The Role of Metallic and Acid Sites of Ru-Nb-Si Catalysts in the Transformation of Levulinic Acid to -Valerolactone

An innovative approach to select urban-rural sites for Urban Heat Island analysis: the case of Turin (Italy)

Francesca Bassani[a,b,*], Valeria Garbero[b], Davide Poggi[a], Luca Ridolfi[a], Jost von Hardenberg[a], and Massimo Milelli[c]

[a] *Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, 10129 Turin, Italy*
[b] *Department of Meteorology, Climate and Air Quality, Arpa Piemonte, 10135 Turin, Italy*
[c] *CIMA Foundation, 17100 Savona, Italy*

Abstract

A novel metric – the Mean Temperature Difference (MTD) – is proposed for the selection of urban-rural pairs of stations needed in the Urban Heat Island (UHI) quantification. This metric highlights the thermal pattern typical of each weather station with respect to the average one of the area of interest. Afterwards, Principal Component Analysis is adopted to cluster stations into subsets exhibiting similar thermal behaviors. The joint use of MTD and PCA allows one to classify stations objectively and without the need of preliminary assumptions about the station landscapes. An application to the metropolitan area of Turin (Italy) and a comparison with validated methods to select urban-rural pairs demonstrate that the proposed approach is easily interpretable and reliable also when the study area exhibits a non-trivial landscape categorization.

Keywords: Urban Heat Island, urban-rural pairs, MTD

1. Introduction

The meteorological phenomenon known as Urban Heat Island (UHI) is one of the main effects produced by increasing urbanization (Landsberg, 1981; Tzavali et al., 2015) and a significant example of anthropogenic climate modification (Arnfield, 2003). UHI refers to the warmer temperatures experienced by a city with respect to its rural surrounding area, mainly due to the different thermal properties between urbanized and natural lands, anthropogenic heat emissions, human-induced pollution and limited wind blowing among buildings (Oke, 1973, 1976; Rizwan et al., 2008). In the long and well-documented urban heat island literature (Stewart, 2011), UHI has been commonly quantified as the difference, in terms of air temperatures, between pairs of urban and rural measurement sites (Oke, 1973; Kim and Brown, 2021) or between a spatial average of several urban and/or several rural stations (e.g., Hoffmann and Schlünzen, 2013). This difference

★The formal publication is available at https://doi.org/10.1016/j.uclim.2022.101099
*Corresponding author
Email address: `francesca.bassani@polito.it` (Francesca Bassani)

is crucial in determining the UHI intensity and requires choosing a non trivial definition of which stations are "urban" and "rural". In his work, Stewart (2007) highlighted the difficulty in the definition of the urban-rural dichotomy, because the demarcation between "urban" and "rural" is artificial and many relevant local-scale aspects should be taken into account. Recent studies tried to address this critical issue by proposing new methods that (i) highlight different thermal behaviors in urban-rural pairs – e.g., the approaches based on the thermal day-to-day variation (Karl et al., 1995; Gough, 2008; Mohsin and Gough, 2012; Tam et al., 2015; Wu et al., 2017; Anderson et al., 2018) or the mean daily excursion (Milelli, 2016) – or (ii) identify the stations called "peri-urban", i.e. those located close to the urban-rural interface, by focusing on the day-to-day warm and cold transitions (Gough, 2020).

Another important approach to classify the stations is the Local Climate Zones (LCZs) Classification System proposed by Stewart and Oke (2012). By using criteria concerning aspects that control the local surface climates, this climate-based tool classifies the landscape (i.e., a local-scale area of land) in 17 regions characterized by uniform surface cover, structure, material and human activity. The classification covers both built and natural environments and each zone is characterized by a distinctive near-surface temperature regime.

Despite the variety of methods, a key point is that all of them need a preliminary classification of the stations. In order to overcome this possible source of arbitrariness, we propose a novel method, which we call the "Mean Temperature Difference" (hereinafter MTD). The MTD is a data-based approach aiming to recognize and differentiate the thermal behavior of the urban context with respect to its surrounding less populated area. The former is different – in terms of thermal response – from the latter, mainly because of its predominantly impervious land cover type and the presence of sheltering constructions, which trap heat during the day and release it during the night resulting in higher night-time temperatures.

The strength of the proposed MTD method is the capability to objectively identify these different thermal behaviors, without assuming *a priori* which sites pertain to the urban and rural categories. In this sense, the MTD approach can complete and be of support to the Local Climate Zones Classification. In fact, as stated by Stewart and Oke (2012), the intention of the LCZs is not to supplant the categories "urban" and "rural" in the heat island issues, but to provide a more conscious and constrained use of these categories when describing the local conditions of the stations.

Starting from a group of stations – heterogeneous in terms of LCZs – and adopting the Principal Component Analysis (Jolliffe and Cadima, 2016) as a clustering method, the proposed approach is able to objectively and clearly identify the different thermal behavior of the stations. It allows a clear distinction of what is the typical "urban" pattern, the different "rural" one and also what does not fall into either categories.

No choice about whether a site is "urban" or "rural" is made *a priori* and no parameters to calibrate enter the procedure, so that its results are clear, immediate and easy to apply. This makes the method objective, totally data-based and unrelated to any preliminary landscape classification.

To show the features of the proposed method, we apply it to the city of Turin (Italy) and its surrounding area, which is characterized by a quite complex morphology (orographic and hydrographic heterogeneity, different land uses, etc.), making it suitable to test the proposed metric.

The paper is organized as follows. Section 2 describes the MTD metric. Section 3 reports the results of the application of the proposed method to Turin area and highlights its advantages. In Section 4 we discuss the applicability of the MTD. Section 5 shows the comparison of the MTD-based approach with existing methods to select proper urban-rural pairs. Finally, some conclusions are drawn.

2. Proposed method

The idea behind the MTD metric is to detect similar behaviors among stations and it is based on two main steps: (i) the evaluation of a metric characterizing the thermal behavior of each measurement site, and (ii) the adoption of the Principal Component Analysis (Jolliffe, 2002; Wilks, 2011), in order to capture common performances of such metric and to cluster the stations into distinct groups.

We start describing the first step. The variable considered for each weather station S is the monthly-averaged hourly temperature $T^S_{i,M}$, where subscripts $i = 1, 2, .., 24$ and $M = Jan, Feb, ..., Dec$ refer to the hours and months, respectively. For example, the monthly temperature value at 01:00 hours for January (i.e., $T^S_{1,Jan}$) refers to the climatological average over all years of all the temperatures of January registered at 01:00. The metric MTD is defined, for each hour i, each month M and each station S, as:

$$\mathrm{MTD}^S_{i,M} = T^S_{i,M} - \overline{T^S_{i,M}} - \langle T^S_{i,M} - \overline{T^S_{i,M}} \rangle \tag{1}$$

in which the overbar refers to a temporal average over all months and hours of $T^S_{i,M}$, and $\langle\cdot\rangle$ represents the spatial mean among all stations included in the study area (i.e., $\langle\cdot\rangle = \sum_j (T_{i,M})_j / N_S$, where j ranges from 1 to the number of stations N_S). The first two terms at the right-hand side of Eq. (1) define the anomaly of temperatures at each hour and month compared to their temporal average for that station. The last two terms remove the mean anomaly across all stations for that hour and month. Positive values of MTD indicate that the station S, at a certain hour and month, is characterized by higher air temperature anomalies than the mean of the other stations, while the opposite occurs for negative MTD.

In the second step of the proposed method, the MTD values are organized in a matrix to which the Principal Component Analysis (PCA) is applied. The matrix (hereinafter **MTD**) has dimensions 288xN_S:

its rows contain the values $\text{MTD}_{i,M}$ corresponding to the 24 hours for every month (24x12=288), and each column refers to a station S: therefore, **MTD** defines a cloud of 288 points in a N_S-dimensional space. PCA is commonly used in the atmospheric science and it is considered a robust tool in climatology and meteorology (e.g., Lorenz, 1956; Hannachi et al., 2007; Demšar et al., 2013). As described by Wilks (2011), this mathematical technique aims at reducing the dimensionality of a large set of data to another data set, which contains a linear combination of the original variables. The analysis can be conducted on the correlation matrix or on the covariance matrix. PCA applied to the correlation matrix weights all the standardized variables equally, because all have variance equal to the unity; instead, the analysis on the covariance matrix emphasizes the principal components having the largest variances (Wilks, 2011). Therefore, we performed PCA on the covariance matrix of **MTD** so that the information about the variance is included in the clustering of stations. PCA arranges the original dimensions of the data matrix **MTD** onto a new orthogonal space, such that the new axes are oriented in the directions explaining largest variance in the data. These new directions are called principal components and they are chosen in such a way that the greatest variance of the data lies along the first direction (namely, the first principal component), the second greatest variance on the second direction, and so on. The principal components correspond to the eigenvectors of the covariance matrix of **MTD**, while the eigenvalues are a proxy of the variance explained along each principal direction. It follows that, ordering the eigenvalues in descending order from largest to smallest, it is likely that the subspace mapped by the first m principal directions explains most of the variability of the data contained in the **MTD** matrix. That is, it is sufficient to consider this m-dimensional subspace to describe the main features of the original N_S-dimensional space. The quality of the description provided by the m-th subspace can be assessed by comparing the sum of the m eigenvalues – corresponding to the m eigenvectors considered – and the cumulated variance explained by all the eigenvectors, computed as the sum of all eigenvalues. Typically, in the present application the first two principal components (i.e., m=2) were sufficient to describe the thermal behavior of the stations, allowing to cluster them on a simple plane. As a consequence, the interpretation of the analysis is straightforward and objective.

3. Case study: Turin (Italy)

3.1. Stations and data

Turin is located in the North-West region of Italy, at latitude 45.071 N and longitude 7.687 E. The metropolitan area of Turin has a population of almost 1.5 million inhabitants, covering an area of about 600 km^2. The city is at about 100 km (air distance) far from the highest peak of the Alps, at a mean elevation above the sea level of 250 meters and it is surrounded by hills up to 600 m high in the Eastern sector, as shown in Fig. 1a.

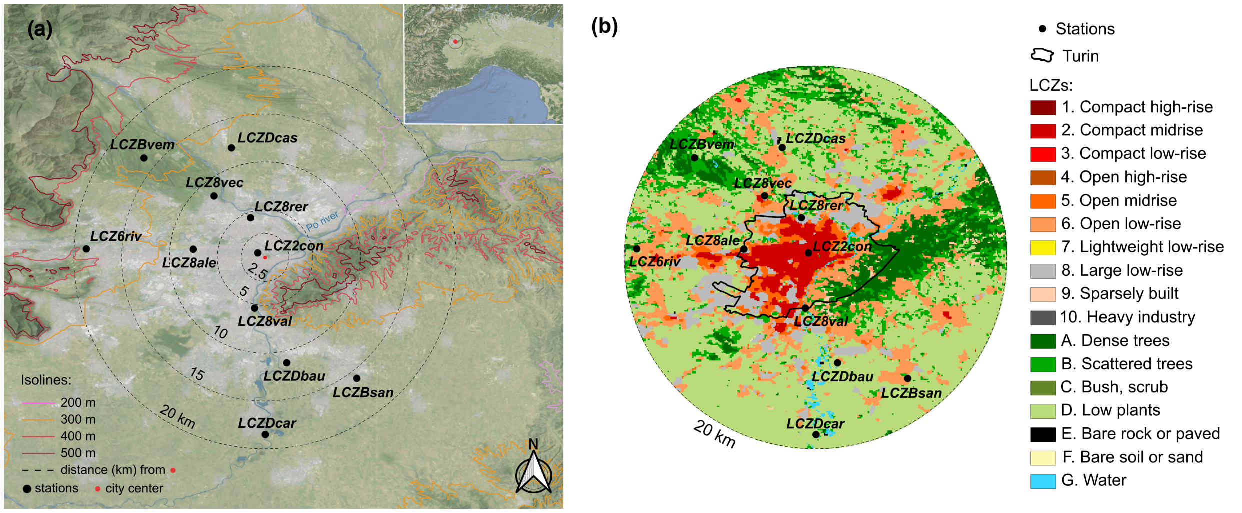

Figure 1: Panel **(a)**: terrain map of the metropolitan area of Turin (North-West of Italy, in the inset), with the principal rivers highlighted in blue and the urbanized area represented in gray. The distance in kilometers of each weather station (black dots) from the city center (red dot, Piazza Castello: lat 45.071 N, lon 7.687 E) is marked with the black dashed isolines, while the colored continuous isolines indicate the elevation above the sea level (meters). Panel **(b)**: LCZ map of the studied area, from Demuzere et al. (2020) (WUDAPT database, Ching et al., 2018).

The Po river flows in the South-East of the city and separates the most urbanized area, which is mainly located on the western bank of the river, from the hills in the East (see Fig. 1a). The Köppen Climate Classification (Köppen and Geiger, 1936) puts Turin into the Humid Subtropical Climate, namely Cfa (C = warm temperature, f = fully humid, a = hot summer). According to this, the climate in Turin is warm and temperate with significant rainfall all over the year. The 11 stations considered in the analysis (see Tab. 1) provide hourly near-surface temperature data. They belong to the network of the Regional Agency for the Protection of the Environment of Piedmont Region (Arpa Piemonte) and are distributed around the city of Turin, with about 20km as maximum distance from the city center (see Fig. 1a). For this study, the selected stations are chosen on the basis of the longest temporal series available, from January 1^{st}, 2007 to December 31^{st}, 2020 (14 years).

According to the Local Climate Zones (LCZ) map (see Fig. 1b), provided for Europe at 100m spatial resolution by Demuzere et al. (2019, 2020) (WUDAPT database, Ching et al., 2018), six stations fall under the "Built types" category of Stewart and Oke (2012): Consolata (*LCZ2con*), Rivoli (*LCZ6riv*), Alenia (*LCZ8ale*), Vallere (*LCZ8val*), Venaria Ceronda (*LCZ8vec*) and Reiss Romoli (*LCZ8rer*). The stations of Santena-Banna (*LCZBsan*), Venaria La Mandria (*LCZBvem*), Bauducchi (*LCZDbau*), Carmagnola (*LCZDcar*) and Caselle

Table 1: Weather stations for the temperature measurements used in the analysis, sorted alphabetically by their short names, with lat-lon coordinates in decimal degrees, elevation above the sea level (a.s.l., meters) and Local Climate Zones types and definitions (from Stewart and Oke, 2012).

Station name	Station short name	Lat (N)	Lon (E)	Elevation (m a.s.l.)	LCZ type and definition
Consolata	*LCZ2con*	45.0758	7.6783	290	2: Compact midrise
Rivoli	*LCZ6riv*	45.0800	7.4989	362	6: Open low-rise
Alenia	*LCZ8ale*	45.0797	7.6108	320	8: Large low-rise
Vallere	*LCZ8val*	45.0181	7.6750	239	8: Large low-rise
Venaria Ceronda	*LCZ8vec*	45.1353	7.6325	253	8: Large low-rise
Reiss Romoli	*LCZ8rer*	45.1125	7.6708	270	8: Large low-rise
Santena-Banna	*LCZBsan*	44.9447	7.7819	238	B: Scattered trees
Venaria La Mandria	*LCZBvem*	45.1750	7.5592	337	B: Scattered trees
Bauducchi	*LCZDbau*	44.9610	7.7086	226	D: Low plants
Carmagnola	*LCZDcar*	44.8861	7.6861	232	D: Low plants
Caselle	*LCZDcas*	45.1856	7.6508	300	D: Low plants

(*LCZDcas*) are categorized as "Land cover types" (Stewart and Oke, 2012), as shown in Tab. 1.

However, the LCZ classification is sometimes the result of unsupervised choices and may lead to assignments which not always match the analysts' expertise. In the following section, the stations are reexamined in light of the MTD approach. It emerges that sometimes the LCZs are too local to fully characterize the thermal behavior of the sites and do not consider the effects induced by local surrounding conditions or the large-scale context around a station, e.g., the distance to the city center or the proximity to the reliefs.

3.2. Results: Urban Heat Island

The MTD patterns obtained from Eq. (1) are shown in Fig. 2. The emerging patterns already feature common behavior among groups of stations at a first glimpse. Firstly, three stations – Consolata (*LCZ2con*), Reiss Romoli (*LCZ8rer*) and Alenia (*LCZ8ale*) – show a central cold area, characterized by negative MTD values during daytime, while at night hours the temperature anomalies are positive (see the first row of Fig. 2). Secondly, the stations Caselle (*LCZDcas*), Rivoli (*LCZ6riv*) and Vallere (*LCZ8val*), displayed in the second row of Fig. 2, do not show a well defined hot/cold blob. Finally, in the remaining panels of Venaria Ceronda (*LCZ8vec*), Venaria La Mandria (*LCZBvem*), Bauducchi (*LCZDbau*), Santena-Banna (*LCZBsan*) and Carmagnola (*LCZDcar*) an inverse pattern clearly emerges, characterized by positive MTD values during daytime and negative ones during night times.

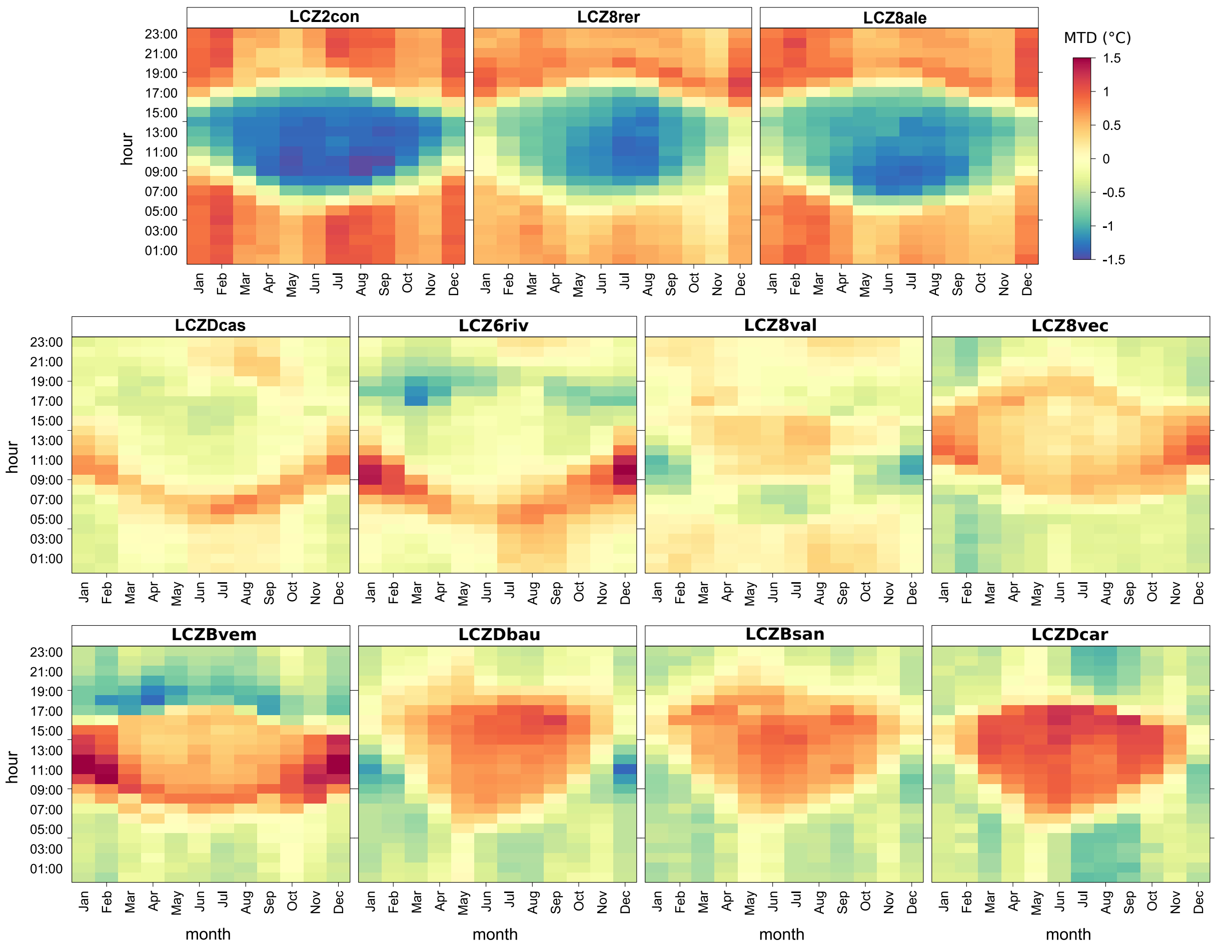

Figure 2: Mean Temperature Difference (MTD): each panel, labeled with the short name associated to each weather station S, represents the $\mathrm{MTD}^{S}_{i,M}$ computed with Eq. 1 (hours i are reported on the y-axis and months M on the x-axis).

Figure 3 shows the results of the applications of the Principal Component Analysis to the **MTD** matrix, allowing the different behaviors of the stations to be distinguished. As described in Section 2, the first two principal components clearly emerge. In Fig. 3a the percentage of the explained variance is plotted on the left y-axis, while its cumulative values are represented as the right ordinate. The first two principal components (p.c.) explain most of the variance in the data (about $\simeq 92\%$) and so they are sufficient to cluster the stations: in particular, the first p.c. accounts for $\simeq 74\%$, while the second p.c. for about 18%. The projection of each station onto the first and the second principal components are reported along the x- and y-axis of Fig. 3b, respectively.

The physical meaning of these principal components is clear looking at panels (c) and (d) of Fig. 3, showing the two signals. Let us focus on the signal described by the first principal component (Fig. 3c). It is characterized by negative temperature anomalies during daytime hours and positive ones during evening

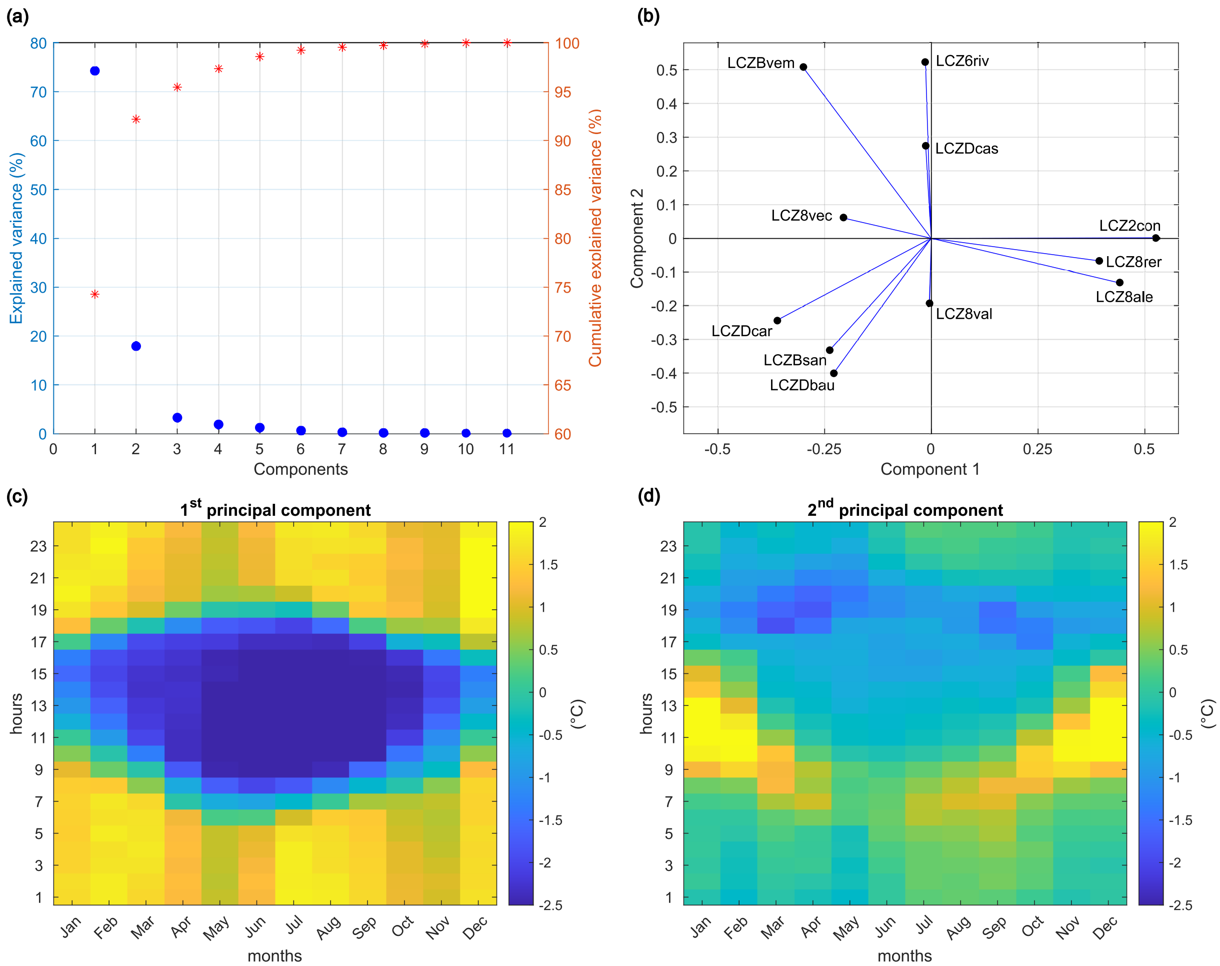

Figure 3: PCA on the matrix **MTD**: **(a)** percentage of explained variance for each of the 11 components (blue dots, left axis) and their cumulative values (red asterisks, right axis); **(b)** space of the first (x-axis) and second (y-axis) principal components; **(c)**,**(d)** representation of the two principal components of the MTD (color scales in degrees Celsius).

and night. This is the typical pattern embedded in the Urban Heat Island phenomenon: in the morning and early-afternoon, the UHI is low and can even become negative in some cases (Memon et al., 2009), resulting in the so-called daytime Urban Cool Island (Theeuwes et al., 2015). Then, when the solar radiation decreases, the urban area retains more heat and cools more slowly than the rural surroundings (Theeuwes et al., 2017), resulting in positive anomalies of temperatures. We deduce from this pattern that the first principal component – which corresponds to the highest eigenvalue ($\simeq 74\%$ of explained variance) – refers to the most evident characteristic differentiating the stations: urban vs. rural thermal behavior. The projection of each station onto the first principal component (x-axis of Fig. 3b), either it is positive or negative, determines whether a site is characterized by one or the other thermal behavior: the stations of Consolata (*LCZ2con*), Reiss Romoli (*LCZ8rer*) and Alenia (*LCZ8ale*) are characterized by substantially positive values, while negative projections correspond to Venaria La Mandria (*LCZBvem*), Venaria Ceronda (*LCZ8vec*), Carmagnola (*LCZDcar*), Santena-Banna (*LCZBsan*) and Bauducchi (*LCZDbau*). Instead, an almost null projection onto the first p.c. means that the thermal behavior cannot be assigned to urban or rural patterns. This is the case of Rivoli (*LCZ6riv*), Caselle (*LCZDcas*) and Vallere (*LCZ8val*).

The general pattern described by the first p.c. (Fig. 3c) can be made station-specific by multiplying it by the projection of the station of interest onto the first principal component, thanks to the fact that the principal components form an orthonormal base. As an example, consider two stations characterized by a positive and a negative projection onto the first p.c., namely Consolata (*LCZ2con*, positive) and Carmagnola (*LCZDcar*, negative), and examine one temporal slot, e.g., 01UTC in January. We define $E_{i,M}^{PC}$ as the value corresponding to the considered instant in time (hour $i = 1$ and month $M = Jan$), derived from the representation of the first principal component ($PC = 1$) in Fig. 3c: $E_{1,Jan}^{PC=1} = 1.63$. The value of the projection of Consolata onto the first principal component is 0.54 (see Fig. 3b). The MTD associated with this station is $\mathrm{MTD}_{1,Jan}^{LCZ2con} = 0.90°\mathrm{C}$ (see panel *LCZ2con* in Fig. 2), and it can be obtained by multiplying $E_{1,Jan}^{1}$ by 0.54: $\mathrm{MTD}_{1,Jan}^{LCZ2con} \simeq 1.63 \cdot 0.54 = 0.88°\mathrm{C}$. The centesimal digits missing for obtaining the exact value 0.90°C (in Fig. 2) derive from the additional contribution of the other principal components. To sum up, at 01UTC in January the temperature at Consolata is $\simeq 0.90°\mathrm{C}$ warmer than the average of all other stations. This behavior reflects the UHI effect and it is mainly due to the urban characteristics of this site, which is confirmed also by its LCZ class. Instead, the station of Carmagnola is characterized by a negative projection onto the first principal component, equal to -0.35 (Fig. 3b): $\mathrm{MTD}_{1,Jan}^{LCZDcar} \simeq 1.63 \cdot (-0.35) = -0.57°\mathrm{C}$. As above, a good approximation of the actual $\mathrm{MTD}_{1,Jan}^{LCZDcar} = -0.38°\mathrm{C}$ displayed in Fig. 2 would be obtained adding $E_{1,Jan}^{2} \cdot (-0.25)$, where (-0.25) is the projection of Carmagnola onto the second principal component (Fig. 3b).

The previous example shows that the correspondence between the projection of the stations onto the first principal component and their thermal pattern is very clear. A first group, characterized by positive projections (*LCZ2con*, *LCZ8rer* and *LCZ8ale*), resembles exactly the signal shown in Fig. 3c, where warmer temperatures are experienced during night. Therefore, its thermal behavior is associated with a typical urban pattern in the UHI effect, as already discussed by Milelli (2016) and Garbero et al. (2021). The LCZ-based classification assigned to the stations belonging to this group confirms these findings, since the combined effect of buildings and the mostly paved surface cover – typical of LCZ classes number 2 and 8 – greatly influences the surface energy and radiation balance (Oke, 1982).

A second group of stations exhibits the opposite temperature pattern with respect to the one shown in Fig. 3c, having a negative projection onto the first p.c.: *LCZBvem*, *LCZ8vec*, *LCZDcar*, *LCZBsan* and

LCZDbau. We associate this thermal pattern with the rural surroundings of the city, characterized by colder temperatures during the night and by higher early morning heating rate than over the city (Johnson, 1985; Theeuwes et al., 2015). As before, some considerations can be drawn in light of the LCZs assignment. In agreement with the authors' expertise, the described rural thermal pattern is coherent with the land cover types B (scattered trees) and D (low plants) associated with the stations of Santena-Banna (*LCZBsan*), Venaria La Mandria (*LCZBvem*), Bauducchi (*LCZDbau*) and Carmagnola (*LCZDcar*). Note that *LCZBsan*, *LCZDbau* and *LCZDcar* were adopted as rural stations also in Milelli (2016) and Garbero et al. (2021). However, the LCZs assignment of Venaria Ceronda (*LCZ8vec*), namely the "Large low-rise" built type, seems too local to fully characterize its rural thermal pattern emerging from PCA. Actually, the LCZ class number 8 would relate this station to a mostly paved surface with few or no trees, but the PCA shows that its thermal behavior is instead more aligned with the rural class.

Figure 3b also shows that the projection onto the first principal component is almost zero for *LCZ6riv*, *LCZDcas* and *LCZ8val*. This means that for these three stations the contribution of the first principal component ($E_{i,M}^{PC=1}$) weights less than the second p.c. and, therefore, the signal characterizing these sites looks like panel (d) of Fig. 3. Before focusing on the meaning behind the second principal component, the near-absence of the projection onto the first principal component in *LCZ6riv*, *LCZDcas* and *LCZ8val* reveals that Rivoli, Caselle and Vallere exhibit an intermediate thermal behavior with respect to the other stations characterized by substantially positive or negative projections. Note that this intermediate behavior is not necessarily homogeneous among these stations. PCA only highlights that their thermal pattern differs from all other sites characterized by positive or negative projections and, therefore, these three stations should be considered carefully for UHI studies. It is also important to point out that also the LCZs assignment for *LCZ6riv*, *LCZDcas* and *LCZ8val* is questionable and appears too local to take into account the real conditions affecting the temperatures measured by the sensors. The site of Rivoli (*LCZ6riv*) is designated as "Open low-rise" built type, but its thermal behavior is not classified as urban by the PCA probably because of the proximity to the reliefs and the distance from the city of Turin. Caselle station (*LCZDcas*) is assigned to the "low plants" land cover type D, but the PCA does not classifies it as rural. Actually, *LCZDcas* is located in the perimeter area of an airport, at about 200 m from the airstrip and only 600 m far from the closest town and the effective presence of low plants is true in the immediate surroundings of the site only. Finally, the "Large low-rise" built type associated with the station of Vallere (*LCZ8val*) does not show an urban thermal behavior, because it is located close to a park. Therefore, the LCZs assignments are not always able to clearly distinguish the urban and rural thermal patterns, since we observed that same LCZ type corresponds in some cases to different thermal behaviors.

Let now focus on the signal described by the second principal component (Fig. 3d), which explains about the 18% of the total variance and captures other aspects (with respect to the first component) related to the stations. This signal is characterized by positive anomalies of temperature after sunrise and negative ones when the solar radiation decreases. We note that southern sites, such as Vallere (*LCZ8val*), Bauducchi (*LCZDbau*), Santena-Banna (*LCZBsan*) and Carmagnola (*LCZDcar*), are characterized by a negative projection onto the second principal component, meaning that these stations experience a later warming in the morning and an earlier cooling in the evenings. This behavior can be ascribed to the different thermal regime existing between the northern and southern portion of the considered area. In the North, the stations are closest to the reliefs and more subject to ventilation, while in the South their location in the Po valley yields to a more frequent inversion in the usual vertical temperature gradient. The colder air near the ground induces a delay in the warming up in the morning and an earlier cooling down in the evening, and it is associated with foggy conditions, as frequently observed in that area (Cassardo et al., 2002). In particular, the Southern stations registered a mean (over the 14 years of the analysis) of 88 foggy days/year, while the Northern ones 17 days/year only (Arpa Piemonte, 2020). Therefore, we suggest that the second principal component is related to the geographical position of the stations, mainly to their elevation – and so proximity to reliefs – and latitude. The high correlation between the projection onto the second principal component and (i) the elevation of the sites (Pearson's coefficient of correlation $\simeq 0.86$) and (ii) their latitude (correlation $\simeq 0.75$) supports our hypotheses.

4. Applicability of MTD

The Urban Heat Island intensity varies from city to city and its quantification is largely affected not only by the geography and climate of the site, but also by the datasets available to researchers. In this context, the proposed method is conceived to be general and applicable even when the data are not as rich as it is for the case study analyzed here.

The first matter to address is the temporal scale of the UHI analysis. In the last decades, different scales were focused on, ranging from climatic scales (e.g., Rosenzweig et al., 2005; Parker, 2010) to seasonal analyses – e.g., summer heat waves (Founda and Santamouris, 2017) or waves in winter months (Giridharan and Kolokotroni, 2009) – to the negative effects of UHI during night hours in health and welfare studies (Tan et al., 2010). There is no single choice, but the selection of the time scale has to agree with the aims of the specific Urban Heat Island study of interest.

If there is no particular purpose other than the characterization of the thermal behavior of stations, the annual time scale represents the most appropriate choice to fully grasp the thermal pattern. In any case, one of the main advantages of the proposed MTD is to be as general as possible and, therefore, adaptable to any temporal scale of interest.

Once the time scale has been chosen – we considered the annual scale in the Turin case study described in the previous section – a second question concerns the duration of the available measures. In order to test this aspect, we applied the MTD method by increasingly reducing our range of data (i.e., 14 years, 13 years, and so on) and we observed that only one year of observations is enough for the MTD to work. In Appendix A we show that the first principal component of the PCA exhibits the same thermal pattern associated to an urban or rural behavior as in Section 3.2. It follows that the metric MTD appears capable of exploiting the data very effectively, even if obviously the longer the period of observations is, the more the results will not be affected by the particular conditions observed in the considered 12 months.

The time resolution of measures is another key aspect. In the Turin test case, we adopted the hourly time step, which is one of the most widely used in UHI literature (Santamouris, 2007; Oh et al., 2020; Kim and Brown, 2021). However, the robustness of the MTD has been tested also against a coarser temporal resolution: by considering a 3 hour time step (e.g., Pakarnseree et al., 2018, used this sampling time). In this study, we consider a subset of our original data with temperature measurements at 00, 03, 06, 09, 12, 15, 18 and 21 UTC. Again, the method proves to work very well, since the resulting clustering of stations is equal to that obtained with the hourly temperatures (results are shown in Appendix A).

An important question about the applicability of the MTD concerns the minimum number of weather stations required for the method to work. Turin has a relatively consistent number of measurement sites, but this may not be the case for other cities. In order to test this point, we performed a detailed sensitivity analysis, by re-evaluating the MTD performances using different subsets of the original 11 stations (see Appendix A). By excluding the stations identified as rural by our method (in Section 3.2), the PCA still identifies an urban thermal behavior and a different one. On the contrary, when considering the rural stations only, the main pattern described by the signal is different and cannot be related to urbanity/rurality. This result is a warning that the considered stations are not a good choice in selecting urban/rural pairs for UHI.

Finally, we evaluated the method on a different dataset and geographic domain: the city of Cuneo, in North-Western Italy. Cuneo is located at a higher mean elevation above the sea level (about 550 m a.s.l.), has a smaller number of inhabitants than Turin (about 60000) and has only two weather stations available, namely Cuneo Camera di Commercio *LCZ2ccc* and Cuneo Cascina Vecchia *LCZ6ccv* (see Appendix A). In the surroundings of the city, the station of Boves *LCZDbov* (575 m a.s.l.) is the only one suitable for this kind of analysis. In addition, the hourly temperatures are available for 18 months only (from July 2019 to December 2020). The Köppen Climate Classification (Köppen and Geiger, 1936) puts Cuneo into the

Temperate Oceanic Climate, namely Cfb. Being at the foot of the Alps, Cuneo receives more snow during winter than Turin (Arpa Piemonte, 2020). Even in this completely different domain, the MTD works very well (see Appendix A): *LCZ2ccc* is deemed as urban, *LCZDbov* as rural and *LCZ6ccv* exhibits an intermediate behavior between the other two.

The minimum number of stations for the application of our method is two, namely the intrinsic number to the Urban Heat Island definition, provided that the selected sites exhibit a different thermal behavior (highlighted in the signal of the PCA) in terms of urbanity or rurality.

5. Discussion and Conclusions

The proposed method aims to cluster common behaviors among the available measurement stations, in order to detect the most representative urban-rural pairs for Urban Heat Island quantification in the studied area. The example of Turin shows that the MTD turns out to be an effective metric able to grasp the main differences – in terms of thermal behavior – among the stations.

Given the widely recognized difficulty of the proper selection of urban-rural pairs, the metric which we propose can complement the methods already existing in literature, and provides an additional tool in the UHI research topic for the landscape classification. In this line, it is instructive to compare the results (for the Turin area) of our approach with those of three consolidated methods: (i) the Day-to-Day variation introduced by Karl et al. (1995) and further developed by Gough (2008), (ii) the mean daily excursion described by Milelli (2016) and (iii) the ratio between warm and cold day transitions recently presented by Gough (2020).

The Day-To-Day (DTD) temperature variation detects urban stations when a site exhibits increasing day-to-day variation in the daytime maximum temperature. Figures 4a-b show the results of this metric: DTD is evaluated as the absolute difference between the temperatures of adjacent days for a given period of time (e.g., month) and is calculated both for daily temperature minimum (nighttime, DTD T_{min}) and daily temperature maximum (daytime, DTD T_{max}). According to Oke (1981, 1982), urban stations exhibit lower nocturnal temperature variability because urbanized areas trap the radiative energy, inducing a slower convective heat loss than the surrounding rural areas. Therefore, the effects of urbanization are associated with the lowest DTD T_{min}. Results shown in Fig. 4a highlight a first cluster corresponding to the three urban stations – i.e., Consolata (*LCZ2con*), Reiss Romoli (*LCZ8rer*) and Alenia (*LCZ8ale*) – and this classification is consistent with what we found by the proposed MTD method. The DTD metric for T_{min} identifies a second cluster characterized by higher values of the day-to-day variation, but it is quite difficult to separate possible intermediate behaviors, at least in an objective way. It follows that the DTD method classifies all other stations – Venaria La Mandria (*LCZBvem*), Venaria Ceronda (*LCZ8vec*), Carmagnola (*LCZDcar*), Santena-Banna

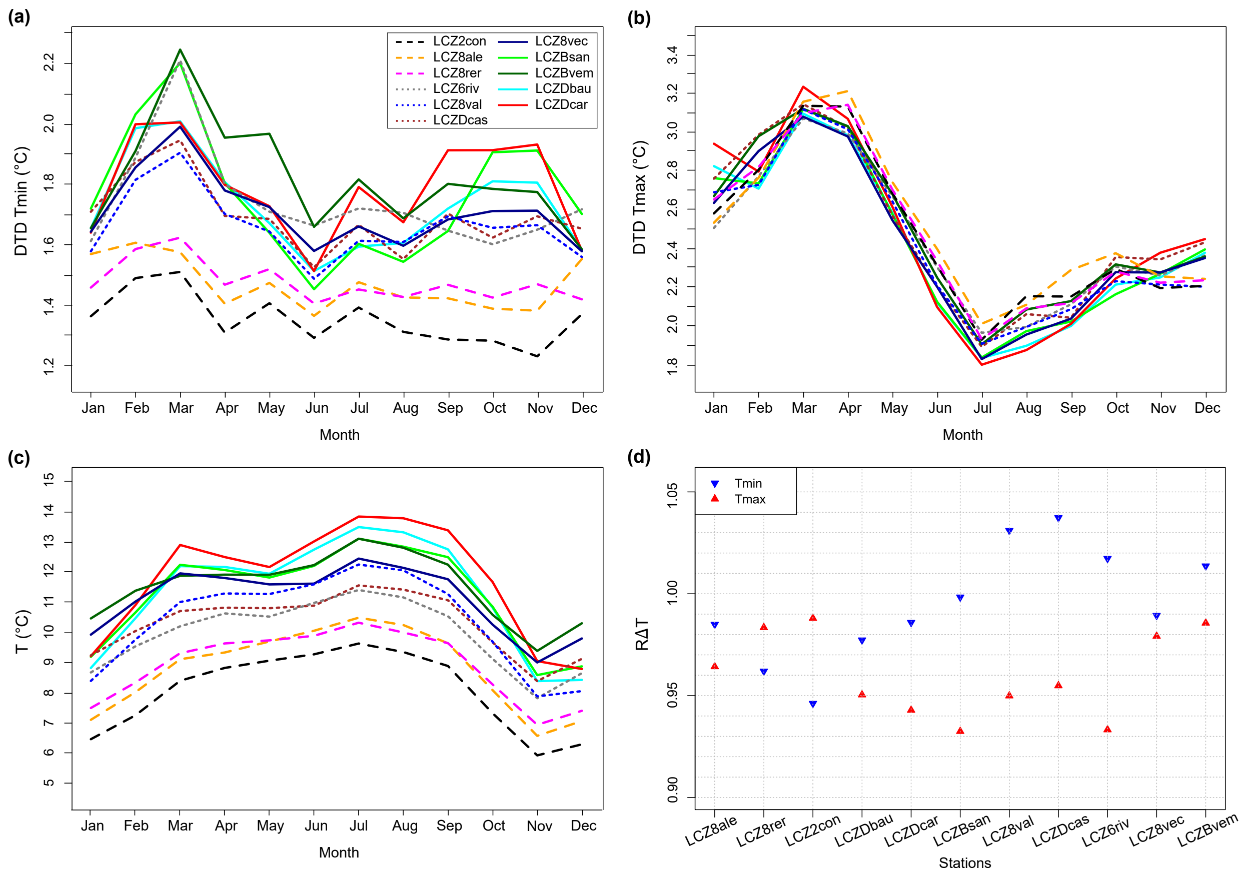

Figure 4: Panels **(a)** and **(b)**: Day-To-Day (DTD), i.e., the average monthly DTD variation of nighttime ($T_{\min}$ in **(a)**) and day ($T_{\max}$ in **(b)**) temperatures (Anderson et al., 2018). The different line styles refer to the thermal behaviors characterizing the stations, obtained through the Mean Temperature Difference (MTD). The dashed lines correspond to stations with positive projection onto the first principal component of the PCA, associated with an urban thermal pattern; the continuous lines refer to stations which projection onto the first p.c. is negative (rural thermal pattern); the dotted lines correspond to the stations characterized by an almost null projection onto the first p.c. Panel **(c)**: mean daily excursion of temperature, i.e., the monthly average of $T_{\max} - T_{\min}$ (Milelli, 2016); this panel refers to the same legend reported in **(a)**. Panel **(d)**: warm to cold transition ratio (RΔT) for the minimum (downwards blue triangles) and maximum (upwards red triangles) temperature of the day (Gough, 2020).

(*LCZBsan*), Bauducchi (*LCZDbau*), Rivoli (*LCZ6riv*), Caselle (*LCZDcas*) and Vallere (*LCZ8val*) – as rural and, differently from our MTD approach, seems unable to grasp the intermediate behavior. As described in Anderson et al. (2018), even smaller differences between the sites emerge when we consider DTD $T_{\max}$ (see Fig. 4b).

The second term of comparison we consider is the mean daily excursion proposed by Milelli (2016), calculated as the difference between the monthly averaged maximum and minimum temperatures (Fig. 4c).

Here, the three stations Consolata (*LCZ2con*), Reiss Romoli (*LCZ8rer*) and Alenia (*LCZ8ale*) are clearly marked by a limited daily excursion, indicating a non-sufficient cooling during the night and therefore – according to the UHI definition – they are related to a urban landscape. This is in agreement with what detected with the MTD metric. However, even in this case a non clearly distinguished group of stations shows an intermediate behavior, i.e., a gradual transition from the low to the high daily excursion groups emerges. See for example the stations of Caselle (*LCZDcas*), Rivoli (*LCZ6riv*), Vallere (*LCZ8val*) and Venaria Ceronda (*LCZ8vec*) in Fig. 4c.

Finally, the application of the ratio between the warm and cold transitions is shown in Fig. 4d. By considering Canadian temperatures, Gough (2020) found a metric sensitive to what he called "peri-urban" landscapes, in particular focusing on the warm to cold transition ratio, RΔT, calculated for minimum temperatures $T_{\min}$. Figure 4c illustrates the results both for $T_{\min}$ and $T_{\max}$. Gough (2020) identified the threshold for $T_{\min}$ (i.e., RΔT= 1.05) above which a group of stations is deemed peri-urban. If we adopt this threshold, no station falls above this limit; therefore, in the case of Turin, the value RΔT= 1.05 appears not to be adequate to detect intermediate thermal behaviors. This is not surprising because of the different climate in Canada. Using the outcomes of our MTD approach, a new *ad hoc* threshold equal to 1.016 would allow one to classify Vallere (*LCZ8val*), Caselle (*LCZDcas*) and Rivoli (*LCZ6riv*) in a different thermal behavior, which we call intermediate since it differs from the urban and rural but has no internal coherence. However, a slightly different value (lower or higher than 1.016) would provide very different results: e.g., if RΔT= 1.01 also the station of Venaria la Mandria (*LCZBvem*) would pertain to the intermediate behavior, while for RΔT= 1.02 the only stations with an intermediate pattern would be Vallere (*LCZ8val*) and Caselle (*LCZDcas*). Note that RΔT ($T_{\min}$) for the remaining stations – ascertained as urban (Alenia (*LCZ8ale*), Reiss Romoli (*LCZ8rer*) and Consolata (*LCZ2con*)) and rural (Bauducchi (*LCZDbau*), Carmagnola (*LCZDcar*) and Santena-Banna (*LCZBsan*)) according to the MTD – form two well separated groups and therefore clearly differentiate from the intermediate landscapes, as in Gough (2020).

In a nutshell, the consolidated methodologies, when are applied to the Turin area, agree on the classification of the urban and rural stations, and identify the same urban-rural sites detected by our Mean Temperature Difference method. However, likely due to the complexity of the Turin landscape, the attribution of intermediate thermal behaviors is not straightforward and the consolidated methods seem not to be able to give an objective and unique characterization of this pattern. In contrast, the proposed Mean Temperature Difference seems to be suitable in this area: in light of the Principal Component Analysis, the three stations Rivoli (*LCZ6riv*), Caselle (*LCZDcas*) and Vallere (*LCZ8val*) are characterized by near-zero projections onto the first principal component. In this way, the subjectivity is minimized, since no thresholds or graphic interpretations are needed, differently from the other methods existing in literature. In fact, since the first principal component is associated with the urbanity or rurality of a site, a missing projection onto this component implies a thermal behavior which is neither clearly urban or rural, but rather an intermediate one which is not necessarily characterized by an internal coherence. This observation is also confirmed by the different Local Climate Zones associated to the stations *LCZ6riv*, *LCZDcas* and *LCZ8val*. Their thermal behavior is not captured by the first principal component, but it is synthesized by the second p.c. emerging from the PCA, which is associated to other geographically-based features characterizing the stations.

To summarize, the combined use of MTD metric and PCA represents a robust tool to characterize the sites in the Urban Heat Island context. The method has been proven (i) to well reproduce the thermal behavior of the metropolitan area of Turin, (ii) to agree with existing and widely validated methods for the distinction between the urban and rural stations, and (iii) to be easily interpretable.

Aware of the impossibility to totally eliminate some kind of subjectivity in the task of selecting urban-rural pairs, we aim at providing an additional tool to discern the landscape categories for the Urban Heat Island quantification. The metric which we suggest can be combined with the existing methods, especially when the study area does not offer a trivial categorization into urban or rural stations.

Acknowledgments

The authors thank the Regional Agency for the Protection of the Environment of Piedmont Region (Arpa Piemonte) for the data used in this paper. Barbara Cagnazzi and Daniele Gandini are acknowledged for having provided the fog statistics. This work is funded by the RISK-GEST Project-PITEM RISK, Interreg 2014-2020 Alcotra IT-FR, the MISTRAL 2017-IT-IA-0144 Program Connecting Europe Facility (CEF) and the 2019-2021 Agreement between National Department of Civil Protection and Arpa Piemonte.

Appendix A. Results of the applicability of the MTD

Sensitivity analysis to evaluate to which extent the MTD method is applicable (attached file).

References

Anderson, C.I., Gough, W.A., Mohsin, T., 2018. Characterization of the urban heat island at Toronto: Revisiting the choice of rural sites using a measure of day-to-day variation. Urban Climate 25, 187–195.

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. International Journal of Climatology: a Journal of the Royal Meteorological Society 23, 1–26.

Arpa Piemonte, 2020. Annual Climatic Report (in Italian). https://www.arpa.piemonte.it/rischinaturali/tematismi/clima/rapporti-di-analisi/annuale.html. [Online; accessed 22-November-2021].

Cassardo, C., Forza, R., Manfrin, M., Longhetto, A., Qian, M., Richiardone, R., Balsamo, G., et al., 2002. The Urban Meteorological Station of Turin, in: 11ˆ th Symposium on Acoustic Remote Sensing, ISAC/CNR. pp. 311–320.

Ching, J., Mills, G., Bechtel, B., See, L., Feddema, J., Wang, X., Ren, C., Brousse, O., Martilli, A., Neophytou, M., et al., 2018. WUDAPT: An urban weather, climate, and environmental modeling infrastructure for the anthropocene. Bulletin of the American Meteorological Society 99, 1907–1924.

Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A.S., McLoone, S., 2013. Principal component analysis on spatial data: an overview. Annals of the Association of American Geographers 103, 106–128.

Demuzere, M., Bechtel, B., Middel, A., Mills, G., 2019. Mapping Europe into local climate zones. PloS one 14, e0214474.

Demuzere, M., Bechtel, B., Middel, A., Mills, G., 2020. European LCZ map. https://urlsand.esvalabs.com/?u=https%3A%2F%2Ffigshare.com%2Farticles%2Fdataset%2FEuropean_LCZ_map%2F13322450%2F1&e=78898b00&h=9a0f73a7&f=y&p=n. [Online; accessed 15-June-2021].

Founda, D., Santamouris, M., 2017. Synergies between Urban Heat Island and Heat Waves in Athens (Greece), during an extremely hot summer (2012). Scientific reports 7, 1–11.

Garbero, V., Milelli, M., Bucchignani, E., Mercogliano, P., Varentsov, M., Rozinkina, I., Rivin, G., Blinov, D., Wouters, H., Schulz, J.P., et al., 2021. Evaluating the urban canopy scheme TERRA_URB in the COSMO model for selected European cities. Atmosphere 12, 237.

Giridharan, R., Kolokotroni, M., 2009. Urban heat island characteristics in London during winter. Solar Energy 83, 1668–1682.

Gough, W., 2008. Theoretical considerations of day-to-day temperature variability applied to Toronto and Calgary, Canada data. Theoretical and Applied Climatology 94, 97–105.

Gough, W.A., 2020. Thermal signatures of peri-urban landscapes. Journal of Applied Meteorology and Climatology 59, 1443–1452.

Hannachi, A., Jolliffe, I.T., Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: A review. International Journal of Climatology: A Journal of the Royal Meteorological Society 27, 1119–1152.

Hoffmann, P., Schlünzen, K.H., 2013. Weather pattern classification to represent the urban heat island in present and future climate. Journal of Applied Meteorology and Climatology 52, 2699–2714.

Johnson, D., 1985. Urban modification of diurnal temperature cycles in Birmingham, UK. Journal of Climatology 5, 221–225.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer New York.

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374, 20150202.

Karl, T.R., Knight, R.W., Plummer, N., 1995. Trends in high-frequency climate variability in the twentieth century. Nature 377, 217–220.

Kim, S.W., Brown, R.D., 2021. Urban heat island (UHI) intensity and magnitude estimations: A systematic literature review. Science of The Total Environment , 146389.

Köppen, W., Geiger, R., 1936. Das geographische System der Klimate Handbuch der Klimatologie. Ed. W. Köppen and R. Geiger 1.

Landsberg, H.E., 1981. The urban climate. Academic press.

Lorenz, E.N., 1956. Empirical orthogonal functions and statistical weather prediction .

Memon, R.A., Leung, D.Y., Liu, C.H., 2009. An investigation of urban heat island intensity (UHII) as an indicator of urban heating. Atmospheric Research 94, 491–500.

Milelli, M., 2016. Urban heat island effects over Torino. COSMO Newsletter 16, 1–10.

Mohsin, T., Gough, W.A., 2012. Characterization and estimation of urban heat island at Toronto: impact of the choice of rural sites. Theoretical and Applied Climatology 108, 105–117.

Oh, J.W., Ngarambe, J., Duhirwe, P.N., Yun, G.Y., Santamouris, M., 2020. Using deep-learning to forecast the magnitude and characteristics of urban heat island in Seoul Korea. Scientific reports 10, 1–13.

Oke, T.R., 1973. City size and the urban heat island. Atmospheric Environment (1967) 7, 769–779.

Oke, T.R., 1976. The distinction between canopy and boundary-layer urban heat islands. Atmosphere 14, 268–277.

Oke, T.R., 1981. Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. Journal of climatology 1, 237–254.

Oke, T.R., 1982. The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society 108, 1–24.

Pakarnseree, R., Chunkao, K., Bualert, S., 2018. Physical characteristics of Bangkok and its urban heat island phenomenon. Building and Environment 143, 561–569.

Parker, D.E., 2010. Urban heat island effects on estimates of observed climate change. Wiley Interdisciplinary Reviews: Climate Change 1, 123–133.

Rizwan, A.M., Dennis, L.Y., Chunho, L., 2008. A review on the generation, determination and mitigation of Urban Heat Island. Journal of environmental sciences 20, 120–128.

Rosenzweig, C., Solecki, W.D., Parshall, L., Chopping, M., Pope, G., Goldberg, R., 2005. Characterizing the urban heat island in current and future climates in New Jersey. Global Environmental Change Part B: Environmental Hazards 6, 51–62.

Santamouris, M., 2007. Heat island research in Europe: the state of the art. Advances in building energy research 1, 123–150.

Stewart, I.D., 2007. Landscape representation and the urban-rural dichotomy in empirical urban heat island literature, 1950–2006. Acta Climatologica et Chorologica 40, 111–121.

Stewart, I.D., 2011. A systematic review and scientific critique of methodology in modern urban heat island literature. International Journal of Climatology 31, 200–217.

Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bulletin of the American Meteorological Society 93, 1879–1900.

Tam, B.Y., Gough, W.A., Mohsin, T., 2015. The impact of urbanization and the urban heat island effect on day to day temperature variation. Urban Climate 12, 1–10.

Tan, J., Zheng, Y., Tang, X., Guo, C., Li, L., Song, G., Zhen, X., Yuan, D., Kalkstein, A.J., Li, F., et al., 2010. The urban heat island and its impact on heat waves and human health in Shanghai. International journal of biometeorology 54, 75–84.

Theeuwes, N.E., Steeneveld, G.J., Ronda, R.J., Holtslag, A.A., 2017. A diagnostic equation for the daily maximum urban heat island effect for cities in northwestern Europe. International Journal of Climatology 37, 443–454.

Theeuwes, N.E., Steeneveld, G.J., Ronda, R.J., Rotach, M.W., Holtslag, A.A., 2015. Cool city mornings by urban heat. Environmental Research Letters 10, 114022.

Tzavali, A., Paravantis, J.P., Mihalakakou, G., Fotiadi, A., Stigka, E., 2015. Urban heat island intensity: A literature review. Fresenius Environmental Bulletin 24, 4537–4554.

Wilks, D.S., 2011. Statistical methods in the atmospheric sciences. volume 100. Academic press.

Wu, F.T., Fu, C., Qian, Y., Gao, Y., Wang, S.Y., 2017. High-frequency daily temperature variability in China and its relationship to large-scale circulation. International Journal of Climatology 37, 570–582.