formation of urban heat islands: a machine learning approach. *Journal of Big Data* 5, 2. <https://doi.org/10.1186/s40537-018-0113-z>.
- Zhang, K., Cao, C., Chu, H., Zhao, L., Zhao, J., Lee, X., 2023. Increased heat risk in wet climate induced by urban humid heat. *Nature* 617, 738–742. <https://doi.org/10.1038/s41586-023-05911-1>.
- Zhang, P., Imhoff, M.L., Bounoua, L., Wolfe, R.E., 2012. Exploring the influence of impervious surface density and shape on urban heat islands in the northeast United States using MODIS and Landsat. *Can. J. Rem. Sens.* 38 (4), 111.
- Zhang, Y., Wang, Y., Ding, N., Yang, X., 2022. Spatial pattern impact of impervious surface density on urban heat island effect: a case study in xuzhou, China. *Land* 11 (12), 2135. <https://doi.org/10.3390/land11122135>.
- Zheng, Y., Li, W., Fang, C., Feng, B., Zhong, Q., Zhang, D., 2023. Investigating the impact of weather conditions on urban heat island development in the subtropical city of Hong Kong. *Atmosphere* 14 (2), 257. <https://doi.org/10.3390/atmos14020257>.
- Zhou, B., Lauwaet, D., Hooyberghs, H., De Ridder, K., Kropp, J.P., Rybski, D., 2016. Assessing seasonality in the surface urban heat island of london. *J. Appl. Meteorol. Climatol.* 55, 493–505.
- Zhou, X., Chen, H., 2018. Impact of urbanization-related land use land cover changes and urban morphology changes on the urban heat island phenomenon. *Sci. Total Environ.* 635 <https://doi.org/10.1016/j.scitotenv.2018.04.091>, 1467–147.
- Zipper, S.C., Schatz, J., Kucharik, C.J., Loheide II, S.P., 2017. II, Urban heat island-induced increases in evapotranspirative demand. *Geophys. Res. Lett.* 44, 873–881. <https://doi.org/10.1002/2016GL072190>.