Remote Sensing of Environment 331 (2025) 115045

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Is satellite land surface temperature an appropriate proxy for intra-urban variability of daytime heat stress?

Ferdinand Briegel [a,b,*], Joaquim G. Pinto [a], Andreas Christen [b]

[a] *Institute of Meteorology and Climate Research Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
[b] *Chair of Environmental Meteorology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg im Breisgau, Germany*

ARTICLE INFO

Edited by Marie Weiss

Keywords:
Landsat Level-2
Land surface temperature
Outdoor thermal comfort
Urban climate
Urban heat
UTCI
Machine learning

ABSTRACT

Adaptation of urban areas to heat extremes requires adequate information on intra-urban variability patterns of outdoor thermal comfort (OTC). Remotely sensed Land Surface Temperatures (LST) are often used to map heat hotspots in urban areas. However, this approach has limitations as LST and OTC are influenced by different physical processes. This study investigates the relationship between satellite-derived Landsat Level-2 LST data and pedestrian-level Universal Thermal Climate Index (UTCI) predictions from a microscale thermal comfort model across Freiburg, Germany. A cluster analysis of the differences is performed, and multiple random forest models are trained using different combinations of LST, ERA5-Land reanalysis, and local-specific urban morphology and land cover data as predictors.

While a linear relationship between LST and UTCI exists under non-heat stress conditions (UTCI <26 °C) and in vegetated or open areas, this becomes non-linear and spatially inconsistent under heat stress, particularly in compact urban environments. The growing divergence between LST and UTCI along an urbanization gradient ranging from −1 K to +9 K highlights the significant impact of urban morphology on the LST-UTCI relationship, leading to substantial intra-urban variability. This variability appears to persist even within similar urban typologies (e.g. LCZs/clusters), with only limited reduction in spatial variability. Random forest models confirm these findings: those based solely on LST or global-scale predictors struggle to capture intra-urban UTC variability, while models incorporating local urban morphology and land cover data outperform them (even without LST input). This suggests that the contribution of LST to neighborhood-scale UTC modeling is limited under certain conditions and environments.

1. Introduction

Urban areas are at the forefront of the challenges posed by climate change (Dodman et al., 2022). Among the most pressing issues is the growing frequency, intensity, and duration of extreme heat in many cities, which affects human well-being, critical infrastructure, and the built environment (Mora et al., 2017; Cissé et al., 2022). Addressing these challenges is essential to ensure the safety, resilience, and well-being of urban populations worldwide. However, assessing the impact of extreme heat on human well-being is inherently complex due to the multifaceted nature of urban environments (Krayenhoff et al., 2021).

One well-documented aspect of urban thermal environments is the Urban Heat Island (UHI) effect (Oke, 1973). The UHI describes differences in air or surface temperature (T_a, T_{sfc}, respectively) between urban areas and their rural surroundings, which are primarily caused by changes in land-cover, urban form, urban materials and by anthropogenic heat emissions (Oke, 1982). UHIs can be classified into four types: subsurface, surface, canopy, and boundary layer UHIs (Oke et al., 2017). Among these, the canopy layer UHI is the most relevant to human thermal comfort as it directly represents the outdoor pedestrian-level environment in which human activities take place (Oke et al., 2017). This means that information on heat exposure within the canopy layer is crucial for the adaptation and mitigation of urban areas to heat stress (Stewart et al., 2021). Nevertheless, many studies focus on the surface UHI to map urban heat patterns and potential hotspots (e.g. Zhou et al., 2019) by using land surface temperature (LST).

LST is the critical variable for studying the surface UHI (Voogt and Oke, 2003; Stewart et al., 2021). Derived from thermal infrared sensors

---

[*] Corresponding author at: Institute of Meteorology and Climate Research Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

E-mail address: ferdinand.briegel@kit.edu (F. Briegel).

https://doi.org/10.1016/j.rse.2025.115045
Received 15 April 2025; Received in revised form 18 September 2025; Accepted 22 September 2025

on satellite platforms, such as Landsat (Earth Resources Observation and Science (EROS) Center, 2020) and MODIS (Wan et al., 2021), or aircrafts, LST provides spatially extensive measures of surface thermal conditions. This fact and the global availability of satellite-based LST have made Landsat and MODIS LST products a popular tool for identifying urban thermal patterns and assessing urban heat risks at coarser scales (Peng et al., 2012; Li et al., 2023b). MODIS LST data, acquired from sensors with a wide off-nadir viewing capability, has a relatively coarse spatial resolution of around 1 km. This limits its effectiveness in capturing intra-urban variability in LST and OTC. In contrast, Landsat offers higher-resolution thermal data (100 m), making it more suitable for urban-scale analysis. However, due to its nadir-viewing geometry, Landsat is less capable of capturing off-nadir surface information. Nevertheless, Landsat LST data's higher resolution offers greater potential to assess intra-urban variability of street-level outdoor human thermal comfort (OTC). OTC, on the other hand, is influenced by a combination of T_a, humidity, wind, and the three-dimensional radiative environment expressed in mean radiant temperature (T_{mrt} -Epstein and Moran, 2006; Potchter et al., 2018). It is typically assessed using thermal indices (de Freitas and Grigorieva, 2017) such as the Universal Thermal Climate Index (UTCI - Błażejczyk et al., 2013) or Physiological Equivalent Temperature (PET - Höppe, 1999), which integrate environmental and physiological factors into an equivalent temperature which can be classified in different stress categories. Unlike LST, which captures surface conditions of roofs, tree canopies, and other (mainly) horizontally surfaces, OTC metrics are measured/modelled at pedestrian level (1.1 m a.g.l.) (Fig. 1), reflecting the actual three-dimensional conditions experienced by pedestrians, such as below the tree canopy of within narrow urban street canyons (Staiger et al., 2019). This means, that LST has limitations when used to represent canopy UHI or OTC due to the preferential nadir view of LST observation systems (Roth et al., 1989). Nazarian et al. (2022) summarize the limitations of LST from satellites for the estimation of OTC in three aspects: (i) LST only represents surfaces visible to the sensor, introducing the problem of anisotropy (Voogt and Oke, 1998; Cao et al., 2019), (ii) LST mainly maps on horizontal surfaces with high exposure to the sky (e.g. high sky view factor), and (iii) the spatial resolution of LST is too coarse to resolve individual urban structures such as buildings or trees. In addition, T_{sfc} of urban street canyon facets is strongly influenced by their height-to-width ratio and orientation, and differs significantly between rooftops and within the canyon (Adderley et al., 2015; Saher et al., 2021). As the proportion of buildings increases, therefore, satellite-based LST data products are increasingly dominated by higher T_{sfc} on rooftops, and their coarse resolution cannot account for differences within the urban canyon. In summary, satellite-based LST provides an integrated temperature signal across rooftops, streets, vegetation canopy, and other surfaces, while missing most 'vertical' and below-canopy surfaces, and thus represents an indirect and incomplete measure of human exposure to heat stress, particularly in street canyons or under tree canopies where human activity is concentrated (Christen et al., 2012; Adderley et al., 2015; Krayenhoff et al., 2020; Cheung et al., 2021; Stewart et al., 2021; Chakraborty et al., 2022; Li et al., 2023a).

Although LST is widely used to map thermal situation and potential hot/cold spots within cities (Renard et al., 2019; Zhou et al., 2019; Manoli et al., 2020; de Almeida et al., 2021), and the drawbacks of it are known (Nazarian et al., 2022), few studies combined remote-sensed LST with spatial fields of precise OTC data to investigate whether LST is a valuable proxy for OTC and whether LST values can be "corrected" or scaled to match OTC and its high intra-urban variability. The lack of high-resolution OTC data over large domains, resolving buildings and vegetation, complicates such assessments. A few studies tried to investigate OTC by temperature-humidity indices based on remote sensing and station data (Feng et al., 2020, 2023; Mijani et al., 2020; Patle and Ghuge, 2024; Fahy et al., 2025). Feng et al. (2020) proposed a modified temperature-humidity index to estimate thermal comfort in Nanjing, China, but without validation on in-situ measurements or urban climate models. In a subsequent work, Feng et al. (2023) conducted a combined analysis using the remotely sensed temperature-humidity index and station data. They evaluated the remotely sensed temperature-humidity index with the in-situ measured thermal comfort index PET and found a R^2 of 0.83. Mijani et al. (2020) developed a model to estimate the discomfort index (Thom, 1959) for Teheran, Iran, using LST and measurements from meteorological stations. In addition, Feng et al. (2023) validated the temperature-humidity index on OTC data but with a limited number of meteorological stations. However, temperature-humidity indices such as the discomfort index neglect effects of the radiative environment and wind, which play a substantial role in OTC, e.g. through ventilation, shading, or radiation trapping. Wang et al. (2020) used LST data from MODIS and weather station data to calculate UTCI, which they mapped for different cities at a resolution of 1 km. It

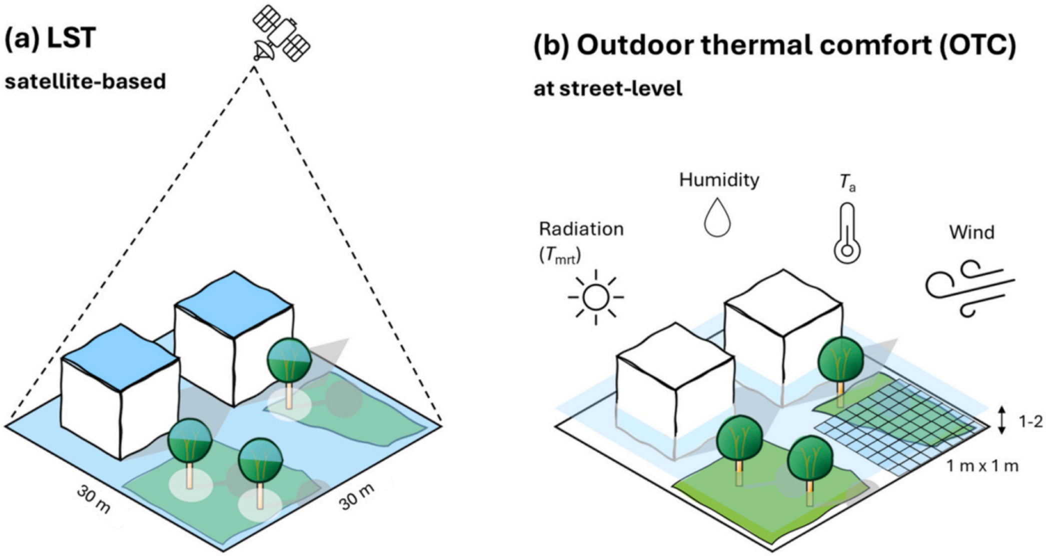

Fig. 1. Schematic representation of the surfaces representing land surface temperature (LST) (a) and outdoor human thermal comfort (OTC) (b). In (a), the blue colour represents the areas seen by a nadir-view remote sensor, such as Landsat, including all horizontal surfaces of urban areas (as well as tree canopies), as seen from above. In (b), the actual OTC at street-level is shown, where rooftops are usually excluded. Figure based on Oke et al. (2017). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

should also be noted that using MODIS to calculate UTCI at a resolution of 1 km cannot accurately represent street-level intra-urban variability of OTC. A more sophisticated study by Li et al. (2023a, 2023b) compared Landsat 8 LST data with high resolution T_{mrt} data and concluded that the relationship is particularly strong on roofs, but not on street-level. In a recent study, Fahy et al. (2025) examined the relationship between measured OTC data and LST, finding only a weak correlation between the two in urbanized areas.

The discrepancy between LST and OTC raises questions about the reliability of LST as a proxy for OTC (Voogt, 2008; Martilli et al., 2020; Venter et al., 2021; Nazarian et al., 2022; Fahy et al., 2025). However, bridging this discrepancy between LST and OTC through a better understanding of their relationship in complex urban environments and their intra-urban variability could be crucial to improve the assessment of global urban heat stress products and the delivery of Earth observation-based services and products for heat intervention and urban planning. This would be preferrable, as OTC-measurements or building-resolved OTC-modeling are complex, resource-intensive, and often unaffordable, while satellite LST products are readily available. Although there are studies in which measured or modelled OTC is compared to LST, these studies take mostly a more ‘qualitative’ approach, comparing limited measurement data from single days or stations to spatial LST data. Furthermore, these studies do not clarify how the variability of the LST-OTC relationship varies within similar urban environments, such as within the same LCZs, or between areas with similar building structures. To overcome these limitations and achieve a more ‘quantitative’ understanding of the LST-OTC-relationship across and within various diverse urban environments, we compare LST observations from Landsat Level-2 (Earth Resources Observation and Science (EROS) Center, 2020) with high-resolution OTC maps and data from a street-level station network. UTCI is used as the thermal index to quantify OTC. OTC maps are obtained by applying the HTC-NN model of Briegel et al. (2024) at 1 m spatial resolution. The HTC-NN is a machine-learning emulator of different numerical urban climate models and allows rapid prediction of high-resolution pedestrian-level UTCI. In addition, both the high-resolution OTC maps and Landsat LST data are further evaluated using data from a street-level station network.

Based on LST data and the HTC-NN model, we perform a systematic investigation between daytime LST and OTC using the city of Freiburg, a midsize town in the Upper Rhine Valley in southwest Germany, for which in-situ measurements of OTC are available from a street-level sensor network (Zeeman et al., 2024; Feigel et al., 2025; Plein et al., 2025). In particular, we examine the intra-urban variability of the LST-OTC relationship and the potential of LST to serve as a proxy for OTC. To this end, LST is compared with OTC on a spatial and temporal base. A cluster analysis is then conducted to identify areas of higher or lower agreement, and to assess the variability of the LST-OTC relationship in different urban environments. In addition, land cover classes and other urban morphometric parameters are examined to investigate how they contribute to the differences between LST and OTC. In a final step, different random forest models are developed that attempt to estimate OTC from LST, to assess whether LST can be used to directly predict UTCI, the corresponding UTCI stress classes, and what additional information/data is required to properly represent the intra-urban variability of OTC. The model results will be compared to the HTC-NN model outputs and to ERA5-HEAT data (Di Napoli et al., 2020) as baseline, which is expected to be less accurate due to its spatial resolution and missing urban parameterization. In addition, the HTC-NN and random forest model are then compared to simultaneous in-situ measurements from the street-level sensor network.

The rest of the paper is structured as follows. Section 2 ‘Data and Methods’ describes the data and methods used, followed by Section 3 ‘Results’ which presents the results of our study. Section 4 ‘Discussion’ discusses the results and section 5 concludes.

2. Data and methods

This section describes the conceptual approach, and how the LST, UTCI, and ERA5 datasets are obtained. Fig. 2 shows a flowchart of the LST-UTCI relationship analysis with UTCI modeling using the HTC-NN (step 1 / Section 2.1), cluster analysis (step 2 / Section 2.2.1), random forest modeling (step 3 / Section 2.2.2), and the evaluation procedure (step 4 / Section 2.2.3).

2.1. Research area and data

2.1.1. Research area and study period

The research area of this study is the city of Freiburg im Breisgau, a mid-sized city in the southwest of Germany. Freiburg is located in the Upper Rhine Valley and is one of the hottest cities in Germany with an annual mean T_{a} of 11.0 °C, mean precipitation of 887 mm (1990–2020), and monthly mean $T_{\mathrm{a}} > 20$ °C in July (DWD Climate Data Center (CDC), 2024). In summer, Freiburg is highly exposed to human heat stress (Briegel et al., 2024) and the Upper Rhine Valley has the highest current and projected heat stress in Germany (Hundhausen et al., 2023). The research area covers the greater area of Freiburg with an extent of 10 km × 7 km and consists of roughly 47,000 30 m × 30 m grid cells. Both the built-up and rural areas of Freiburg are studied, as we want to investigate not only urban specific relationships between LST and UTCI, but also from suburban, rural, and vegetated areas close to the city. The research area is masked by the extent of the urbanized area of Freiburg. To the east of Freiburg lies the ridge of the Black Forest which is not considered in the analysis. However, some parts of the research area cover the urbanized valley at its entrance. In addition, the lack of geospatial data availability causes some areas close to the city center to be missing.

The study period covers the years 2021 to 2024.

2.1.2. Land surface temperature data

The land surface temperature data used in this study are remotely sensed Landsat Level-2 observations (Earth Resources Observation and Science (EROS) Center, 2020). Landsat Level-2 observations provide LST in an 8-day interval at an overpass at approximately 10:20 UTC (11:20 CET/local time). Landsat Level-2 LST data has a spatial resolution of 30 m × 30 m. Although the nominal resolution of the thermal infrared sensor is 100 m × 100 m, the LST product is downscaled to 30 m × 30 m based on the emissivity layer, which has a native resolution of 30 m × 30 m. Due to cloudy weather conditions, not all overpasses could be used for analysis. For the overall comparison between LST and UTCI, a spatial LST data availability threshold of 80 % per timestep is applied to ensure that only timesteps with high spatial coverage (cloud-free days) are analyzed. The 80 % threshold corresponds to more than 37,700 valid (cloud-free) 30 m × 30 m Landsat LST grid cells, resulting in 46 observations (timesteps) in summer and 77 for all seasons. For the subsequent cluster analysis (see section 2.2), the LST data availability threshold per timestep is set at 95 % (> 44,600 30 m × 30 m Landsat LST grid cells with valid data) resulting in only 22 observations. The threshold was set to 95 %, as the study aims to map the entire urbanized area of Freiburg, and if an observation of a grid cell is missing, the grid cells cannot be considered in the cluster analysis. An overview of the dates for which LST data was used is given in Table A2.

2.1.3. Thermal comfort data

Human outdoor thermal comfort data are obtained by applying the HTC-NN model by Briegel et al. (2023, 2024), (Fig. 2 - Step 1). The HTC-NN is a machine learning emulator of numerical urban climate models designed for rapid, high-resolution prediction of human OTC in complex urban environments, developed and evaluated for Freiburg, Germany. The model directly predicts two-dimensional fields of UTCI with a resolution of 1 m × 1 m resolving buildings and vegetation. An overview of the different UTCI stress classes is given in Table A1 (Appendix A).

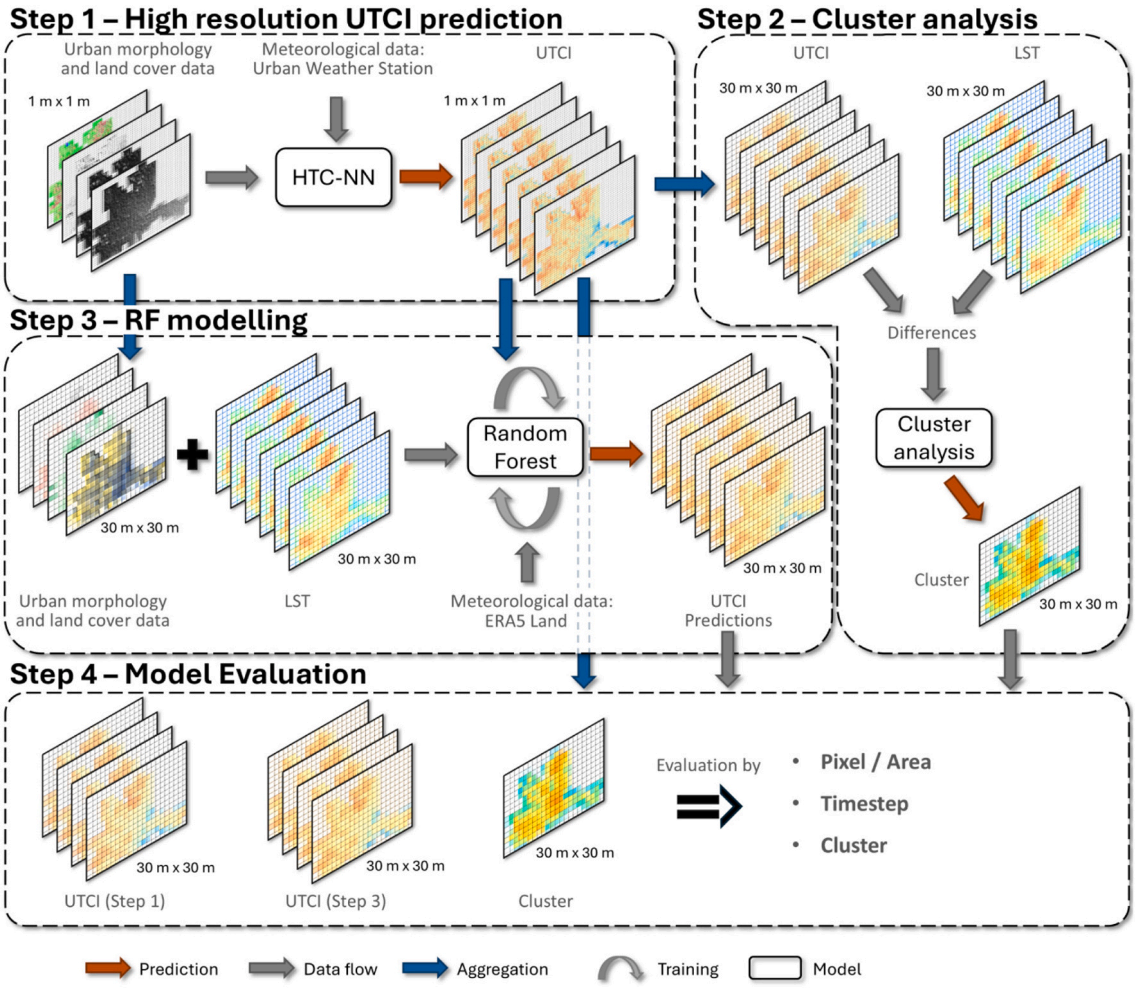

Fig. 2. Schematic representation of the LST analysis. In the first step (1), HTC-NN is used to predict high resolution UTCI maps (1 m), which are then aggregated to 30 m × 30 m and the differences to LST are calculated for the cluster analysis (2). In addition, different random forest models are trained using different combination of aggregated local-specific data (urban morphology, land cover), LST, and ERA5-Land data (3). In step 4, the different models are evaluated.

Spatial input data of the HTC-NN include spatially resolved land cover classes (e.g., buildings, trees, paved surfaces), digital elevation and surface models (for both buildings and vegetation), and sky view factor maps, all at 1 m resolution. Sky view factor maps are generated using the UMEP scheme (Lindberg et al., 2018). The model is forced with T_a, incoming long- and shortwave radiation, relative humidity, air pressure, and wind speed and direction. The forcing data are provided by an automated urban weather station located on a rooftop (approx. 60 m a.g.l.) in the city center of Freiburg (Briegel et al., 2023). The HTC-NN model is thoroughly validated using numerical urban climate model outputs, as well as measured UTCI data from a dense sensor network and from mobile measurements. The overall Mean Absolute Error (MAE) for UTCI is 2.3 K. For further details on model development and configuration, spatial and forcing data, and the validation procedure see Briegel et al. (2023, 2024). As the spatial resolution of the Landsat LST data is 30 m × 30 m, the 1 m resolution UTCI maps produced by the HTC-NN are resampled and averaged to match the coarser resolution of the Landsat LST data. As this study aims to investigate the potential of LST as a proxy for pedestrian-level OTC, 1 m pixels covered by buildings are excluded from the analysis and not considered for aggregation.

2.1.4. *ERA5-land and ERA5-HEAT reanalysis data*

The ERA5-HEAT dataset (Di Napoli et al., 2020) is used as a benchmark for the random forest UTCI models forced by LST and ERA5-Land data (Muñoz Sabater, 2019) (see Section 2.2). ERA5-HEAT is a global dataset of UTCI and mean radiant temperatures derived from ERA5 reanalysis data for open areas using the methods proposed by Di Napoli et al. (2021). ERA5-HEAT is available with an hourly resolution and a spatial resolution of 0.25° x 0.25° (∼27.75 km at the equator). From the ERA5-HEAT dataset, the grid cell centered at 48.00°N / 7.75°E is used as it covers the entire urbanized area of Freiburg (48.00°N, 7.80°E) and is mainly located in the Upper Rhine Valley, which represents the climate in the city of Freiburg better than neighboring grid cells covering the Black Forest mountains.

The ERA5-Land dataset (Muñoz Sabater, 2019) is used to obtain additional predictors for the random forest models (see Section 2.2). ERA5-Land data is a high-resolution reanalysis dataset with a horizontal resolution of 0.01° x 0.01° (∼ 9 km at the equator) and is available at hourly resolution. The grid cell used in this study is centered at 48.00°N / 7.80°E. From the ERA5-Land data, T_a, relative humidity (RH), incoming short- and longwave radiation (Kdown and Ldown), and U are extracted and used as additional predictors for the random forest models.

2.2. *Analysis of relationship between LST and OTC*

The following section describes the cluster analysis of the differences between LST and UCTI. the random forest modeling approach, and the evaluation of the random forest models (Fig. 2 Steps 2–4).

2.2.1. *Cluster analysis*

To better investigate and understand potential spatial relationships and systematic differences between LST and OTC, a cluster analysis of the differences (LST − UTCI) across grid cells and time steps is performed (Fig. 2 - Step 2). Differences of LST and the aggregated UTCI maps are calculated for approximately 45,000 grid cells (30 m × 30 m) and for each of the 22 overpasses (data availability threshold of 95 %, see section 2.1). The 44,600 × 22 matrix is then used for the cluster analysis and to assign each grid cell to a cluster.

The optimal number of clusters (*k*) is first determined using the NBClust package (Charrad et al., 2014) in R (R Core Team, 2024), which evaluates 30 different clustering indices (e.g., silhouette, gap statistic). Based on the majority rule across these indices, the optimal *k* is identified as 3 representing urban, suburban, and rural areas. However, the three clusters do not allow for the distinction of different urban environments and are therefore not suitable for an assessment of intra-urban variability. To further investigate the relationships between LST and OTC in different urban contexts, we perform hierarchical clustering as an additional step. This approach allows the identification of possible subgroups within the primary clusters. Hierarchical clustering is performed using a Euclidean distance matrix as the similarity measure and Ward's minimum variance method as the linkage criterion (Ward, 1963). This method identifies three primary clusters, with evidence of two additional sub-clusters. These sub-clusters suggest finer-scale patterns that may provide further insights into the urban spatial dynamics of LST and OTC.

We also investigated the relationship between LST and UTCI in terms of the different local climate zones (LST - Stewart and Oke, 2012). However, as Freiburg mainly comprises LCZ 2 (see Table A3), our focus is on the derived clusters, which are based directly on the LST-UTCI relationship and thus better represent its physical properties.

2.2.2. *Random forest modeling approach*

Random forest modeling is carried out to better assess whether LST is a good proxy for OTC in urban areas and whether LST can be used to derive accurate OTC products (Fig. 2 - Step 3). Therefore, several random forest models (Breiman, 2001) with different spatial and temporal input data (predictors) are developed and compared. An overview of the different model configurations is given in Table 1. The baseline model is forced with LST only (rf_LST). However, this simple linear model with only LST as a predictor with $\widehat{UTCI} = f(LST)$, cannot accurately capture the heterogeneity in the LST-UTCI relationship, as indicated by the wide 95 % confidence interval (Fig. A1). This suggests that additional predictors are needed to account for that large (intra-urban) variability. Another baseline model is forced with ERA5-Land data only (rf_ERA5), using only one grid cell for the entire model domain (see ERA5-Land section). In addition, a model is trained with a combination of LST and ERA5-Land data (rf_LST + ERA5). These three model configurations are forced by the globally available data products LST and ERA5-Land, and are further extended by incorporating specific data on local urban morphology and landcover (Table 1) such as land cover class (LCC) fractions (Briegel and Christen, 2025), average sky view factor, and frontal area index of each individual 30 m × 30 m grid cell (rf_LST + Urb, rf_ERA5 + Urb, and rf_All, respectively). Sky view factor and frontal areas index are computed by using the UMEP scheme (Lindberg et al., 2018). In total, six model combinations are explored, allowing a comprehensive analysis of the predictor contributions.

Each of these six models is trained by two approaches: one focusing on continuous UTCI values through regression and the other on UTCI stress class classification (Table A1), resulting in a total of 12 models. For each of the individual 12 models, a random search optimization is performed to find the hyperparameter set with the highest accuracy (lowest Root Mean Square Error - RMSE). The random search is conducted individually for each model, as the number of predictors is different for each model. Overall, the search range of the number of trees varies from 200 to 800, the number of predictors randomly sampled at each split from 1 to 4, and the minimum number of data points in a node required to split the node further from 1 to 10. A detailed table with the hyperparameter set is given in Appendix A (Table A4). The dataset is divided into train and test data to evaluate the performance of the final model. The train data covers 35,365 areas (30 m × 30 m) and 62 timesteps and the test data covers 11,792 areas and 24 timesteps resulting in ∼2 million data points for training and ∼ 250,000 data points for testing. For robust hyperparameter optimization, the train data is further divided into 5 folds and a 5-fold cross-validation is applied to each hyperparameter set. The best hyperparameter set is used to train the final model by training on the entire train data. Final model is evaluated by testing against the test data. All modeling is done in R (R Core Team, 2024) using the ranger package (Wright and Ziegler, 2017).

2.2.3. *Random Forest model evaluation*

The UTCI heat stress classification accuracy of the random forest models is evaluated using the Overall Accuracy (ACC) and the True Skill Score (TSS – Stull, 2015). ACC describes simply the percentage of all correct stress class predictions. The TSS is defined as follows:

$$\mathrm{TSS} = \frac{a}{a+c} - \frac{b}{b+d} \tag{1}$$

with true positives (*a*), false positives (*b*), false negatives (*c*), and true negatives (*d*). TSS closer to +1 indicates more accurate prediction, TSS closer to 0 indicates less accurate prediction.

To evaluate the model skill and performance and to assess the ability of LST as a proxy for OTC, the predictions of the rf_LST model are compared with those of the other random forest models and additionally with the ERA5-HEAT dataset as a baseline (Di Napoli et al., 2020, 2021), which is described in section 2.1.

3. Results

3.1. *Comparison of LST and UTCI*

Fig. 3 shows the averaged UTCI of each 30 m grid cell as a function of LST for the summer season (Fig. 3a) and all seasons (Fig. 3b). Pearson's correlation coefficients between all 30 m × 30 m grid cells of LST and

Table 1

Different models with predictor combinations. All globally available data besides LST are extracted from ERA5-Land data. The six model configurations are carried out as regression (UTCI) and as classification (UTCI stress class).

Variables	Random forest models with different predictor combinations					
	Globally available data			Specific local data on urban morphology and LCC		
	rf_LST	rf_ERA5	rf_LST + ERA5	rf_LST + Urb	rf_ERA5 + Urb	rf_All
LST	✓		✓	✓		✓
ERA5-Land[1]		✓	✓		✓	✓
Data on urban morphology and LCC[2]				✓	✓	✓

[1] air temperature, incoming shortwave radiation, incoming longwave radiation, relative humidity, wind speed.

[2] LCC: land cover class fractions (buildings, trees, low vegetation, paved, water), sky view factor, frontal area index.

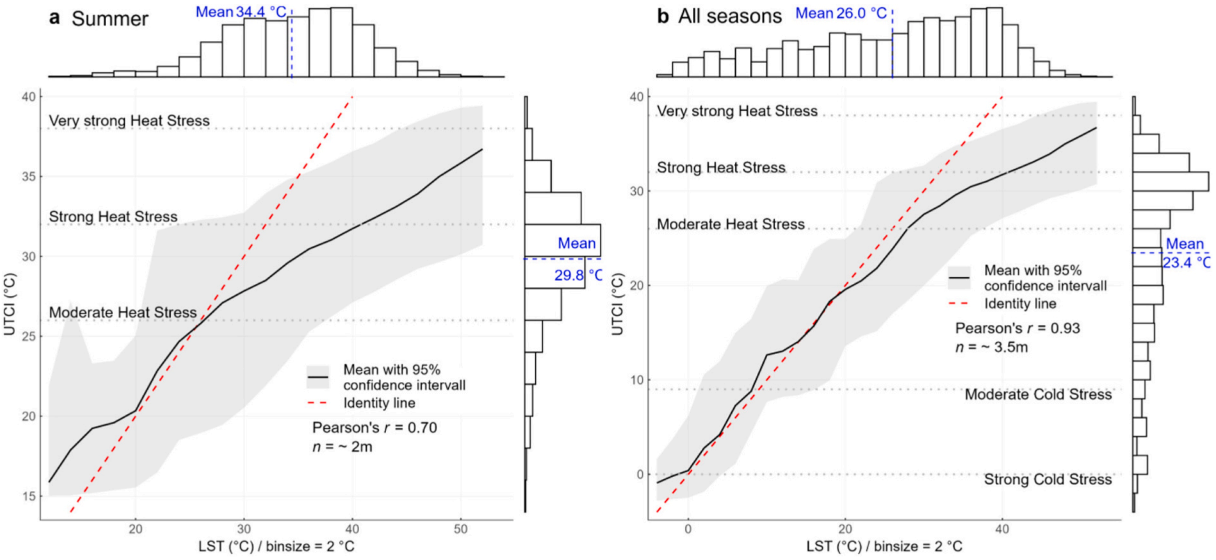

Fig. 3. UTCI as a function of Landsat LST for 2 million summer observations (a) and for 3.5 million annual observations (b) with a bin size of 2 °C LST. Black lines show mean value and gray areas the 95 % confidence interval. The histograms on top and at the right show the LST and UTCI distributions, respectively. Note, x and y axis differ for a and b. The average 95 % confidence intervals are 10.3 K and 11.1 K for (a) and (b), respectively.

UTCI for summer and all seasons are 0.70 and 0.93 respectively. The average UTCI of all grid cells across all time steps is 23.5 °C for all seasons and 29.8 °C in summer, while the average LST is 26.0 °C for all seasons and 34.4 °C in summer. Fig. 3b shows that UTCI and LST have a linear relationship between 0 °C and 20 °C UTCI. Between 20 °C and 25 °C UTCI, the slope decreases, indicating a transition to a non-linear relationship. Above 25 °C UTCI, the relationship flattens further, showing a decreasing response of UTCI to increases in LST. The relationship between UTCI and LST is almost linear for non-heat stress situations (UTCI <26 °C), but non-linear in general and for heat stress

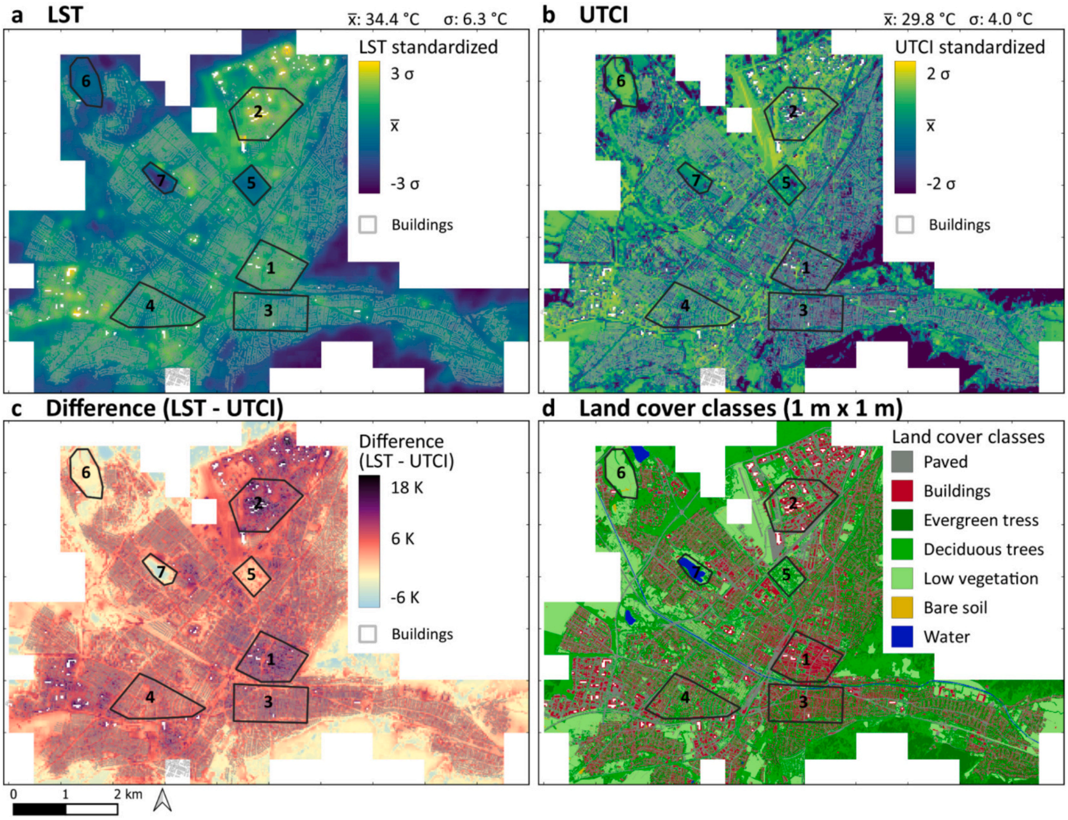

Fig. 4. Standardized maps of averaged LST (a) and UTCI (b) for 46 timesteps in summer. The difference between LST and UTCI (not standardized) is shown in (c), and the land cover classes are shown in (d). Maps (a)-(c) have a spatial resolution of 30 m × 30 m and (d) of 1 m × 1 m. The labeled areas in the maps indicate different urban morphologies with (1) densely built-up city center, (2) industrial area, (3) and (4) residential areas with a mixture of buildings and trees, (5) park area with low vegetation and trees, (6) open areas with low vegetation, and (7) park area with a lake. Standardization is done by subtracting the mean from each value and then dividing by the standard deviation.

situations (UTCI ≥26 °C). In the summer period (Fig. 3a) no linear relationship between UTCI and LST can be seen. In addition, the 95 % confidence interval spans up to 10 K, indicating that the large intra-urban variability of UTCI cannot be expressed by LST.

Figs. 4a, b show standardized maps of the averaged LST and UTCI and Fig. 4c the corresponding average differences (LST – UTCI) for the 46 timesteps of the summer period. An overview of the different land cover classes in Freiburg is shown in Fig. 4d, with more detailed maps in Fig. S1 (Supplement). LST has different spatial patterns than UTCI. LST is highest in industrial areas (2) and the densely built-up city center (1) and lowest in areas with a high proportion of trees or forest such as parks (5) and open areas with unsealed or water surfaces (6, 7). Residential areas (3, 4) have lower LST than sealed industrial areas, but higher LST than unsealed/forested areas. UTCI, on the other hand, has partly reversed spatial patterns, with the highest values in industrial (2) and open areas (6, 7), and the lowest values in densely built-up areas (1) and residential areas with densely built-up and trees (3). Areas with a mixed land cover have intermediate UTCI values (4). Fig. 4c shows the differences and the highest positive differences are found in densely built-up areas (1), industrial areas (2), and residential areas (3). The lowest (negative) differences are found in areas with large water bodies (7) and open, unsealed areas (6). These patterns can also be seen in Fig. 5, which shows LST, UTCI, and the corresponding differences as functions of the land cover classes buildings, paved, trees, and low vegetation for the 30 m × 30 m grid cells. LST increases as the fraction of buildings and paved surfaces increases, and decreases as the fractions of vegetated areas such as low vegetation or trees increases (Fig. 5a). On the other hand, UTCI follows a pattern more related to pedestrian-level shading, with UTCI increasing in open areas with higher fractions of low vegetation or paved surfaces and decreasing with increasing fractions of buildings and trees (Fig. 5b). The corresponding difference summarizes the relationship between LST and UTCI, with partly inverse pattern for the different land cover classes. Higher building fractions resulting in larger differences and higher tree or low vegetation fractions resulting in smaller differences (Fig. 5c). The only land cover which seems to have no positive or negative influence on the relationship are paved surfaces.

3.2. *Cluster analysis*

The results of the hierarchical cluster analysis of the LST-UTCI differences, mapped across three and five spatial clusters, are shown in Fig. 6a and b and in Table 2.

The statistical analysis revealed that the three-cluster classification most accurately reflects the relationship between LST and UTCI. This classification provides a clear gradient ranging from vegetated to highly urbanized environments. Cluster A comprises the most natural areas and shows the highest levels of grass (31.2 %) and tree (50.8 %) cover, low building density (5.3 %) and a relatively unobstructed view of the sky (SVF = 0.63). This cluster also exhibits the lowest LST (31.0 °C) and the smallest LST–UTCI difference (0.97 K). Cluster B reflects a more mixed urban form, with increased building (17.3 %) and paved surface (28.0 %) fractions, a slightly lower SVF (0.61), and higher LST (35.0 °C) and LST–UTCI differences (4.82 K). Cluster C is the most urbanized, showing the highest building fraction (31 %), the lowest SVF (0.481), and the smallest green cover. It corresponds to the highest LST (36.9 °C) and the largest LST–UTCI difference (8.03 K). While the three-cluster classification provides strong statistical and thematic coherence, the five-cluster solution is used for further analysis as it offers greater spatial differentiation and captures important variations within categories, particularly in vegetated areas and among moderately to densely built-up urban forms.

The vegetation and water cluster (A.I) has the highest proportion of low vegetation (37 %) and trees (49 %), combined with minimal built-up fraction (3 %). Similarly, cluster A.II has a slightly higher proportion of paved surfaces (12 %) but remains predominantly vegetated and has a lower Sky View Factor (0.57 vs 0.74 in A.I). Cluster A.I and A.II show a distinct spatial pattern, dominating areas outside the built-up environment and occurring in large parks and cemeteries. Cluster B represents a mixture of buildings (17 %), paved surfaces (28 %), and vegetation (39 % trees and 15 % low vegetation) and represents transitional areas between vegetated clusters A.I/A.II and densely built-up clusters C.I/C.II, but also low density residential areas. The urbanized clusters (C.I and C.II) are located in the city center, industrial zones, and densely built-up residential areas. These clusters are characterized by a higher proportion of buildings (25 % in C.I, 42 % in C.II), lower low-vegetation cover (7 % and 4 %, respectively), and decreasing tree fractions (38 % and 30 %). The paved fraction also increases from cluster A.I to C.I, but decreases to cluster C.II. The increasing built-up fraction from A.I to C.II highlights a clear transition from natural to urban environments. The morphometric parameters further emphasize this urbanization trend. The sky view factor, which indicates the openness of the urban canopy (including buildings and vegetation), decreases significantly from 0.74 in the vegetation and water clusters A.I to 0.41 in the more urbanized clusters C.I and C.II. Similarly, the frontal area index and mean building

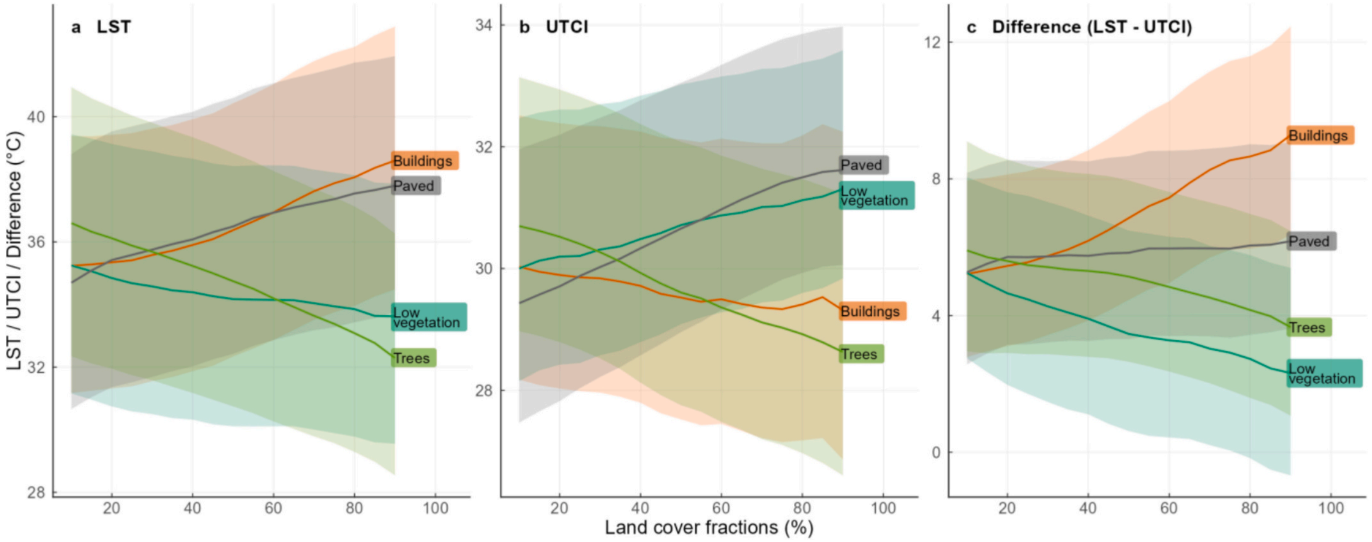

Fig. 5. LST (a), UTCI (b), and the corresponding differences (c) as functions of land cover class fractions for the 30 m × 30 m grid cells. Shaded areas represent the 95 % confidence interval. Note that the y-axes differ.

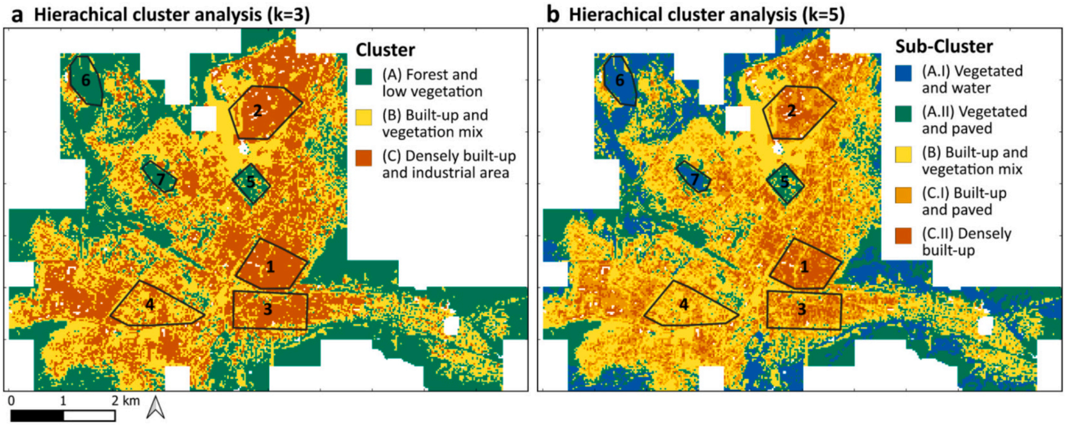

Fig. 6. Maps of the hierarchical cluster analysis of the LST-UTCI differences cut by 3 (a) and 5 (b) clusters. Cluster A.I and A.II are sub-clusters of cluster A, and C.I and C.II are sub-clusters of cluster C. Spatial resolution is 30 m × 30 m. The labeled areas in the maps are the same as in Fig. 4 and indicate different urban morphologies with (1) densely built-up city center, (2) industrial area, (3) and (4) residential areas with a mixture of buildings and trees, (5) park area with low vegetation and trees, (6) open areas with low vegetation, and (7) park area with a lake.

Table 2
Results of the cluster-analysis distinguished by temperature types, land cover fraction, and morphometric parameters. The temperature values give the average of each cluster and the values in brackets (±) below the average temperature values are the standard deviations.

	Cluster and sub-cluster				
	(A) Forest and low vegetation area		(B) Built-up and vegetation mix	(C) Densely built-up and industrial area	
	(A.I) Vegetated and water	(A.II) Vegetated and paved		(C.I) Built-up and paved	(C.II) Densely built-up
Temperature types					
LST	30.1 °C (± 5.7 °C)	31.5 °C (± 6.5 °C)	35.0 °C (± 6.4 °C)	36.5 °C (± 6.5 °C)	37.7 °C (± 6.7 °C)
UTCI	30.9 °C (± 3.9 °C)	29.6 °C (± 4.5 °C)	30.2 °C (± 4.3 °C)	29.3 °C (± 4.4 °C)	27.9 °C (± 4.4 °C)
Difference (LST – UTCI)	−0.8 K (± 3.2 K)	2.0 K (± 3.2 K)	4.8 K (± 3.3 K)	7.1 K (± 3.2 K)	9.7 K (± 3.5 K)
Land cover fractions					
Buildings	3 %	7 %	17 %	25 %	42 %
Low vegetation	37 %	28 %	15 %	7 %	4 %
Paved	7 %	12 %	28 %	29 %	24 %
Trees	49 %	52 %	39 %	38 %	30 %
Water	5 %	1 %	< 0 %	< 0 %	< 0 %
Morphometric parameters					
Sky view factor	0.74	0.57	0.61	0.52	0.41
Frontal area index	0.02	0.06	0.13	0.19	0.26
Mean building height	1.0 m	2.5 m	5.8 m	7.8 m	10.4 m
Mean canopy height	6.3 m	9.2 m	6.4 m	6.8 m	6.8 m

height increase, indicating a greater presence of vertical structures in urbanized areas.

A clear LST gradient is observed from clusters A.I to C.II, with LST increasing from 30.1 °C in the vegetated Cluster A.I to 37.7 °C in the densely built-up Cluster C.II. Conversely, UTCI follows an inverse pattern, with the highest value recorded in Cluster A.I (30.9 °C) and the lowest in Cluster C.II (27.9 °C). The difference between LST and UTCI increases with urbanization, ranging from −0.8 K in A.I to 9.7 K in C.II. The differentiation between sub-clusters (A.I vs. A.II, C.I vs. C.II) shows the importance of a more detailed classification beyond broad urban-rural categorizations. Although the cluster can account for some of the intra-urban variability of LST-UTCI relationship, the within-cluster variability is still high (Fig. A2). For example, the average 95 % confidence interval span is 7.5 K within cluster C.II, compared to 10.3 K across the overall relationship.

Besides the cluster analysis, an overview of the averages of LST, UTCI, and their corresponding differences by Local Climate Zone (LCZ) is given in Table A3, showing a similar increasing difference from rural to more urbanized LCZs.

3.3. Random Forest modeling

The performances of the regression and classification random forest models are evaluated for predicting continuous UTCI values and UTCI stress classes for different combinations of predictors, including models with and without LST. The predictions are made for 11:00 local time (CET) and match the spatial extent and resolution of the LST data and aggregated UTCI data from the HTC-NN.

A detailed comparison of the results for the regression random forest models and the different predictor combinations is given in Table 3 (Table A5 for classification models). In addition, the comparison of the predictions of the HTC-NN (1 m resolution), HTC-NN predictions aggregated to 30 m (excluding building footprints), and the LST data with the in-situ UTCI measurements is given in Table A6. Since ACC and

Table 3

Model evaluation results: Root Mean Square Error (RMSE – regression models only), overall accuracy of predicted UTCI classes (ACC), and True Skill Score (TSS) for the different UTCI classes: no thermal stress (8 °C ≤ UTCI <26 °C), cold stress (UTCI ≤8 °C), for all heat stress classes (UTCI ≥26 °C), moderate heat stress (26 °C ≤ UTCI <32 °C), and ≥ strong heat stress (UTCI ≥32 °C).

Model	RMSE (UTCI)	ACC	TSS (no thermal stress)	TSS (cold stress)	TSS (heat stress)	TSS (moderate heat stress)	TSS (≥ strong heat stress)
rf_LST	3.52 K	74 %	0.81	0.86	0.86	0.56	0.48
rf_ERA5	2.90 K	80 %	0.88	0.86	**0.92**	0.63	0.58
rf_LST + ERA5	2.83 K	82 %	**0.89**	0.86	**0.92**	0.62	**0.70**
rf_LST + Urb	3.11 K	76 %	0.81	0.81	0.89	0.61	0.54
rf_ERA5 + Urb	2.60 K	81 %	0.85	0.86	0.87	0.62	0.68
rf_All	**2.34** K	**83 %**	**0.89**	**0.86**	0.91	**0.66**	0.68
ERA5-HEAT	3.71 K	77 %	0.83	0.47	**0.92**	0.45	**0.81**

TSS of the regression and classification models are similar, but regression models allow continuous prediction of UTCI values, providing additional flexibility over classification models, the following analyses are carried out focusing on the regression model approaches.

3.3.1. *Overall Accuracies of random forest predictions*

Among the models using only globally available predictors (LST and ERA5-Land), the combined rf_LST + ERA5 regression model performs best, with an RMSE of 2.83 K and an ACC of 82 %. However, the rf_ERA5 model performs almost similarly, with a slightly higher RMSE of 2.90 K and an ACC of 80 %, indicating that the inclusion of LST data slightly improves the prediction accuracy. The slight improvement is particularly evident for strong heat stress conditions, where rf_LST + ERA5 achieves a TSS of 0.70 compared to 0.58 for rf_ERA5. In contrast, the rf_LST model, which relies solely on LST, has the lowest performance of the meteorological models, with an RMSE of 3.52 K and an ACC of 74 %.

The addition of specific local data on urban morphology and land-cover as predictors to the models leads to mixed results. While rf_ERA5 + Urb and rf_LST + Urb achieve lower RMSEs (2.60 K and 3.11 K) than their non-local-specific counterparts, the improvements of ACC are marginal, increasing to 81 % and 76 % respectively. The effect on classification performance is more nuanced. The addition of local-specific predictors does not improve TSS values for general heat stress

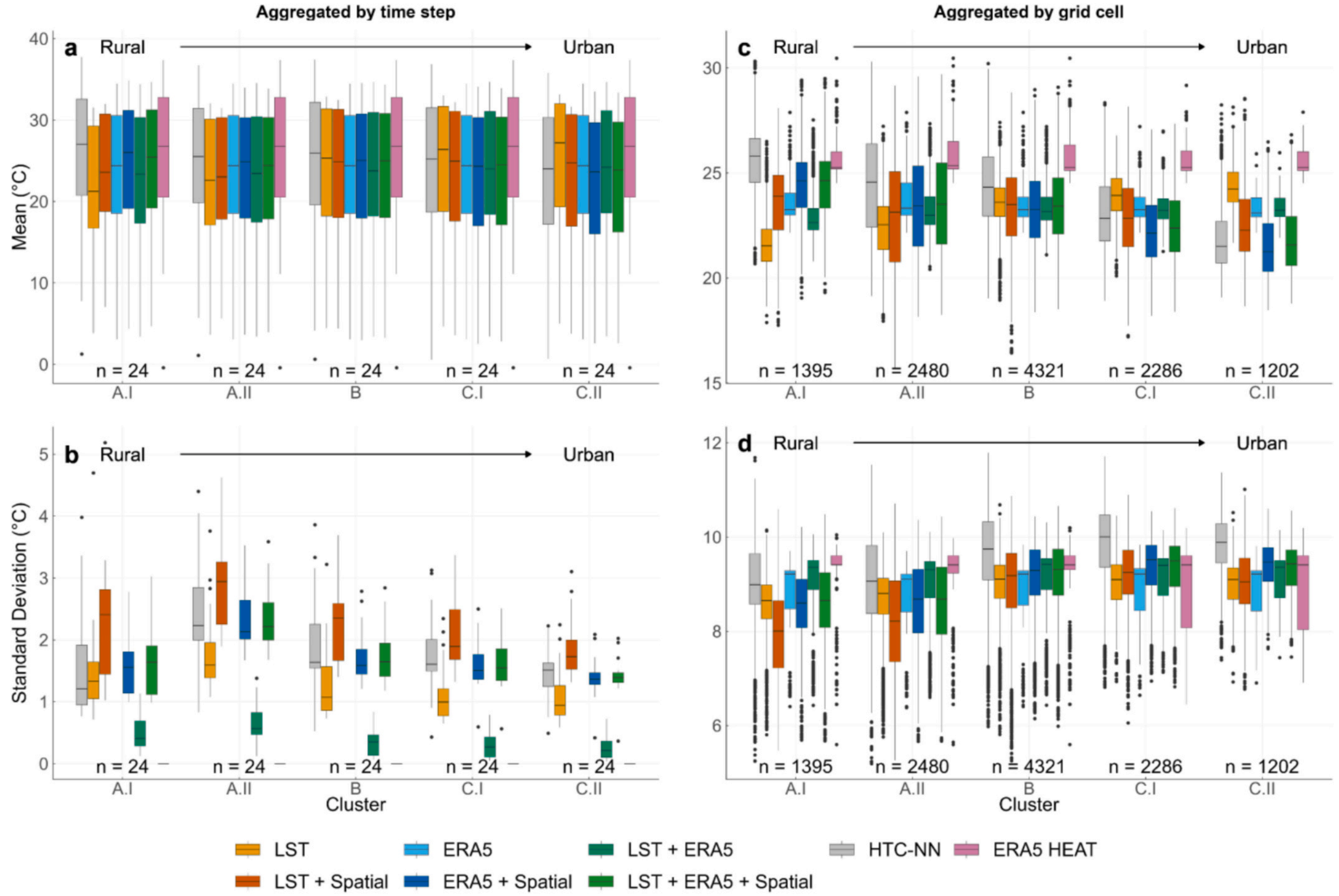

Fig. 7. Boxplots of the mean and standard deviation of UTCI of the test data (HTC-NN), the predictions on the test data of the different random forest models, and the ERA5-HEAT dataset, aggregated by date (a, b) and by grid cell (c, d) for the different clusters A (rural) to C (urban). In (a) and (b) each observation represents the mean and standard deviation of all grid cells per time step. In (b), for example, this means that it represents the spatial variability of each timestep. In (c) and (d), each observation represents the mean and standard deviation of all time steps per grid cell. This means that (d) represents the temporal variability within each grid cell.

(UTCI ≥26 °C), but lead to higher TSS values for moderate heat stress (26 °C ≤ UTCI <32 °C) and strong heat stress (UTCI ≥32 °C). This means that models with local-specific data on urban morphology and LCC provide a more detailed representation of heat stress variability, particularly in distinguishing between different heat stress classes. The rf_All model, which integrates all available predictors, achieves the highest ACC (83 %) and the lowest RMSE (2.34 K), demonstrating that the combination of globally available data and local-specific data leads to the most robust predictions. The ERA5-HEAT dataset, used as a baseline for comparison, has a higher error than the random forest models with an RMSE of 3.71 K and a similar accuracy to the LST forced models with an ACC of 77 %. ERA5-HEAT systematically overestimates UTCI (see also Fig. 8b and c), resulting in poor performance in identifying cold or no stress conditions, but better classification of moderate and strong heat stress. Despite its higher RMSE, ERA5-HEAT achieves a similar TSS to rf_All for general heat stress conditions, showing that although its absolute predictions are less accurate, it remains a competitive dataset for heat stress classification.

3.3.2. *Representation of the intra-urban variability of OTC*

Although ERA5-HEAT and the random forest models incorporating only globally availably data appear to have similar TSS and ACC to the random forest models additionally incorporating local-specific data, they cannot accurately capture the intra-urban variability of OTC. In

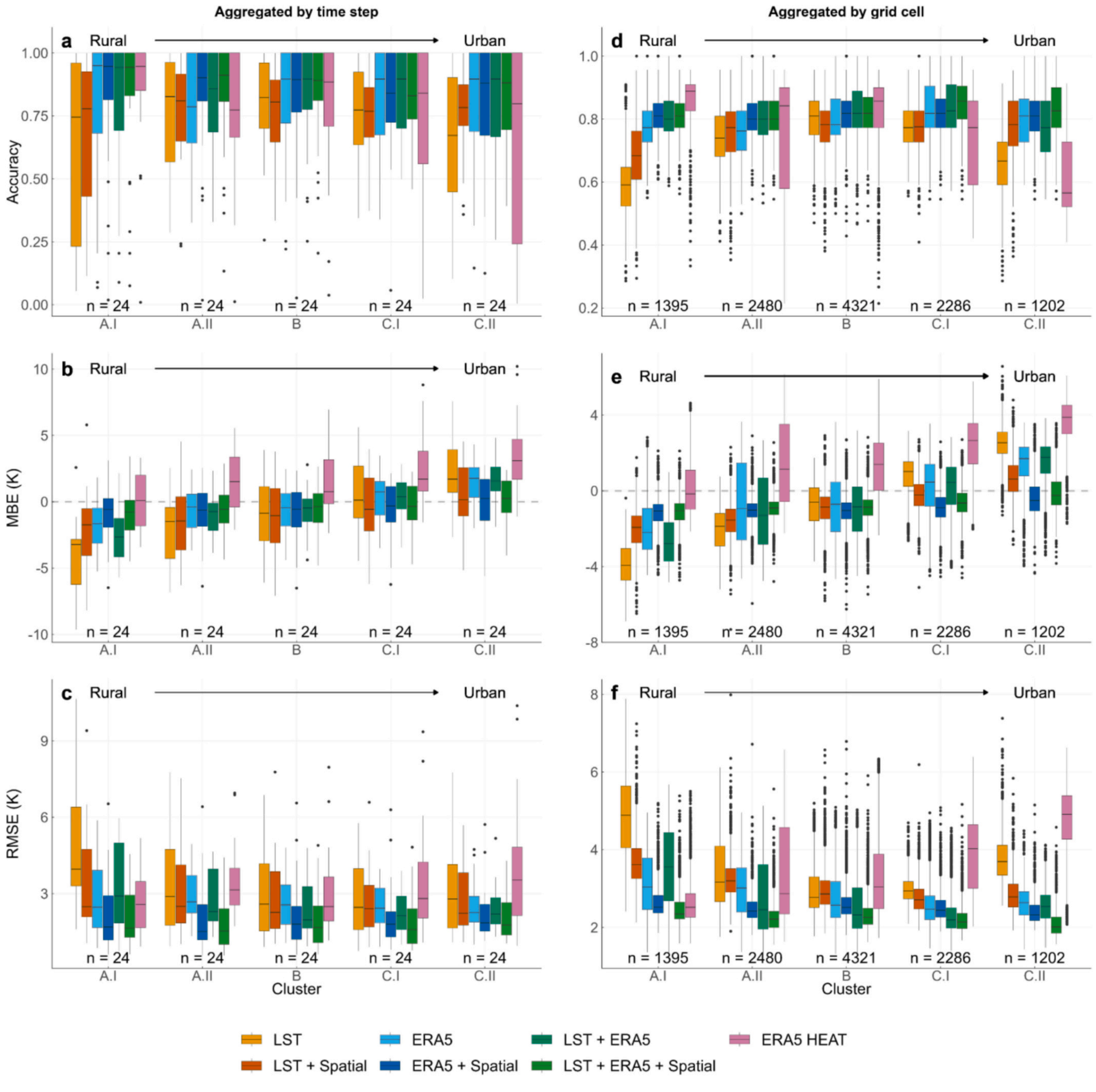

Fig. 8. Boxplots of the overall accuracy (ACC), Mean Bias Error (MBE), and Root Mean Square Error (RMSE) of the random forest model predictions and the ERA5-HEAT dataset, aggregated by date (a, b, c) and by area (d, e. f) for the different clusters A (rural) to C (urban). In (a-c), each observation represents the correspondings metric of all grid cells per time step. In (d-f), each observation represents the corresponding metric for all time steps per area, reflecting the temporal variability within each area.

fact, the rf_ERA5 model has no intra-urban variability as forcing is globally for the entire model domain and no further information at the 30 m resolution is provided. The intra-urban variability of the random forest predictions can be assessed from Fig. 7, which shows the mean and standard deviation aggregated by time step and grid cell (30 m × 30 m), from Fig. 8, which shows the evaluation metrics ACC, Mean Bias Error (MBE), and RMSE aggregated by time step and grid cell, and from Fig. A3, which shows predictions of an exemplary timestep of the different RF models. Both figures show these metrics for the different spatial clusters. For example: time step aggregation in Fig. 7b means that each point in these boxplots represents the standard deviation of all areas at one time step. In addition, Table 4 and Table A8 show the TSS by UTCI stress class and cluster.

The random forest models forced with additional local-specific data on urban morphology and LCC capture the intra-urban heterogeneity much better than the models forced only with globally available data across all clusters. The rf_LST model has the highest RMSE, lowest ACC, and an inconsistent MBE across the different cluster (Figs. 7 and 8). The addition of ERA5-Land data in rf_LST + ERA5 seems to improve the performance, reducing the RMSE and increasing ACC. However, the representation of OTC intra-urban variability deteriorates as the standard deviation for each time step decreases (Fig. 7b). The rf_ERA5 predictions and the ERA5-HEAT dataset have no intra-urban variability as the standard deviation for each time step is 0, meaning that all grid cells have the same value (Fig. 7b), making these predictions not suitable to map intra-urban variability of OTC. In summary, random forest models forced only with globally available data remain inferior to local-specific informed random forest models, which consistently achieve lower RMSE, more consistent MBE (Fig. 8), and more similar distributions of mean and standard deviation compared to the HTC-NN “ground truth” across all clusters (Fig. 7). The rf_ERA5 + Urb and rf_All models perform best, as their boxplots of mean and standard deviation best match those of HTC-NN (Fig. 7), and the RMSE and MBE are homogeneously distributed across all clusters (Fig. 8). This means that the addition of local-specific data on urban morphology and LCC makes the prediction more consistent across the different clusters and thus enables a better mapping of intra-urban variability of OTC. Based on this, further investigation is focused on the rf_All model and the ERA5-HEAT as baseline.

The predictor importance of the rf_All model is given in Table A7. Although meteorological variables are the most important predictors (accounting for ~68 % of the variance), spatial variables such as land cover classes and sky view factors are essential for capturing intra-urban OTC variability. Spatial predictors modify the global forcing patterns, enabling neighborhood-scale variability in RF model outputs.

The performance of the rf_All model and the ERA5-HEAT dataset in predicting UTCI and its associated thermal stress classes and their respective predictive capabilities across different urban clusters is given in Table 4. The rf_All model shows a high and consistent ACC, ranging from 81 % to 85 % across all clusters. The RMSE decreases with increasing urban complexity, from 2.53 K in cluster A.I to 2.18 K in cluster C.II, showing the improved predictive performance in all clusters but particularly in more built-up environments. In terms of classification skill, the TSS for heat stress remains stable around 0.90 across clusters A.I to C.I/C.II. For moderate heat stress, the TSS increases with urbanization, indicating a better differentiation between moderate and strong heat stress in densely built-up areas. Conversely, the ERA5-HEAT dataset exhibits an inverse trend. Its RMSE increases from 2.78 K to 4.80 K with increasing urbanization (from cluster A.I to C.II), indicating reduced accuracy in densely built-up areas. Similarly, ACC decreases from 83 % to 59 %, reflecting reduced classification performance. This downward trend is also observed for TSS across all thermal stress categories. The higher TSS for strong heat stress of the ERA5-HEAT is caused by a systematic overestimation (Fig. 8c and e). This decrease in model performance from rural to urban areas highlights the expected limitations of the ERA5-HEAT dataset in capturing the nuanced thermal dynamics of densely populated built-up areas.

The results show that the rf_All model has a superior performance compared not only to the ERA5-HEAT dataset, but also to the other random forest models forced with globally available data such as LST or ERA5-Land only. In particular, in densely built-up areas, the rf_All model, which also incorporates local-specific data on urban morphology and LCC, provides a more accurate representation of the thermal processes and thus of the OTC conditions.

4. Discussion

4.1. LST UTCI relationship

The results show that there is no clear linear relationship between LST and UTCI in urban areas and that this non-linearity cannot be solved easily due to the large intra-urban variability of OTC. As there are only a

Table 4

Model selection results by cluster for the rf_All model and the ERA5-HEAT dataset with Root Mean Square Error (RMSE), overall accuracy of predicted UTCI classes (ACC), and True Skill Score (TSS) for the different UTCI classes: no thermal stress ($8\ ^{\circ}\mathrm{C} \leq \mathrm{UTCI} < 26\ ^{\circ}\mathrm{C}$), cold stress ($\mathrm{UTCI} \leq 8\ ^{\circ}\mathrm{C}$), for all heat stress classes ($\mathrm{UTCI} \geq 26\ ^{\circ}\mathrm{C}$), moderate heat stress ($26\ ^{\circ}\mathrm{C} \leq \mathrm{UTCI} < 32\ ^{\circ}\mathrm{C}$), and $\geq$ strong heat stress ($\mathrm{UTCI} \geq 32\ ^{\circ}\mathrm{C}$).

		Cluster and sub-cluster				
		(A) Forest and low vegetation area		(B) Built-up and vegetation mix	(C) Densely built-up and industrial area	
		(A.I) Vegetated and water	(A.II) Vegetated and paved		(C.I) Built-up and paved	(C.II) Densely built-up
rf_All						
RMSE (UTCI)		**2.53 K**	**2.30 K**	**2.38 K**	**2.25 K**	**2.18 K**
ACC		81 %	**81 %**	**83 %**	**85 %**	**83 %**
	no stress	0.90	**0.86**	**0.90**	**0.91**	**0.86**
	cold stress	**0.95**	**0.82**	**0.92**	**0.85**	**0.74**
TSS	heat stress	0.91	0.88	0.91	**0.93**	**0.92**
	moderate heat stress	0.60	**0.63**	**0.65**	**0.71**	**0.71**
	≥ strong heat stress	0.68	0.66	0.69	0.70	0.68
ERA5-HEAT						
RMSE (UTCI)		2.78 K	3.68 K	3.43 K	4.02 K	4.80 K
ACC		**83 %**	73 %	77 %	70 %	59 %
	no stress	**0.93**	0.81	0.86	0.80	0.70
	cold stress	0.69	0.53	0.48	0.39	0.32
TSS	heat stress	**0.97**	**0.89**	**0.94**	0.92	0.87
	moderate heat stress	**0.65**	0.43	0.53	0.38	0.21
	≥ strong heat stress	**0.88**	**0.80**	**0.84**	**0.79**	**0.72**

limited number of studies that directly compare LST with high-resolution spatial fields of OTC, our findings can only be partially assessed against those in the literature. Li et al. (2023a, 2023b) compared Landsat 8 LST data with high-resolution T_{mrt} data, while Fahy et al. (2025) compared Landsat 8 LST data with measured PET data. However, whereas Li et al. (2023a, 2023b) conducted a spatial comparison, their focus was on T_{mrt} rather than UTCI. In contrast, Fahy et al. (2025) only had transects of measurement data, not spatially resolved data on PET. We therefore set our findings additionally in context to individual studies of LST and OTC in urban areas.

The LST-OTC relationship is influenced by several physical processes related to the nature of street-level OTC and remotely sensed LST, as well as the meteorological background conditions. Street-level OTC is predominantly influenced by the radiative environment and shading patterns during the day (Lindberg et al., 2016; Middel et al., 2021), whereas LST is mainly controlled by land surface and material properties (Chakraborty and Lee, 2019; Schwaab et al., 2021). This means that open areas, whether sealed or covered by low vegetation, have higher UTCI values in the absence of shading, and shaded areas have lower UTCI values regardless of whether they are shaded by trees or buildings (Coutts et al., 2016; Briegel et al., 2024). LST, on the other hand, is influenced by the thermal properties (emissivity) of the surface material, and areas with all types of vegetation have potentially lower LST, while built-up and paved areas have higher LST values (Ferreira and Duarte, 2019). This means that in densely built-up areas, such as city centers, LST measurements are affected by anisotropy and dominated by roofs (to the neglect of other surfaces) and respond to the thermal properties of the roofing material (Stewart et al., 2021). On the other hand, street-level OTC represents the actual thermal conditions in the street canyons within the first 2 m, which are often shaded by the built environment or trees (Li et al., 2023a, 2023b).

The results of our study support these findings by showing significant intra-urban spatial variability in LST, UTCI, and the LST-UTCI relationship, which are closely tied to land cover types and urban morphology. The relationship between LST and UTCI for non-heat stress situations (UTCI <26 °C) appears robust as it is predominantly linear. However, as UTCI increases, particularly above 26 °C UTCI, the relationship becomes increasingly non-linear, with UTCI showing a diminished response to increases in LST. One potential reason for this could be the faster, more gradual heating of roof surfaces compared to the OTC in the street canyon. At a certain T_a or solar irradiance, LST may continue to rise due to the solar heating of rooftops and upper building surfaces, while pedestrian-level OTC in street canyons may stay constant due to increased shading and reduced solar exposure. Li et al. (2023a, 2023b) found similar results for T_{mrt} and concluded that LST and T_{mrt} have a stronger relationship on rooftops rather than in street canyons. In addition to the non-linearity relationship, the span of the 95 % confidence interval of more than 10 K in Fig. 3a indicates significant intra-urban variability in the LST-UTCI relationship, for both the linear and non-linear parts. For instance, at an LST value of 35 °C, the range of UTCI values is 25–35 °C, representing conditions ranging from no thermal stress to strong heat stress. These findings suggest that it is difficult to draw any conclusions from LST towards pedestrian-level OTC.

The cluster analysis and the corresponding land cover class fractions provide useful context for further examination of these results. The higher the building fraction or other urban morphometric indices (Frontal Areas Index or Building Height), the higher the differences between LST and UTCI (7 to 9 K for cluster C) with relatively low UTCI values and relatively high LST values. Conversely, the higher the vegetation fraction, the lower the differences (−1 to +2 K for cluster A) with relatively high UTCI values and relatively low LST values. The cluster analysis reveals a progressive shift from natural to urban environments, accompanied by increasing LST, decreasing UTCI, and increasing LST-UTCI differences. This discrepancy is due to LST being increasingly influenced by roofs as the built fraction increases, primarily reflecting the energy balance at the top of the canopy layer. In contrast, OTC is influenced by the three-dimensional urban form, such as building shading and reduced wind exposure, at the bottom of the canopy layer. Although the cluster analysis reduced the variability of the LST-UTCI relationship within each cluster, substantial variability remains within each cluster. The 95 % confidence intervals range from 7.1 K to 7.5 K across the different clusters (Fig. A2), compared to 10.3 K before clustering. Similar observations are made for LCZ classifications.

Besides the differences in the physical nature of LST measurements and OTC modeling, the native resolution of the thermal infrared sensors must have a strong impact on the LST-OTC relationship. The native resolution is 100 m, and the LST products are downsampled by the emissivity layer to a resolution of 30 m. It must be assumed that the native resolution of 100 m has a strong influence on the representation of intra-urban differences.

In summary, the spatial patterns found in this study are noteworthy as they show that LST maps exaggerate intra-urban differences of OTC, and that surface heat accumulation exceeds its impact on actual OTC. In addition, even within similar urban environments (cluster or LCZ) LST cannot fully capture the high variability of OTC. This implies that LST alone does not capture the intra-urban complexity of OTC, particularly under heat stress conditions. While increasing green space effectively reduces LST, the effect on OTC may be more complex, especially in urbanized areas, where the three-dimensional form plays a crucial role.

4.2. Random Forest modeling

The results of the random forest models also highlight that local-specific information on urban morphology and LCC significantly improves the predictability of the intra-urban variability of OTC. While models relying on coarse-resolution globally available data alone, such as the rf_LST, achieve reasonable classification accuracy, they fail to capture the intra-urban variability of OTC, as model performance varies widely between clusters. As expected, the rf_ERA5 predictions and ERA5-HEAT dataset show no intra-urban variability of OTC. Despite the competitive classification performance of ERA5-HEAT for heat stress conditions, its systematic overestimation leads to poor accuracy in identifying cold or no stress conditions, and its uniform spatial predictions make it unsuitable for fine-scale OTC assessments. The integration of LST with ERA5-Land data in rf_LST + ERA5 slightly improves the model performance, particularly for strong heat stress conditions, as LST provides information at 30 m resolution. However, the issue of OTC intra-urban variability is not adequately addressed. As LST has an almost negative correlation with UTCI in densely built-up areas, the added spatial value is limited. In addition, Landsat Level-2 thermal infrared images have a native resolution of 100 m and are downscaled to 30 m by using the emissivity layer. The native resolution of 100 m may be too coarse to map microscale intra-urban OTC variability. Fahy et al. (2025) developed a PET model that incorporated LST alongside other spatial predictors. They obtained similar results, quantifying the effect of LST on PET as accounting for less than 2 % of the observed variation in PET.

In contrast, models incorporating additional local-specific high-resolution data significantly improve prediction accuracy and consistency. The rf_ERA5 + Urb and rf_All models do not only achieve lower RMSEs, but also provide a more detailed representation of OTC patterns. This improvement is particularly evident in densely built-up areas, where local-specific predictors allow the models to be more effective in distinguishing between the moderate and strong heat stress classes. The observed reduction in RMSE with increasing urbanization further emphasizes the ability of local-specific-informed models to map OTC in complex urban environments. The rf_All model outperforms all other approaches, achieving the highest classification accuracy and the best agreement with the test data of the HTC-NN model. However, the rf_ERA5 + Urb model has almost similar performance, but is not based on LST with fixed time stamps (11:20 CET) and temporally infrequent availability (every 8 days), which raises the question whether LST are necessary to train such a neighborhood scale OTC model. ERA5-Land reanalysis data are available globally at hourly resolution, which would make a model more flexible and better suited to assess OTC for both day and night, as forecasts are not limited to one point in time (11:20 CET local time) during the day. A cost-benefit analysis of whether LST is advantageous for this type of model must conclude that this is not the case.

In summary, as in section 4.1 on the LST-OTC relationship, the results of the random forest modeling highlight the difficulty of reproducing intra-urban variability of OTC by relying solely on LST. Moreover, the results emphasize the need to incorporate local-specific information on urban morphology and LCC, rather than LST, into predictive models in order to accurately represent micro- and neighborhood-scale thermal processes and ensure reliable assessments of OTC.

4.3. Limitations

Although the findings from this study are robust, it is difficult to extrapolate the results to other cities for several reasons. The different background climate conditions, urban morphology, materials, vegetation types, and land cover of Freiburg, are not representative of other cities around the world. This means that although the findings of this study may be transposed for cities with similar morphology and land cover in Central Europe, a transfer to other cities would first need assessments, even though we can expect that physical characteristics of the LST-OTC interplay should be similar worldwide.

It should be noted that all these findings are only valid for 10:20 UTC (11:20 CET local time) at clear sky conditions, and that the LST-UTCI relationship may vary for cloudy conditions, throughout the course of the day and be different at night. For the latter, when UHI is strongest, different processes compared to daytime dominate OTC. The role of T_a becomes more prominent and the impact of radiation on OTC changes, as the effect of radiation trapping induced by trees and buildings is more dominant. Previous studies found that T_a can be better estimated from LST at night than during the day, which might result in higher capability in estimating OTC (Naserikia et al., 2024). However, the relationship between LST and T_{mrt} has not yet been investigated during nighttime conditions. Investigation of the temporal and diurnal dynamics of the LST-OTC relationship, particularly during extreme heat events, would be valuable. However, the temporal availability of satellite based LST data with a spatial resolution higher than 100 m / 30 m is currently limited which makes a comprehensive diurnal investigation of LST as proxy of OTC difficult.

In addition, the spatial resolution of the downsampled LST data of 30 m may not be sufficient to map complex OTC patterns in densely built-up areas. Even at 30 m resolution, Landsat LST data does not represent the microscale processes in urban street canyons that are responsible for street-level OTC variability, especially as most vertical surfaces are neglected which are of importance for the radiation term (T_{mrt}) in OTC calculations. High-resolution thermal sensors would greatly enhance the ability to capture fine-scale spatial variability in urban environments. In particular, a higher spatial resolution would provide a better representation of narrow features, such as street canyons, rooftops and small patches of vegetation. Several upcoming satellite missions aim to provide LST data with a spatial resolution below 5 m. This could help to overcome some of the limitations identified in this study regarding the discrepancy between LST and OTC in complex urban environments.

While off-nadir viewing angles could enhance the depiction of vertical surfaces and facets within urban canyons, such as building walls and surfaces beneath the (tree) canopy, the benefits for building-resolved OTC assessments are unclear. Off-nadir observations introduce the issue of thermal anisotropy, whereby the directional dependence of surface-emitted radiation can lead to inconsistencies and biases in LST retrievals (Lagouarde and Irvine, 2008; Adderley et al., 2015; Du et al., 2023). Therefore, although high-resolution LST sensing is unequivocally beneficial for building-resolved OTC mapping, the value of off-nadir viewing is more questionable due to the complexities it introduces and would need a comprehensive assessment.

5. Conclusion

This study systematically evaluates the validity of using satellite-based land surface temperature (LST), derived from Landsat level 2 products, as a proxy for outdoor thermal comfort (OTC) in a complex urban environment. The spatial and statistical relationship between LST and pedestrian-level OTC, represented by the universal thermal climate index (UTCI), is investigated at multiple scales. We compared LST with high-resolution UTCI predictions from the HTC-NN model (Briegel et al., 2024). To assess the predictive capacity of LST for OTC quantitatively, multiple random forest models are trained using LST combined with various other predictors, such as ERA5-Land reanalysis data, land cover characteristics and urban morphological parameters.

The results clearly demonstrate that LST alone is not a reliable proxy for OTC, particularly in densely urbanized areas. Although a linear relationship can be observed between LST and UTCI in non-heat stress conditions (UTCI <26 °C) and in vegetated or open areas, this relationship breaks down under heat stress, particularly in densely built-up environments. Additionally, intra-urban variability persists even within similar urban typologies (e.g. urban clusters or LCZs). In densely built-up areas, LST is increasingly dominated by rooftop surface temperatures, whereas OTC at pedestrian level is influenced by shading, ventilation, and three-dimensional radiative exchange within the urban canyon. Thus, while LST primarily reflects the energy balance at the upper surface of the canopy layer, OTC integrates multiple microclimatic variables, including air temperature, wind speed, humidity and mean radiant temperature, all of which can be affected by the three-dimensional form of the city. More specifically, the progressive increase in the LST–UTCI difference along the urbanization gradient highlights the significant influence of urban morphology on the spatial heterogeneity of OTC. This causes the LST–UTCI relationship to become nonlinear and spatially inconsistent, resulting in substantial intra-urban variability that cannot be captured by LST alone.

These limitations are further highlighted by the random forest models. Those trained solely on LST or global predictors (e.g. ERA5-Land) fail to capture the heterogeneity of OTC within urban areas, particularly in dense urban areas. It is only when local varying predictors, such as building morphology and land cover, are included that the models achieve meaningful improvements in accuracy and spatial representation of variability.

However, several limitations must be considered. The analysis is restricted to clear-sky, mid-morning conditions (11:20 CET local time) and does not capture the full diurnal cycle of thermal comfort due to the fixed satellite overpass time. The transferability of the results to other cities requires further investigation and evaluation given differences in climate, urban morphology, and land cover. Nevertheless, the physical principles underlying the LST–UTCI relationship are likely to be similar worldwide. Future work should explore the potential of higher-resolution, temporally flexible LST datasets, and assess the diurnal and night-time dynamics of the LST-OTC relationship. At night, LST may better reflect the urban heat island (UHI) effect, and thus OTC conditions.

In summary, LST should not be used as a standalone proxy for OTC in urban areas it has limited ability to capture intra-urban OTC variability and it weakly related to pedestrian-level heat stress in densely built-up areas. To ensure accurate, actionable OTC assessments, predictive models must integrate urban morphology, land cover, and meteorological context.

CRediT authorship contribution statement

Ferdinand Briegel: Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – original draft. **Joaquim G. Pinto:** Supervision, Resources, Funding acquisition, Writing – review & editing. **Andreas Christen:** Supervision, Conceptualization, Writing – review & editing.

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.

Acknowledgments

F.B. is funded by the Helmholtz program “Changing Earth” at KIT. A. C. received funding as part of the ERC Synergy Grant ‘urbisphere’, funded by the European Research Council (ERC-SyG) within the European Union’s Horizon 2020 research and innovation program under grant agreement no. 855005, which also supported the collection of validation data (sensor network) used in this research. J.G.P. thanks the AXA Research Fund for support. The authors would also like to thank Tobias Leismann from constellr for providing data.

Appendix A. Appendix

Table A1
UTCI stress classes and the corresponding UTCI ranges.

Stress class	UTCI range (°C)
Extreme cold stress	< −40
Very strong cold stress	−40 – −27
Strong cold stress	-27 – −13
Moderate cold stress	-13–0
Slight cold stress	0–9
No thermal stress	9–26
Moderate heat stress	26–32
Strong heat stress	32–38
Very strong heat stress	38–46
Extreme heat stress	≥ 46

Table A2
Overview of the dates with valid data for the cluster analysis, the additional dates of overall comparison in summer, and the dates of the rest of the year.

	May	**June**	**July**	**August**	**September**
Cluster analysis	2022-05-04	2021-06-11	2022-07-15	2021-08-14	2021-09-22
	2022-05-20	2022-06-30	2022-07-16	2021-08-21	2022-09-02
	2022-05-29	2023-06-01	2022-07-24	2022-08-09	2023-09-04
		2023-06-25	2023-07-11	2022-08-25	2023-09-12
				2023-08-11	2023-09-20
					2023-09-29
+ Overall comparison (summer)	2023-05-15	2021-06-02	2021-07-20	2022-08-01	2021-09-06
	2023-05-31	2021-06-18	2022-07-08	2022-08-08	2022-09-01
		2022-06-13	2023-07-10	2022-08-16	2023-09-05
		2022-06-21		2022-08-24	
		2022-06-29		2023-08-19	
		2023-06-08		2023-08-20	
		2023-06-09			
		2023-06-16			
		2023-06-17			
		2023-06-24			
+ Overall comparison (remaining seasons)					
January	2022-01-12, 2022-01-13, 2023-01-07				
February	2022-02-05, 2022-02-13, 2023-02-08, 2023-02-09, 2023-02-16				
March	2021-03-07, 2021-03-30, 2022-03-09, 2022-03-10, 2022-03-25, 2022-03-26				
April	2021-04-24, 2022-04-18, 2023-04-05, 2023-04-21, 2023-04-22				
October	2021-10-01, 2021-10-08, 2021-10-24, 2022-10-04, 2022-10-27, 2023-10-07, 2023-10-22				
November	2023-11-23				
December	2022-12-06, 2022-12-07, 2022-12-31				

Table A3
The temperature values give the average of each Local Climate Zone (Demuzere et al., 2022). The values in brackets (±) below the average temperature values are the standard deviations, n describes the number of grid cells in the LCZs.

LCZ	UTCI	LST	Diff
2	28.9 °C	38.2 °C	**9.3 °C**
(*n* = 584)	(± 4.1 °C)	(± 5.6 °C)	(± 3.9 °C)
3	28.9 °C	37.5 °C	8.6 °C
(*n* = 642)	(± 4.0 °C)	(± 5.6 °C)	(± 3.6 °C)
5	29.0 °C	36.0 °C	7.0 °C
(*n* = 2304)	(± 4.0 °C)	(± 5.5 °C)	(± 3.5 °C)
6	29.9 °C	35.1 °C	5.2 °C
(*n* = 26,405)	(± 3.8 °C)	(± 5.7 °C)	(± 3.8 °C)
8	30.9 °C	38.9 °C	8.1 °C
(*n* = 4677)	(± 3.8 °C)	(± 6.0 °C)	(± 4.2 °C)
A	28.1 °C	28.4 °C	**0.3 °C**
(*n* = 5883)	(± 4.2 °C)	(± 5.2 °C)	(± 3.3 °C)
B	30.4 °C	31.9 °C	1.5 °C
(*n* = 4629)	(± 3.7 °C)	(± 5.5 °C)	(± 3.9 °C)
D	31.5 °C	34.4 °C	2.9 °C
(*n* = 1706)	(± 3.4 °C)	(± 6.1 °C)	(± 4.6 °C)

Table A4
Hyperparameter sets of the random forest models with number of trees, mtry: the number of predictors randomly sampled at each split, and min node size: minimum number of data points in a node required to split the node.

Model	Number of trees	mtry	min node size
rf_LST	800	1	9
rf_ERA5	800	4	7
rf_LST + ERA5	400	3	7
rf_LST + Urb	800	2	9
rf_ERA5 + Urb	400	6	1
rf_All	800	5	3

Table A5
Model selection results Root Mean Square Error (RMSE – regression models only), overall accuracy of predicted UTCI classes (ACC), and True Skill Score (TSS) for the different UTCI classes: no thermal stress (8 °C ≤ UTCI <26 °C), cold stress (UTCI ≤8 °C), for all heat stress classes (UTCI ≥26 °C), moderate heat stress (26 °C ≤ UTCI <32 °C), and ≥ strong heat stress (UTCI ≥32 °C).

Model	RMSE (UTCI)	ACC	TSS (no thermal stress)	TSS (cold stress)	TSS (heat stress)	TSS (moderate heat stress)	TSS (≥ strong heat stress)
rf_LST	–	74 %	0.80	0.86	0.86	0.55	0.48
rf_ERA5	–	80 %	0.88	0.86	0.92	0.63	0.58
rf_LST + ERA5	–	81 %	0.88	0.86	0.91	0.63	0.67
rf_LST + Urb	–	77 %	0.82	0.80	0.89	0.60	0.60
rf_ERA5 + Urb	–	83 %	0.90	0.83	0.92	0.69	0.67
rf_All	–	**84 %**	**0.91**	**0.86**	**0.93**	**0.72**	0.70
ERA5-HEAT	3.71 K	77 %	0.83	0.47	0.92	0.45	**0.81**

Table A6
Root Mean Square Errors (RMSE) of UTCI predictions of the machine learning model HTC-NN with 1 m and 30 m resolution and of the different random forest models in comparison with UTCI in-situ measurements. $N = 27$.

UTCI RMSE (K)	
Model Combination	Random Forest Model
rf_LST	4.77 K
rf_ERA5	3.88 K
rf_LST + ERA5	3.89 K
rf_LST + Urb	4.19 K
rf_ERA5 + Urb	2.70 K
rf_All	2.72 K
HTC-NN (1 m)	2.49 K
HTC-NN (30 m)	2.53 K

Table A7
Predictor importance assessed using the Mean Decrease in Impurity (MDI) method provided by the H2O Random Forest algorithm.

Variable	Importance (%)
T_a	37.86 %
LST	28.58 %
Ldown	16.58 %
Kdown	8.70 %
RH	4.46 %
Sky View Factor	2.31 %
Wind	0.73 %
Low vegetation	0.25 %
Trees	0.24 %
Frontal Area Index	0.12 %
Buildings	0.09 %
Paved	0.08 %
Water	< 0.01 %

Table A8
Model selection results by cluster for the remaining models not shown in Table 4 with Root Mean Square Error (RMSE), overall accuracy of predicted UTCI classes (ACC), and True Skill Score (TSS) for the different UTCI classes: no thermal stress (8 °C ≤ UTCI <26 °C), cold stress (UTCI ≤8 °C), for all heat stress classes (UTCI ≥26 °C), moderate heat stress (26 °C ≤ UTCI <32 °C), and ≥ strong heat stress (UTCI ≥32 °C).

		Cluster and sub-cluster				
		(A) Forest and low vegetation area		(B) Built-up and vegetation mix	(C) Densely built-up and industrial area	
		(A.I) Vegetated and water	(A.II) Vegetated and paved		(C.I) Built-up and paved	(C.II) Densely built-up
rf_LST						
RMSE (UTCI)		5.01 K	3.56 K	3.04 K	**3.00 K**	3.85 K
ACC		59 %	74 %	**80 %**	78 %	65 %
	no stress	0.76	0.84	0.86	0.79	0.66
	cold stress	**0.94**	0.84	0.91	0.85	0.74
TSS	heat stress	0.81	0.90	**0.90**	0.85	0.77
	moderate heat stress	0.23	0.62	**0.70**	0.59	0.30
	≥ strong heat stress	0.10	0.34	0.60	**0.73**	0.72
rf_ERA5						
RMSE (UTCI)		3.21 K	3.10 K	2.87 K	**2.64 K**	2.68 K
ACC		78 %	77 %	80 %	**83 %**	80 %
	no stress	0.88	0.83	0.90	**0.92**	0.86
	cold stress	**0.94**	0.82	0.91	0.85	0.74
TSS	heat stress	0.92	0.88	0.93	**0.94**	0.93
	moderate heat stress	0.62	0.57	0.65	**0.68**	0.63
	≥ strong heat stress	0.58	0.56	0.58	0.62	**0.68**
rf_LST + ERA5						
RMSE (UTCI)		3.70 K	2.96 K	2.71 K	**2.39 K**	2.59 K
ACC		81 %	80 %	83 %	**84 %**	78 %
	no stress	0.86	0.85	0.91	**0.92**	0.87
	cold stress	**0.94**	0.82	0.91	0.85	0.74
TSS	heat stress	0.89	0.89	0.93	**0.95**	0.93
	moderate heat stress	0.61	0.61	**0.65**	0.64	0.54
	≥ strong heat stress	0.71	0.69	0.70	0.72	**0.74**
rf_LST + Urb						
RMSE (UTCI)		3.75 K	3.28 K	2.99 K	**2.79 K**	2.92 K
ACC		68 %	75 %	**77 %**	**77 %**	57 %
	no stress	0.82	0.82	**0.83**	0.79	0.76
	cold stress	0.74	0.78	**0.87**	0.83	0.72
TSS	heat stress	0.88	**0.90**	**0.90**	0.87	0.86
	moderate heat stress	0.46	0.60	0.63	0.64	**0.65**
	≥ strong heat stress	0.35	0.50	**0.62**	060	0.52
rf_ERA5 + Urb						
RMSE (UTCI)		2.71 K	2.54 K	2.65 K	2.56 K	**2.48** K
ACC		81 %	80 %	81 %	**82 %**	80 %
	no stress	**0.88**	0.84	0.87	0.85	0.80
TSS	cold stress	**0.95**	0.82	0.92	0.85	0.74
	heat stress	**0.89**	0.86	0.88	0.87	0.84

(*continued on next page*)

Table A8 (*continued*)

	Cluster and sub-cluster				
	(A) Forest and low vegetation area		(B) Built-up and vegetation mix	(C) Densely built-up and industrial area	
	(A.I) Vegetated and water	(A.II) Vegetated and paved		(C.I) Built-up and paved	(C.II) Densely built-up
moderate heat stress	0.62	0.61	0.62	**0.64**	0.61
≥ strong heat stress	0.68	0.67	**0.69**	**0.69**	0.68

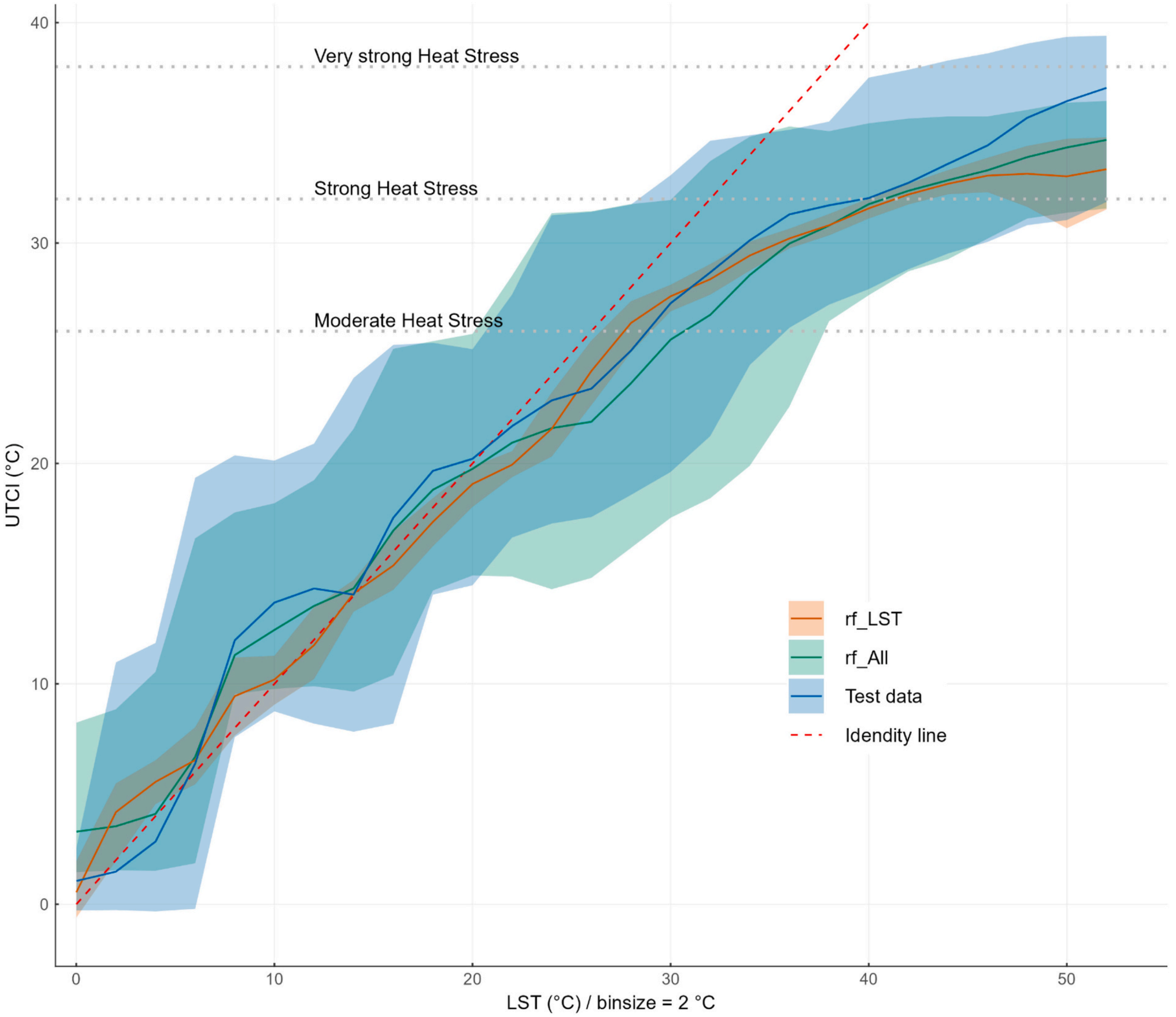

Fig. A1. UTCI as a function of Landsat LST (blue), a simple model forced with LST only (rf_LST, orange), and a model forced with LST, ERA5-Land reanalysis data and data on urban morphology and land cover (rf_All, green) of the random forest modeling test data with a bin size of 2 °C LST. Solid lines show mean values and shaded areas show the 95 % confidence interval. The rf_LST predictions do not represent the wide 95 % confidence interval of the LST-UTCI relationship, whereas the rf_All model better captures this variability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

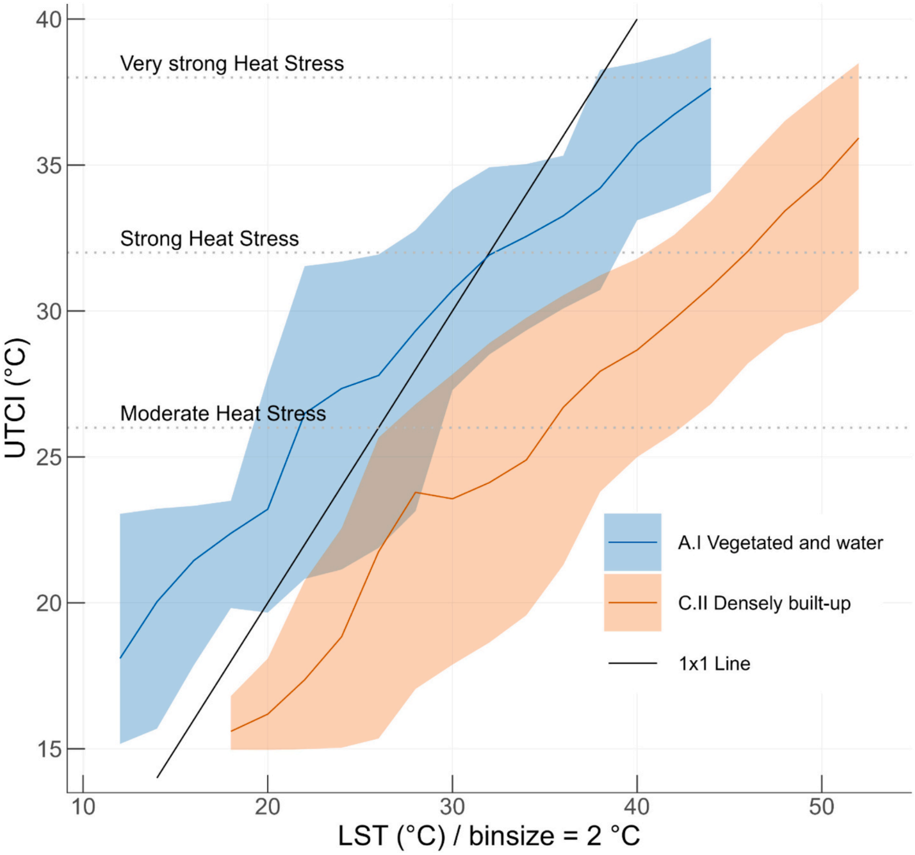

Fig. A2. UTCI as a function of Landsat LST for 2 million summer observations with a bin size of 2 °C LST This is exemplary of clusters A.I. and C.II. The lines show the mean value, and the shaded areas show the 95 % confidence interval. The relationships of the remaining clusters fall between the lines of A.I. and C.II. The average 95 % confidence intervals are 7.1 K and 7.5 K for A.I and C.II, respectively.

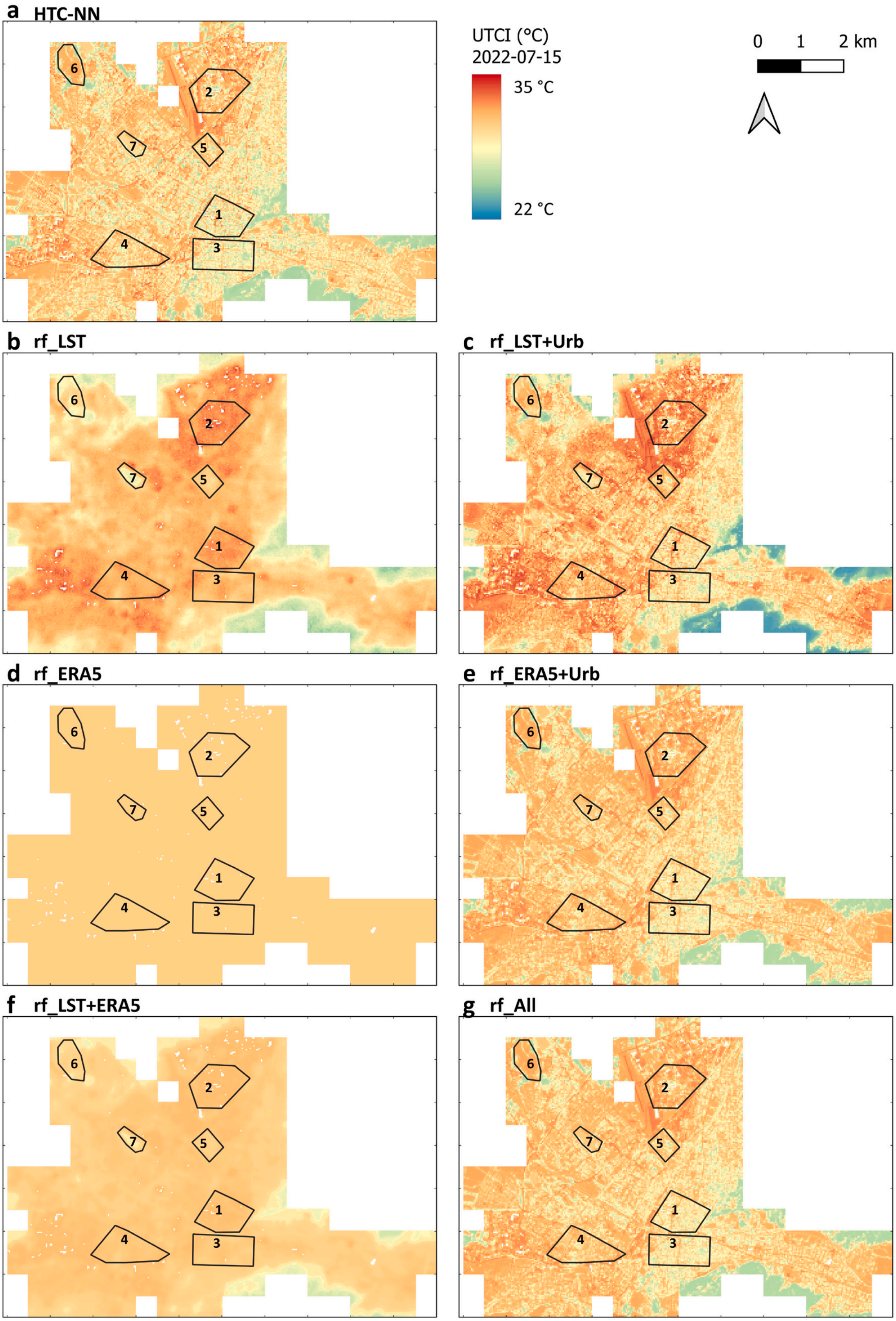

Fig. A3. UTCI prediction at 30 m × 30 m resolution of (a) the HTC-NN (averaged from 1 m), and the different random forest models (b-g) for one exemplary timestep (2022-07-15 11:20:00 local time).

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2025.115045.

Data availability

The HTC-NN and Landsat LST data can be found here: https://doi.org/10.5281/zenodo.17224146.

References

Adderley, C., Christen, A., Voogt, J.A., 2015. The effect of radiometer placement and view on inferred directional and hemispheric radiometric temperatures of an urban canopy. Atmos. Meas. Tech. 8, 2699–2714. https://doi.org/10.5194/amt-8-2699-2015.

Błażejczyk, K., Jendritzky, G., Bröde, P., Fiala, D., Havenith, G., Epstein, Y., Psikuta, A., Kampmann, B., 2013. An introduction to the universal thermal climate index (UTCI). Geogr. Pol. 86, 5–10. https://doi.org/10.7163/GPol.2013.1.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Briegel, F., Christen, A., 2025. Land cover classes for Freiburg. Germany (Dataset). https://doi.org/10.5281/zenodo.14855706.

Briegel, F., Makansi, O., Brox, T., Matzarakis, A., Christen, A., 2023. Modelling long-term thermal comfort conditions in urban environments using a deep convolutional encoder-decoder as a computational shortcut. Urban Clim. 47, 101359. https://doi.org/10.1016/j.uclim.2022.101359.

Briegel, F., Wehrle, J., Schindler, D., Christen, A., 2024. High-resolution multi-scaling of outdoor human thermal comfort and its intra-urban variability based on machine learning. Geosci. Model Dev. 17, 1667–1688. https://doi.org/10.5194/gmd-17-1667-2024.

Cao, B., Liu, Q., Du, Y., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Trigo, I.F., Zhan, W., Yu, Y., Cheng, J., Jacob, F., Lagouarde, J.-P., Bian, Z., Li, H., Hu, T., Xiao, Q., 2019. A review of earth surface thermal radiation directionality observing and modeling: historical development, current status and perspectives. Remote Sens. Environ. 232, 111304. https://doi.org/10.1016/j.rse.2019.111304.

Chakraborty, T., Lee, X., 2019. A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. Int. J. Appl. Earth Obs. Geoinf. 74, 269–280. https://doi.org/10.1016/j.jag.2018.09.015.

Chakraborty, T., Venter, Z.S., Qian, Y., Lee, X., 2022. Lower urban humidity moderates outdoor heat stress. AGU Adv. 3, e2022AV000729. https://doi.org/10.1029/2022AV000729.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014. NbClust: an R package for determining the relevant number of clusters in a data set. J. Stat. Softw. 61. https://doi.org/10.18637/jss.v061.i06.

Cheung, P.K., Jim, C.Y., Hung, P.L., 2021. Preliminary study on the temperature relationship at remotely-sensed tree canopy and below-canopy air and ground surface. Build. Environ. 204, 108169. https://doi.org/10.1016/j.buildenv.2021.108169.

Christen, A., Meier, F., Scherer, D., 2012. High-frequency fluctuations of surface temperatures in an urban environment. Theor. Appl. Climatol. 108, 301–324. https://doi.org/10.1007/s00704-011-0521-x.

Cissé, G., McLeman, R., Adams, H., Aldunce, P., Bowen, K., Campbell-Lendrum, D., Clayton, S., Ebi, K.L., Hess, J., Huang, C., Liu, Q., McGregor, G., Semenza, J., Tirado, M.C., 2022. Health, Wellbeing, and the Changing Structure of Communities. In: Pörtner, H.-O., Roberts, D., Fischlin, A., Howden, M., Méndez, C., Pereira, J.J., Sánchez-Rodríguez, R.A., Semenov, S., Yanda, P., Zatari, T.M. (Eds.), Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK. In Press.

Coutts, A.M., White, E.C., Tapper, N.J., Beringer, J., Livesley, S.J., 2016. Temperature and human thermal comfort effects of street trees across three contrasting street canyon environments. Theor. Appl. Climatol. 124, 55–68. https://doi.org/10.1007/s00704-015-1409-y.

de Almeida, C.R., Teodoro, A.C., Gonçalves, A., 2021. Study of the urban Heat Island (UHI) using remote sensing data/techniques: a systematic review. Environments 8. https://doi.org/10.3390/environments8100105.

de Freitas, C.R., Grigorieva, E.A., 2017. A comparison and appraisal of a comprehensive range of human thermal climate indices. Int. J. Biometeorol. 61, 487–512. https://doi.org/10.1007/s00484-016-1228-6.

Di Napoli, C., Barnard, C., Cloke, H., Pappenberger, F., 2020. Thermal comfort indices derived from ERA5 reanalysis. Copernicus climate change service (C3S) climate data store (CDS) (accessed on 08-01-2025) [WWW document]. 10.24381/cds.553b7518. https://doi.org/10.24381/cds.553b7518.

Di Napoli, C., Barnard, C., Prudhomme, C., Cloke, H.L., Pappenberger, F., 2021. ERA5-HEAT: a global gridded historical dataset of human thermal comfort indices from climate reanalysis. Geosci Data J 8, 2–10. https://doi.org/10.1002/gdj3.102.

Dodman, D., Hayward, B., Pelling, M., Castan Broto, V., Chow, W., Chu, E., Dawson, R., Khirfan, L., McPhearson, T., Prakash, A., Zheng, Y., Ziervogel, G., 2022. Cities, Settlements and Key Infrastructure. In: Pörtner, H.O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA, pp. 907–1040. https://doi.org/10.1017/9781009325844.008.907.

Du, H., Zhan, W., Liu, Z., Scott Krayenhoff, E., Chakraborty, T., Zhao, L., Jiang, L., Dong, P., Li, L., Huang, F., Wang, S., Xu, Y., 2023. Global mapping of urban thermal anisotropy reveals substantial potential biases for remotely sensed urban climates. Sci Bull (Beijing) 68, 1809–1818. https://doi.org/10.1016/j.scib.2023.06.032.

DWD Climate Data Center (CDC), 2024. Multi-annual station means for the climate normal reference period 1981–2010, for current station location and for reference station location.

Earth Resources Observation and Science (EROS) Center, 2020. Landsat 8–9 Operational Land Imager / Thermal Infrared Sensor Level-2, Collection 2 [Dataset]. U.S. Geological Survey.

Epstein, Y., Moran, D.S., 2006. Thermal comfort and the heat stress indices. Ind. Health 44, 388–398. https://doi.org/10.2486/indhealth.44.388.

Fahy, J.C., Bachofen, C., Camponovo, R., Gallinelli, P., Schlaepfer, M.A., 2025. Beyond land surface temperature: identifying areas of daytime thermal discomfort in cities by combining remote sensing and field measurements. Urban Clim. 61, 102460. https://doi.org/10.1016/j.uclim.2025.102460.

Feigel, G., Plein, M., Zeeman, M., Metzger, S., Matzarakis, A., Schindler, D., Christen, A., 2025. High spatio-temporal and continuous monitoring of outdoor thermal comfort in urban areas: a generic and modular sensor network and outreach platform. Sustain. Cities Soc. 119, 105991. https://doi.org/10.1016/j.scs.2024.105991.

Feng, L., Zhao, M., Zhou, Y., Zhu, L., Tian, H., 2020. The seasonal and annual impacts of landscape patterns on the urban thermal comfort using Landsat. Ecol. Indic. 110, 105798. https://doi.org/10.1016/j.ecolind.2019.105798.

Feng, R., Wang, F., Liu, S., Qi, W., Zhao, Y., Wang, Y., 2023. How urban ecological land affects resident heat exposure: evidence from the mega-urban agglomeration in China. Landsc. Urban Plan. 231, 104643. https://doi.org/10.1016/j.landurbplan.2022.104643.

Ferreira, L.S., Duarte, D.H.S., 2019. Exploring the relationship between urban form, land surface temperature and vegetation indices in a subtropical megacity. Urban Clim. 27, 105–123. https://doi.org/10.1016/j.uclim.2018.11.002.

Höppe, P., 1999. The physiological equivalent temperature – a universal index for the biometeorological assessment of the thermal environment. Int. J. Biometeorol. 43, 71–75. https://doi.org/10.1007/s004840050118.

Hundhausen, M., Feldmann, H., Laube, N., Pinto, J.G., 2023. Future heat extremes and impacts in a convection-permitting climate ensemble over Germany. Nat. Hazards Earth Syst. Sci. 23, 2873–2893. https://doi.org/10.5194/nhess-23-2873-2023.

Krayenhoff, E.S., Jiang, T., Christen, A., Martilli, A., Oke, T.R., Bailey, B.N., Nazarian, N., Voogt, J.A., Giometto, M.G., Stastny, A., Crawford, B.R., 2020. A multi-layer urban canopy meteorological model with trees (BEP-tree): street tree impacts on pedestrian-level climate. Urban Clim. 32, 100590. https://doi.org/10.1016/j.uclim.2020.100590.

Krayenhoff, E.S., Broadbent, A.M., Zhao, L., Georgescu, M., Middel, A., Voogt, J.A., Martilli, A., Sailor, D.J., Erell, E., 2021. Cooling hot cities: a systematic and critical review of the numerical modelling literature. Environ. Res. Lett. 16, 53007. https://doi.org/10.1088/1748-9326/abdcf1.

Lagouarde, J.-P., Irvine, M., 2008. Directional anisotropy in thermal infrared measurements over Toulouse city Centre during the CAPITOUL measurement campaigns: first results. Meteorog. Atmos. Phys. 102, 173–185. https://doi.org/10.1007/s00703-008-0325-4.

Li, X., Chakraborty, T., Wang, G., 2023a. Comparing land surface temperature and mean radiant temperature for urban heat mapping in Philadelphia. Urban Clim. 51, 101615. https://doi.org/10.1016/j.uclim.2023.101615.

Li, L., Zhan, W., Hu, L., Chakraborty, T.C., Wang, Z., Fu, P., Wang, D., Liao, W., Huang, F., Fu, H., Li, J., Liu, Z., Du, H., Wang, S., 2023b. Divergent urbanization-induced impacts on global surface urban heat island trends since 1980s. Remote Sens. Environ. 295, 113650. https://doi.org/10.1016/j.rse.2023.113650.

Lindberg, F., Onomura, S., Grimmond, C.S.B., 2016. Influence of ground surface characteristics on the mean radiant temperature in urban areas. Int. J. Biometeorol. 60, 1439–1452. https://doi.org/10.1007/s00484-016-1135-x.

Lindberg, F., Grimmond, C.S.B., Gabey, A., Huang, B., Kent, C.W., Sun, T., Theeuwes, N. E., Järvi, L., Ward, H.C., Capel-Timms, I., Chang, Y., Jonsson, P., Krave, N., Liu, D., Meyer, D., Olofson, K.F.G., Tan, J., Wästberg, D., Xue, L., Zhang, Z., 2018. Urban multi-scale environmental predictor (UMEP): an integrated tool for city-based climate services. Environ. Model Softw. 99, 70–87. https://doi.org/10.1016/j.envsoft.2017.09.020.

Manoli, G., Fatichi, S., Bou-Zeid, E., Katul, G.G., 2020. Seasonal hysteresis of surface urban heat islands. Proc. Natl. Acad. Sci. 117, 7082–7089. https://doi.org/10.1073/pnas.1917554117.

Martilli, A., Roth, M., Chow, W.T.L., Demuzere, M., Lipson, M., Krayenhoff, E.S., Sailor, D., Nazarian, N., Voogt, J., Wouters, H., Middel, A., Stewart, I.D., Bechtel, B., Christen, A., Hart, M.A., 2020. Summer average urban-rural surface temperature differences do not indicate the need for urban heat reduction. https://doi.org/10.31219/osf.io/8gnbf.

Middel, A., AlKhaled, S., Schneider, F.A., Hagen, B., Coseo, P., 2021. 50 grades of shade. Bull. Am. Meteorol. Soc. 102, E1805–E1820. https://doi.org/10.1175/BAMS-D-20-0193.1.

Mijani, N., Alavipanah, S.K., Firozjaei, M.K., Arsanjani, J.J., Hamzeh, S., Weng, Q., 2020. Modeling outdoor thermal comfort using satellite imagery: a principle component analysis-based approach. Ecol. Indic. 117, 106555. https://doi.org/10.1016/j.ecolind.2020.106555.

Mora, C., Dousset, B., Caldwell, I.R., Powell, F.E., Geronimo, R.C., Bielecki, C.R., Counsell, C.W.W., Dietrich, B.S., Johnston, E.T., Louis, L.V., Lucas, M.P., McKenzie, M.M., Shea, A.G., Tseng, H., Giambelluca, T.W., Leon, L.R., Hawkins, E., Trauernicht, C., 2017. Global risk of deadly heat. Nat. Clim. Chang. 7, 501–506. https://doi.org/10.1038/nclimate3322.

Muñoz Sabater, J., 2019. ERA5-land hourly data from 1981 to present. Copernicus climate change service (C3S) climate data store (CDS). (accessed on < 21-11-2024 >) [WWW document]. Doi:10.24381/cds.e2161bac.

Naserikia, M., Nazarian, N., Hart, M.A., Sismanidis, P., Kittner, J., Bechtel, B., 2024. Multi-city analysis of satellite surface temperature compared to crowdsourced air temperature. Environ. Res. Lett. 19, 124063. https://doi.org/10.1088/1748-9326/ad8be4.

Nazarian, N., Krayenhoff, E.S., Bechtel, B., Hondula, D.M., Paolini, R., Vanos, J., Cheung, T., Chow, W.T.L., de Dear, R., Jay, O., Lee, J.K.W., Martilli, A., Middel, A., Norford, L.K., Sadeghi, M., Schiavon, S., Santamouris, M., 2022. Integrated Assessment of Urban Overheating Impacts on Human Life. Earths Future 10, e2022EF002682. https://doi.org/10.1029/2022EF002682.

Oke, T.R., 1973. City size and the urban heat island. Atmos. Environ. 1967 (7), 769–779. https://doi.org/10.1016/0004-6981(73)90140-6.

Oke, T.R., 1982. The energetic basis of the urban heat island. Q. J. R. Meteorol. Soc. 108, 1–24. https://doi.org/10.1002/qj.49710845502.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press. https://doi.org/10.1017/9781139016476.

Patle, S., Ghuge, V.V., 2024. Urban fragmentation approach for assessing thermal environment dynamics: a case study of semi-arid city from a comfort perspective. Urban Clim. 53, 101784. https://doi.org/10.1016/j.uclim.2023.101784.

Peng, S., Piao, S., Ciais, P., Friedlingstein, P., Ottle, C., Bréon, F.-M., Nan, H., Zhou, L., Myneni, R.B., 2012. Surface urban Heat Island across 419 global big cities. Environ. Sci. Technol. 46, 696–703. https://doi.org/10.1021/es2030438.

Plein, M., Feigel, G., Zeeman, M., Dormann, C.F., Christen, A., 2025. Using gradient boosting for gap-filling to analyze temperature and humidity patterns in an urban weather station network in Freiburg. Germany. Urban Clim 62, 102496. https://doi.org/10.1016/j.uclim.2025.102496.

Potchter, O., Cohen, P., Lin, T.-P., Matzarakis, A., 2018. Outdoor human thermal perception in various climates: a comprehensive review of approaches, methods and quantification. Sci. Total Environ. 631–632, 390–406. https://doi.org/10.1016/j.scitotenv.2018.02.276.

R Core Team, 2024. R: A Language and Environment for Statistical Computing.

Renard, F., Alonso, L., Fitts, Y., Hadjiosif, A., Comby, J., 2019. Evaluation of the effect of urban redevelopment on surface urban Heat Islands. Remote Sens (Basel) 11. https://doi.org/10.3390/rs11030299.

Roth, M., Oke, T.R., Emery, W.J., 1989. Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. Int. J. Remote Sens. 10, 1699–1720. https://doi.org/10.1080/01431168908904002.

Saher, R., Stephen, H., Ahmad, S., 2021. Understanding the summertime warming in canyon and non-canyon surfaces. Urban Clim. 38, 100916. https://doi.org/10.1016/j.uclim.2021.100916.

Schwaab, J., Meier, R., Mussetti, G., Seneviratne, S., Bürgi, C., Davin, E.L., 2021. The role of urban trees in reducing land surface temperatures in European cities. Nat. Commun. 12, 6763. https://doi.org/10.1038/s41467-021-26768-w.

Staiger, H., Laschewski, G., Matzarakis, A., 2019. Selection of appropriate thermal indices for applications in human biometeorological studies. Atmosphere (Basel). https://doi.org/10.3390/atmos10010018.

Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93, 1879–1900. https://doi.org/10.1175/BAMS-D-11-00019.1.

Stewart, I.D., Krayenhoff, E.S., Voogt, J.A., Lachapelle, J.A., Allen, M.A., Broadbent, A. M., 2021. Time evolution of the surface urban Heat Island. Earths Future 9, e2021EF002178. https://doi.org/10.1029/2021EF002178.

Stull, R., 2015. Practical Meteorology : An Algebra-Based Survey of Atmospheric Science. Department of Earth, Ocean & Atmospheric Sciences, University of British Columbia. https://doi.org/10.14288/1.0300441.

Thom, E.C., 1959. The discomfort index. Weatherwise 12, 57–61. https://doi.org/10.1080/00431672.1959.9926960.

Venter, Z.S., Chakraborty, T., Lee, X., 2021. Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms. Sci. Adv. 7, eabb9569. https://doi.org/10.1126/sciadv.abb9569.

Voogt, J.A., 2008. Assessment of an urban sensor view model for thermal anisotropy. Remote Sens. Environ. 112, 482–495. https://doi.org/10.1016/j.rse.2007.05.013.

Voogt, J.A., Oke, T.R., 1998. Effects of urban surface geometry on remotely-sensed surface temperature. Int. J. Remote Sens. 19, 895–920. https://doi.org/10.1080/014311698215784.

Voogt, J.A., Oke, T.R., 2003. Thermal remote sensing of urban climates. Remote Sens. Environ. 86, 370–384. https://doi.org/10.1016/S0034-4257(03)00079-8.

Wan, Z., Hook, S., Hulley, G., 2021. MODIS/Terra land surface temperature/emissivity daily L3 global 0.05Deg CMG V061 [data set] NASA EOSDIS land processes distributed active archive center. https://doi.org/10.5067/MODIS/MOD11C1.061.

Wang, C., Zhan, W., Liu, Z., Li, J., Li, L., Fu, P., Huang, F., Lai, J., Chen, J., Hong, F., Jiang, S., 2020. Satellite-based mapping of the universal thermal climate index over the Yangtze River Delta urban agglomeration. J. Clean. Prod. 277, 123830. https://doi.org/10.1016/j.jclepro.2020.123830.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244. https://doi.org/10.1080/01621459.1963.10500845.

Wright, M.N., Ziegler, A., 2017. **Ranger** : a fast implementation of random forests for high dimensional data in *C*++ and *R*. J. Stat. Softw. 77. https://doi.org/10.18637/jss.v077.i01.

Zeeman, M., Christen, A., Grimmond, S., Fenner, D., Morrison, W., Feigel, G., Sulzer, M., Chrysoulakis, N., 2024. Modular approach to near-time data management for multi-city atmospheric environmental observation campaigns. Geoscientific Instrumentation, Methods and Data Systems 13, 393–424. https://doi.org/10.5194/gi-13-393-2024.

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Frolking, S., Yao, R., Qiao, Z., Sobrino, J.A., 2019. Satellite remote sensing of surface urban Heat Islands: Progress, challenges, and perspectives. Remote Sens (Basel) 11. https://doi.org/10.3390/rs11010048.