# Statistical modelling of spatial patterns of the urban heat island intensity in the urban environment of Augsburg, Germany

Annette Straub[a,*], Katja Berger[d], Susanne Breitner[b,e], Josef Cyrys[b], Uta Geruschkat[b], Jucundus Jacobeit[a], Benjamin Kühlbach[a], Thomas Kusch[b,c], Andreas Philipp[a], Alexandra Schneider[b], Robin Umminger[a,c], Kathrin Wolf[b], Christoph Beck[a]

[a] University of Augsburg, Institute of Geography, Augsburg, Germany
[b] Institute of Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany
[c] Environment Science Center, University of Augsburg, Augsburg, Germany
[d] Department of Geography, LMU Munich, Munich, Germany
[e] IBE-Chair of Epidemiology, LMU Munich, Munich, Germany

A B S T R A C T

Spatial and temporal variability of meteorological variables across urban areas due to differences in land surface characteristics is a common phenomenon. Most pronounced is the effect of land cover on air temperature. In this study, parametric and non-parametric statistical approaches (stepwise multiple linear regression, random forests) were applied in order to model sub-daily and daily spatial patterns of the urban heat island intensity in the major city of Augsburg, Southern Germany, and its rural surroundings. A large number of model setups utilizing variables from different land surface data sets as predictors and taking into account different seasonal, daily and meteorological situations was examined. The results were compared concerning different measures of model performance (mean squared skill score, mean squared error, explained variance). For individual setups and situations considerable skill with a mean squared skill score of up to 0.85 was reached. The best performing models were obtained from multiple linear regression for situations with low wind speeds and cloud cover in the morning and evening. Selected models were utilized to derive continuous spatial distributions of the air temperature deviations from a rural reference station. The resulting maps can be useful for various applications, e.g. in the context of urban planning.

## 1. Introduction

In urban areas, air temperature generally reaches higher magnitudes compared to the rural surroundings. This phenomenon, first observed in London by Luke Howard in 1833 (Oke, 1982), and since then known as the urban heat island (UHI) has been investigated thoroughly for several decades (Tyson et al., 1972; Arnfield, 2003; Stewart, 2011; Deilami et al., 2018). However, spatial air temperature differences are not only observable between urban and rural locations but also occur within cities (e.g. Spronken-Smith and Oke, 1998; Fenner et al., 2017; Skarbit et al., 2017; Beck et al., 2018).

These differences in thermal characteristics correspond to varying land cover types and are a result of variations in the energy balances of these surfaces (Oke, 1982). The energetic properties of artificial urban surfaces differ from those of natural surfaces due to

* Corresponding author at: Institute of Geography, University of Augsburg, Alter Postweg 118, D - 86159 Augsburg, Germany.
*E-mail address:* annette.straub@geo.uni-augsburg.de (A. Straub).

distinctions in thermal conductivity, heat capacity, water permeability and albedo (Oke, 1988). Thus, intra-urban vegetation such as single trees, ground vegetation, green roofs, gardens, forests or parks can reduce urban air temperatures, e.g. by shading and increasing the latent heat flux (Oláh, 2012; Bowler et al., 2010). Sufficiently large vegetated areas tend to generate a so-called park cool island (PCI) with lower air temperatures compared to the adjacent built-up areas, sometimes affecting not only the green space itself but also its surroundings (e.g. Jáuregui, 1975; Upmanis et al., 1998; Wang et al., 2018). Another cause for the formation of an UHI are the geometric features of buildings, specified by e.g. the height-to-width ratio of urban street canyons or the sky view factor (Oke, 1981; Unger, 2004). Moreover, anthropogenic heat emissions alter the urban energy balance to a larger extent than the rural one (Oke, 1988; Hinkel et al., 2003). The UHI intensity follows a distinct daily as well as annual cycle with highest values usually occurring several hours after sunset during summer in temperate climates (e.g. Erell and Williamson, 2007; Wienert et al., 2013; Arnds et al., 2017) and during the dry season in tropical climates (e.g. Chow and Roth, 2006; Roth, 2007; Jongtanom et al., 2011). Besides, mesoscale conditions can alter this distinct periodicity, e.g. sea breeze can reduce summerly UHI intensities in coastal cities (e.g. Kim and Baik, 2002; Kim and Baik, 2004). During summer daytime, the formation of an urban cool island is possible with the occurrence of lower urban air temperatures compared to the rural surroundings (Oke, 1993; Yang et al., 2017). Besides, UHI intensities greatly depend on the meteorological conditions where cloudless and calm weather favours the development of large temperature differences between urban and rural areas (e.g. Oke, 1982; Morris et al., 2001; Alonso et al., 2007). Wind, on the other hand, diminishes UHI intensities (e.g. Oke, 1982; Morris et al., 2001; Konarska et al., 2015) and can cause a relocation of the areas with elevated temperatures (Gedzelman et al., 2003; Vicente-Serrano et al., 2005).

Knowledge of the spatially and temporally varying thermal conditions within urban areas is becoming increasingly important especially as advancing urbanization results in an increase of UHI intensities (Pauleit et al., 2005; Kim and Baik, 2004) also discussed within the concept of urban heat archipelagos (e.g. Unger, 2004). Accompanied by current climate change this causes additional heat stress in urban areas (Chapman et al., 2017; Wouters et al., 2017). Elevated outdoor air temperatures have an impact on human thermal comfort, but can also have more severe consequences as they contribute to an increase of heat induced morbidity and mortality (e.g. Basu, 2009; Guo et al., 2014). Therefore, the results of spatio-temporal models for air temperature are of great relevance e.g. for urban planning (e.g. Goosen et al., 2014), the selection of appropriate heat mitigation strategies (e.g. Kim et al., 2018) and the calculation of human-biometeorological indices (e.g. Matzarakis et al., 2008).

Numerous studies used different statistical methods to estimate spatio-temporal UHI development with temporal analysis focusing on seasonality and the daily cycle (e.g. Chow and Roth, 2006; Jongtanom et al., 2011), the influence of weather conditions (e.g. Morris et al., 2001; Kim and Baik, 2004; Alonso et al., 2007; Hoffmann et al., 2012) or changes due to urbanization (e.g. Pauleit et al., 2005; He et al., 2007). Spatial patterns were related to land cover characteristics (e.g. Eliasson and Svensson, 2003; Suomi and Käyhkö, 2012; Konarska et al., 2015; Tong et al., 2018). Multiple linear regression (MLR) based models were applied frequently and often reached considerable goodness of fit as shown e.g. by Unger et al. (2001) for Szeged, Hungary, Bottyán et al. (2005) for Debrecen, Hungary, Alcoforado and Andrade (2006) for Lisbon, Portugal, Szymanowski and Kryza (2012) for Wrocław, Poland, Buttstädt and Schneider (2014) for Aachen, Germany, Heusinkveld et al. (2014) for Rotterdam, Netherlands, Ketterer and Matzarakis (2015) for Stuttgart, Germany as well as Bernard et al. (2017) for Nantes, Angers and La Roche-sur-Yon in western France. While MLR belongs to the linear parametric approaches and is a well-established method in climate science (Wilks, 2006), another, relatively new statistical approach successfully utilized to derive spatio-temporal patterns of air temperature were Regression Trees, e.g. Random Forests (RF). They were introduced by Breiman (2001) and are a non-parametric machine learning method, which is able to model non-linear processes (Breiman, 2001; Hastie et al., 2009). RF often result in high model performance, when applied e.g. by Hart and Sailor (2009) for Portland, Oregon, Hjort et al. (2011) for Turku, Finland, Ho et al. (2014) for Vancouver, Canada, and Makido et al. (2016) for Doha, Qatar, which makes them a promising approach.

An initial study by Beck et al. (2015) showed stepwise MLR and RF to be suited for modelling spatial patterns of intra-urban air temperature in the urban area of Augsburg, Southern Germany. The aim of the present study is the evaluation and comparison of various setups of stepwise MLR- and RF-based statistical models used for estimating spatial patterns of mean hourly and mean daily air temperature deviations in the city of Augsburg and its surroundings from a rural reference station. A large set of land use/land cover characteristics is available as predictors for the models. Different situations discriminated according to season, time of the day and synoptic conditions are taken into account. Finally, selected models are transferred to a regular grid to obtain continuous spatial air temperature distributions.

In Section 2 the study area is characterized, a description of the available data is given and the methodology is described. A presentation of the results follows in Section 3, their discussion in Section 4. Finally, chapter 5 draws some main conclusions.

## 2. Material and methods

### 2.1. Study area

Augsburg is located in Bavaria, Southern Germany, (48.37°N; 10.90°E) (Stadt Augsburg, 2015) on an average elevation of 485 m a.s.l. (LDBV 2016). It has 288.653 inhabitants (effective December 2016, Stadt Augsburg, 2017) or 372.422, respectively, if the directly neighbouring municipalities of Königsbrunn, Stadtbergen, Gersthofen and Friedberg are included (Stadt Königsbrunn, 2018, Stadt Stadtbergen, 2018, Stadt Gersthofen, 2018, Stadt Friedberg, 2018). The elevation within the region of Augsburg varies by 89 m (LDBV, 2016). The study region covers an area of $9 \times 10\,km^2$ (Fig. 1). The city centre and as well some parts near the outskirts are characterized by built-up areas with percentages of sealed surface of > 80% (based on the European Urban Atlas, cf. Section 2.3). Two rivers, framed by vegetated areas, traverse the city from south to north. A large forested area is situated to the southwest of

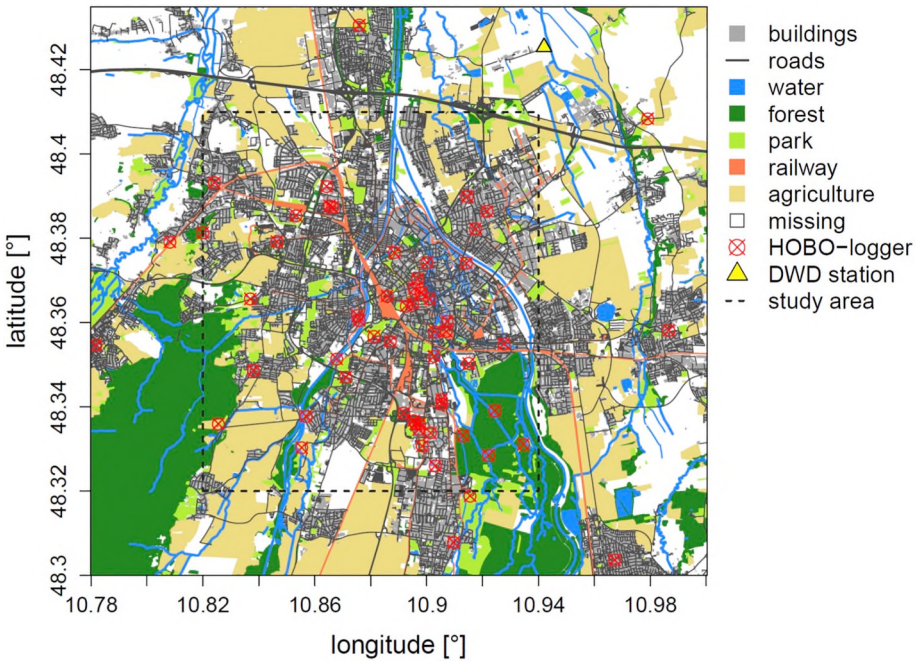


**Fig. 1.** The city of Augsburg and its surroundings, including the study area (within the box with dashed border) with locations of HOBO logger stations as well as the rural reference station ("DWD station") and coverages of selected land surface categories. White areas indicate missing land cover information (land surface data: Open Street Map).

Augsburg in the rural surroundings. Another forest, most of which is included in the study area, lies in the southeast of the city. Fig. 1 shows the setting using land surface data from Open Street Map (cf. Section 2.3). A map of the European Urban Atlas data (cf. Section 2.3) for the same area is presented in Appendix A.

According to the Köppen-Geiger climate classification, Augsburg belongs to type *Cfb*, which corresponds to a warm temperate climate with humid conditions throughout the year (Kottek et al., 2006). Annual mean temperature amounts to 8.4 °C with maximum values in July. Measured UHI intensities in the most densely built-up areas of the study region reach mean values of up to 5 °C in summer for calm and clear nocturnal situations when highest air temperature differences between distinct environments tend to occur (Beck et al., 2018). The most frequent wind direction is southwest.

### *2.2. Meteorological data*

The meteorological database includes a measurement network of HOBO pro v2 U23–001 data loggers (Onset, 2010) recording air temperature and relative humidity with a temporal resolution of four minutes. It is operated in collaboration of the Institute of Epidemiology, Helmholtz Zentrum Munich, German Research Center for Environmental Health and the Institute of Geography, University of Augsburg. The network has been running since December 2012 and comprises altogether 82 measurement sites, currently 49 of them in operation. Data of 77 stations in the period from December 2012 until January 2017 were utilized for the analysis presented here. The measurement sites are located in the urban area of Augsburg and its rural surroundings and represent different types of surface cover and land use (cf. Fig. 1, Appendix B). The instruments are mounted in measurement heights between 1.5 m and 2.3 m above the ground due to reasons of feasibility and secureness (Beck et al., 2018). The data have been quality checked via a validation procedure including a fixed range test and persistence test (Estévez et al., 2011), a step test (Shafer et al., 2000) as well as tests for spatial and temporal outliers (Gandin, 1988). Here, "spatial outliers" refer to values, which are unrealistic with regard to the measurements at other stations within the network at the same time, and "temporal outliers" are unrealistic values with regard to the season in which they were recorded. C.f. Beck et al. (2018) for a more detailed description of the conducted tests. For the further procedure, the data values, which passed all tests, were aggregated to hourly and daily mean values. There are large differences in data availability between the stations, ranging from 5% to 95% of missing observations with an average availability of 50% for the hourly values. This is due to the fact that some stations were relocated or operated for only several months during the study period, while nearly continuous measurements are available at other sites (Beck et al., 2018). For information on the characteristics of the individual measurement sites and their surroundings see Appendix B.

A station from the German Meteorological Service (DWD), providing hourly and daily meteorological data, has been used as rural reference station for the calculation of mean hourly and mean daily (0 a.m. to 12 p.m.) temperature deviations of the HOBO loggers. Besides, hourly and daily observations have been utilized for the characterization of different synoptic conditions based on wind speed and cloud cover. The station is situated north of Augsburg at the airport of Mühlhausen (48.4254°N; 10.9420°E) (DWD Climate

Data Centre, 2017). The measurement site is surrounded by agricultural land. According to the local climate zone (LCZ) concept by Stewart and Oke (2012), the station lies within the LCZ "low plants" (Beck et al., 2018).

Furthermore, air temperature data from automatic weather stations situated at the University and the University of Applied Sciences of Augsburg were available. They were used for comparisons of the HOBO logger data at these respective sites with another, more reliable device, in order to estimate the quality of the HOBO logger data in addition to the previously mentioned quality control. The results show reasonably well agreements between HOBO loggers and the automatic weather stations, with a bias of 0.26 °C at the University and 0.04 °C at the University of Applied Sciences and a slight underestimation of the temperature maxima by the HOBO loggers (cf. Appendix C). Thus, the comparison confirms that the HOBO loggers provide a suitable data basis for this study.

### 2.3. Land surface characteristics

Data on surface characteristics originate from the open access databases of Open Street Map (OSM) (Geofabrik, 2016) and European Urban Atlas (UA) (EEA, 2010), respectively. The OSM data is based on crowd sourcing and thus, it consists of information from mapping with GPS devices as well as freely available satellite data and maps. Due to the dependency on the contribution of volunteers, the horizontal resolution, accuracy and completeness are spatially inhomogeneous (Haklay and Weber, 2008). A free version, which is updated daily but does not contain all layers, is available as vector data (Geofabrik, 2016). Areas with missing information in the OSM dataset amount to approximately 7.8% of the study area (cf. Fig. 1). The UA data are provided by the European Environment Agency (EEA) for European urban areas with > 100,000 inhabitants in the form of vector data. They are based on remote sensing data, topographic maps as well as street navigation data and soil sealing data of the EEA. The UA version utilized in this study is based on data from the years 2005 and 2009, respectively, and fully covers the study region. The size of the smallest mapped element amounts to 0.25 ha (EEA, 2011; EEA, 2016).

Different land use/land cover characteristics from both data sets that are presumably relevant for the spatial air temperature distribution were taken into account. For OSM this comprises the categories "buildings", "roads", "railway", "water", "parks" (urban green spaces), "forest", "meadow" and "agriculture". For UA five categories of "urban fabric" characterizing different percentages of sealed area (> 80%, 50% - 80%, 30% - 50%, 10% - 30% and < 10%, respectively), "Roads", "Railways", "Industrial, commercial, public, military and private units", "Agricultural areas, semi-natural areas and wetlands", "Forest", "Water" and "Green urban areas" were considered. Area percentages of these land use/land cover types were calculated using quadratic buffer zones of varying sizes (edge lengths of 500 m, 250 m, 100 m, 50 m and 25 m, respectively) around the logger stations as well as for a regular grid with spatial resolution of 50 m and additionally for presumed source regions of advective air flow. These source regions were defined as quadratic zones southwest of a given logger location in accordance with the most frequent wind direction in Augsburg, with the logger lying in the north-eastern corner of the respective buffer zone. Thereby, varying edge lengths of 500 m, 250 m, 100 m, 50 m and 25 m, respectively, were applied, too.

The UA data were used to calculate the additional variables "distance to areas with potential cooling effect" defined as water-bodies and vegetated areas as well as "distance to the city centre". As the city of Augsburg is not concentric in shape and the density of built-up areas does not decrease continuously from the centre to the edge of the city, a single central area could not be defined. Instead, densely built-up polygons of considerable size were detected within the centre of each city district via the coverage by the UA land cover class "continuous urban fabric (sealed layer > 80%)" and defined as "densely built-up areas". The variable "distance to the city centre" was calculated as the distance of each HOBO logger or grid cell, respectively, to the nearest "densely built-up area".

Furthermore, the sky view factor (SVF) was calculated from fish-eye photographs for the HOBO logger stations in operation during summer 2017 utilizing the RayMan 1.2 software (Matzarakis et al., 2007) in order to consider geometric features. The SVF is available for 48 stations, which corresponds to 62% of the stations used in this study. Finally, the surface elevation was included, which originates from a digital terrain model from the Bavarian State Office for Survey and Geoinformation with a spatial resolution of 25 m and an accuracy of 2–3 m (LDBV 2016).

### 2.4. Statistical modelling approaches and model setups

#### 2.4.1. Modelling approaches

Two different statistical modelling approaches were applied in order to derive spatial patterns of air temperature deviations from the rural reference station for the study area based on a large number of explanatory variables. These predictors are spatial variables, some of which are included in the form of area percentages by utilization of different buffer zone sizes. A result of that is a close correlation among some of the independent variables. In order to reduce this multicollinearity in the predictor data set and thus the impact of overfitting, a pre-selection of the land use/land cover categories available in different buffer zone sizes was carried out. For this purpose, correlations between air temperature deviations from the reference station and the area percentages of these predictors around the logger sites were calculated and tested for significance ($\alpha = 0.05$ for MLR and $\alpha = 0.1$ for RF, respectively, due to the – according to Breiman (2001) – robustness of RF when including relatively weak predictors). For each land cover type, only the buffer zone size variant with highest significant correlation to the predictand was included in the fitting procedure, types with no significant correlation were omitted, which affected every type in at least one model setup.

Multiple linear regression is a parametric approach that fits a linear equation, which best represents the relation between the independent variables and the predictand, based on the minimization of the sum of the squared residuals (Wilks, 2006). The MLR models were fitted using the R software. Here, a stepwise approach for predictor selection, including both forward selection and backward elimination in combination, was applied, based on the minimization of the Akaike Information Criterion (AIC) (Akaike, 1974).

**Table 1**
Characterization of different modelled situations regarding season, time of day and meteorological conditions.

| Differentiation according to temporal aspects | | | | |
|---|---|---|---|---|
| Season | | | | |
| spring: MAM, summer: JJA, autumn: SON, winter: DJF | | | | |
| Times of day | | | | |
| Variant 1): mean values for fixed time spans [UTC + 1] | | | | |
| morning: 4:00–7:00, noon: 12:00–15:00, evening: 20:00–23:00 | | | | |
| Variant 2): times of day or mean values for time spans, respectively, varying according to the season [UTC + 1] | | | | |
| | MAM | JJA | SON | DJF |
| Morning | 6:00 | 4:00 | 6:00 | 7:00 |
| Noon | 12:00 | 13:00 | 12:00 | 12:00 |
| Evening | 20:00 | 21:00–22:00 | 20:00 | 18:00–19:00 |
| Variant 3) daily mean values | | | | |
| Meteorological situations (including consecutive numbering) based on wind speed [$ms^{-1}$] and cloud cover [octas] | | | | |
| | | Wind speed | | |
| | | $< 2\,ms^{-1}$ | $2–7\,ms^{-1}$ | $> 7\,ms^{-1}$ |
| Cloud cover | 0–3 octas | (1) Calm and clear | (3) Moderate and clear | (5) Windy and clear |
| | > 3 octas | (2) Calm and cloudy | (4) Moderate and cloudy | (6) Windy and cloudy |

Random Forests is a non-parametric, non-linear machine learning approach, based on a large number of uncorrelated, equally distributed decision trees (Breiman, 2001). In this study, the RF models were fitted including 1500 regression trees each. The number of variables used as candidates for each split was determined by an optimization algorithm from the R package "randomForest" (Liaw and Wiener, 2002). It searches for the optimal value starting from the default value of one third of the number of input variables. The minimum size of the terminal nodes was set to 5, which is the default for regression trees (Liaw and Wiener, 2002).

Missing land use/land cover information are treated as a coverage of 0% of the respective type. Thus, for areas with incomplete data the area percentages within the quadratic buffer zones around the logger sites as well as the pixels of the grid for the whole model region do not sum up to 100%. Consequently, for these areas the modelled air temperatures are based only on those variables which are both available area wide (e.g. elevation) and included in the model equation.

### *2.4.2. Model setups*

The statistical models were fitted for different situations separately, taking into account the meteorological season (March to May

**Table 2**
Overview of different model setups regarding the predictors included in the fitting procedure as well as the temporal resolution of the resulting model (sub-daily or daily, respectively). The two columns to the right give the denotation used for the model variants in the presentation of the results (abbreviations: OSM: Open Street Map, UA: Urban Atlas, SVF: Sky view factor).

| Model setups | OSM-based | UA-based |
|---|---|---|
| Setup including all predictors, based on sub-daily mean values for selected (fixed) times of day | "OSM" | "UA" |
| Setup including all predictors, based on sub-daily mean values accounting for seasonal variation | "OSM adj. time" | "UA adj. time" |
| Setup without SVF, based on sub-daily values for selected (fixed) times of day | "OSM w/o SVF" | "UA w/o SVF" |
| Setup summarizing selected land use/land cover categories, based on sub-daily values for selected (fixed) times of day | "OSM built-up" | – |
| Setup utilizing land use/land cover parameters (all predictors) of source areas, based on sub-daily values for selected (fixed) times of day | "OSM source" | – |
| Setup including all predictors, based on daily mean values | "OSM daily" | "UA daily" |

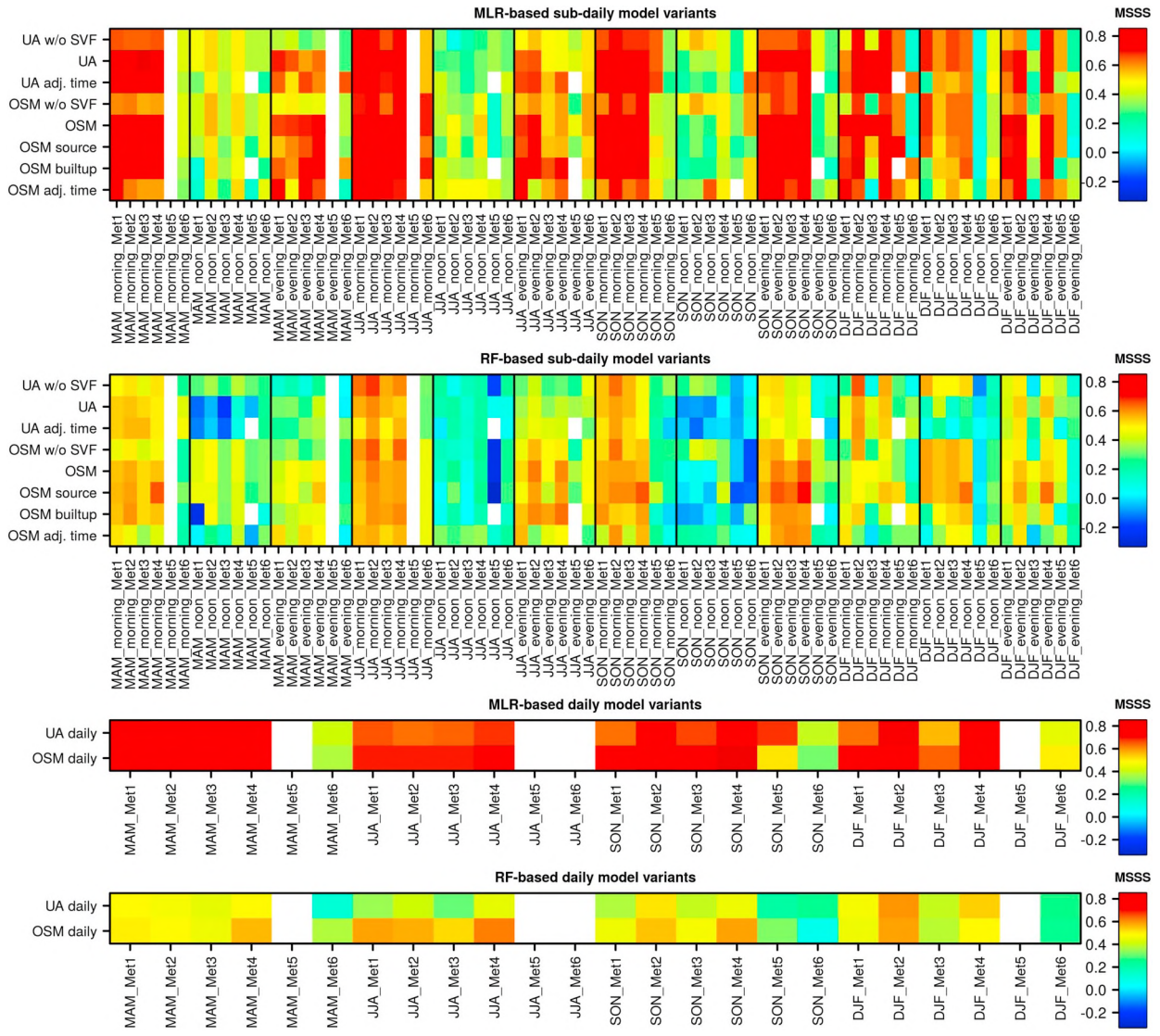


**Fig. 2.** MSSS for all MLR-based and RF-based sub-daily and daily model variants (fitted models), seasonal subsets (MAM, JJA, SON, DJF), times of day (morning, noon, evening) and synoptic conditions (Met1: calm and clear, Met2: calm and cloudy, Met3: moderate wind speed and clear, Met4: moderate wind speed and cloudy, Met5: windy and clear, Met6: windy and cloudy). White signatures indicate that models could not be fitted in these situations due to missing temperature observations.

("MAM"), June to August ("JJA"), September to November ("SON") and December to February ("DJF")), time of day and meteorological conditions (cf. Table 1). While it is also common to include meteorological parameters as temporal explanatory variables (e.g. Shi et al., 2018), a similar stratification of the input data in order to derive models for different situations has been successfully applied in other studies (e.g. Eliasson and Svensson, 2003). In this study, the selected times of day were morning (before sunrise), noon (maximum solar elevation) and evening (after sunset). In different model setups, times of day were differentiated utilizing either mean values for fixed time spans, varying times according to seasonal variations of the hour of sunrise, maximum solar elevation and sunset or daily mean values, respectively. Six different synoptic situations reaching from "calm and clear" to "windy and cloudy" were defined referring to wind speed and cloud cover from the rural reference station. This classification was based on hourly observations for the sub-daily models and on daily observations for the daily models. The division into seasonal subsets as well as the threshold values for the times of day and the weather conditions (cf. Table 1) were derived from the studies by Kuttler et al. (1994), Eliasson and Svensson (2003), Alonso et al. (2007) and Wienert et al. (2013). The upper threshold value for the wind speed results from the empirical equation by Wienert et al. (2013), which calculates a wind speed above which an UHI is not expected to develop based on the number of inhabitants of the city. The resulting models represent estimations of average air temperature deviations between the logger stations and the reference station for each of the different situations. This stratification results in 72 sub-daily (four seasons, three times of the day, six meteorological classes) and 24 daily situations (four seasons, six meteorological

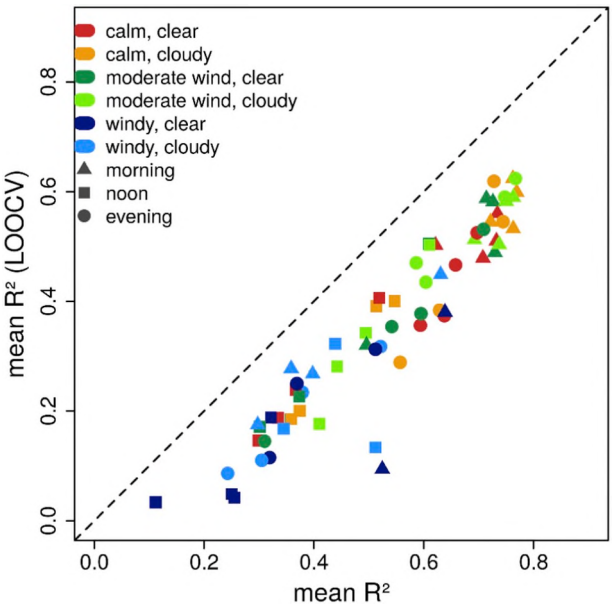


**Fig. 3.** Mean $R^2$ and $R^2$ from leave-one-out cross-validation over all model setups and seasons for MLR-based models.

classes), for each of which the various model setups described in the following passage were tested.

Different model variants including selected, predefined subsets of the independent variables were evaluated for each of the situations described above. Variants incorporating all available types of predictors were fitted separately using OSM- and UA-based land use/land cover data, respectively. A further setup considered all predictors but SVF, because this variable is only available for part of the logger locations and thus cannot be used for estimating continuous spatial air temperature patterns for the whole study area. Another variant included the summarized land cover classes "built-up area" composed of the OSM-based land cover classes "buildings", "roads" and "railway" as well as "low vegetation" composed of "meadow" and "agriculture" in addition to the other OSM-based predictors. Furthermore, a model setup using land cover properties of source regions of advective air flow instead of on-site characteristics was tested for situations with a wind direction between 180° and 270° according to the definition of the source areas (cf. Table 2).

#### *2.4.3. Assessment of model performance and predictor importance*

In this study, the model performance of the fitted models was assessed via different scores, e.g. the explained variance ($R^2$), which is the square of the Pearson correlation coefficient (Wilks, 2006). Besides, the mean square error (MSE) was utilized. It is calculated from the square of the deviations between observed and modelled values. Thus, it takes positive values, where smaller values indicate higher model quality. A drawback of the MSE is its sensitivity to outliers (Wilks, 2006). Another way to assess the model performance are skill scores, which describe the accuracy of a model compared to a reference, e.g. the climatological mean. They become unity for a perfect model and take negative values, if the reference fits better than the respective model. Here, the mean square skill score (MSSS) (Murphy, 1988) was calculated. Furthermore, $R^2$ from leave-one-out cross validation (LOOCV) was calculated for MLR. As MLR is sensitive to overfitting, cross validation is a common method to evaluate the model performance for test data in comparison with the training data. Small differences between $R^2$ of the fitted models and $R^2$ from cross validation indicate good transferability of the model equation to test data and thus confirm the validity of the models (Wilks, 2006).

The relative importance of individual predictors was examined using averaging over orderings for MLR. This method decomposes the explained variance of a model into the contributions of individual predictors to the overall explained variance. It can also be interpreted as the decrease of $R^2$ of a model if the respective predictor was not included (Grömping, 2006; Grömping, 2007). For RF-based models, the predictor importance was assessed utilizing the permutation importance, which is calculated from data excluded from the fitting procedure (out-of-bag data). It is given as the increase of the MSE in percent and can be interpreted as the deterioration of the model if the respective predictor was not included (Breiman, 2001).

## 3. Results

### *3.1. Model performance*

The following section presents the performance of the resulting statistical models using the MSSS, the MSE, and in the case of MLR-based approaches, also the difference between $R^2$ and $R^2$ from LOOCV. At first, differences in the quality of models for different seasons, times of day and weather conditions are shown, then the differences between model setups with regard to predefined

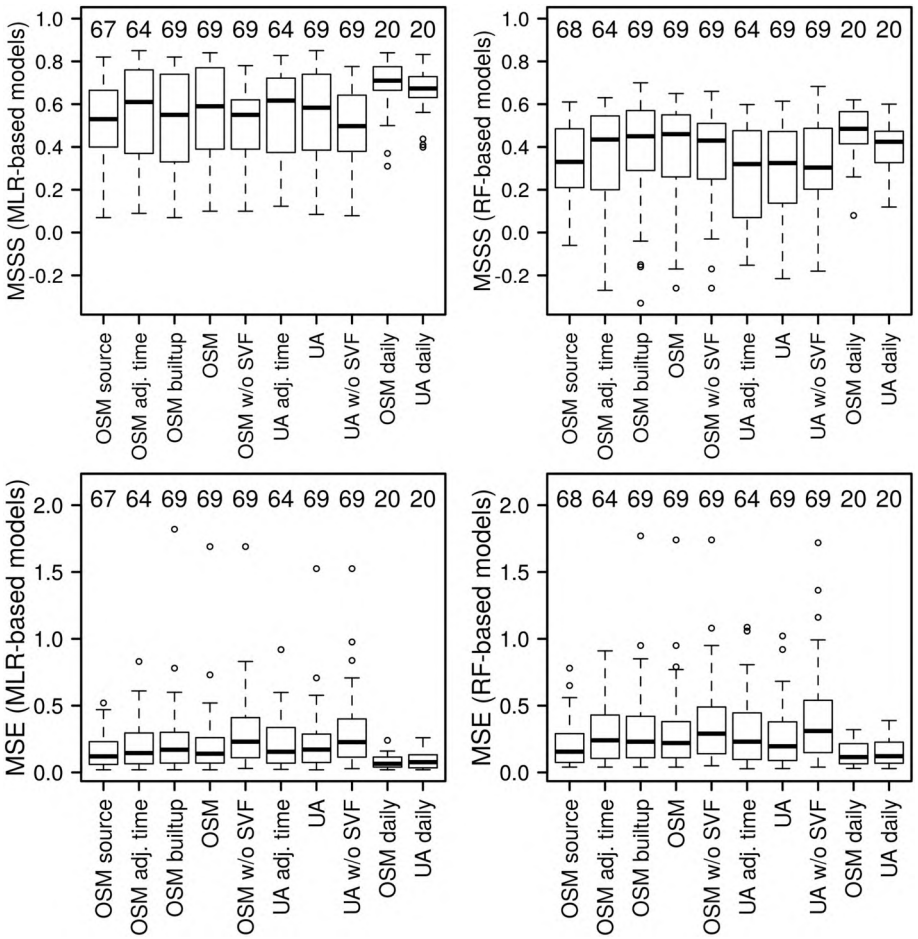


**Fig. 4.** Box-whisker plots of performance measures for MLR- and RF-based model variants: MSSS and MSE [°C] for different OSM- and UA-based model setups as well as for setups utilizing daily and sub-daily values, respectively. The boxes show the median as well as the 1st and 3rd quartile (lower/upper boundary of the box, respectively), whiskers indicate the 1.5-fold interquartile range from the upper/lower boundary of the box. Each box contains the model variants for different seasons (MAM, JJA, SON, DJF) as well as meteorological situations (Met 1 – Met 6) and in the case of sub-daily models different times of day (morning, noon, evening), too. Numbers above the boxes indicate the sample size (number of models) contained in each box.

predictor combinations and temporal resolution are considered. In general, both MLR- and RF-based models show considerable model performances for individual situations and setups (Fig. 2).

*3.1.1. Differences between seasons, times of the day and weather conditions*

Fig. 2 shows the MSSS for all model variants and situations as well as both statistical approaches. For both, MLR and RF, the models with highest performance for sub-daily as well as daily variants occur for low to intermediate wind speeds, especially for situations in spring, summer and autumn. In winter, highest MSSS are reached also for low to intermediate wind speeds, while models for situations with high wind speeds are among those with lowest performance. Differences between clear and cloudy situations are less pronounced. The MSSS for sub-daily variants shows that the best models are obtained for morning and evening, whereas for noon the resulting values of the MSSS are lower (Fig. 2). For some situations, no models could be fitted due to missing air temperature observations within the respective classes.

For sub-daily models, values of the MSSS up to 0.85 for MLR and 0.7 for RF are reached in the fitting procedure. For the daily variants, similar maxima of the MSSS are obtained (Fig. 2). As visible in Fig. 2, not only the MSSS values of the best-performing models are lower for RF than MLR, but also they are lower for most of the other situations and setups. Besides, the best MLR-based models have $R^2$ of up to 0.85 (Appendix D) and MSE of minimum values of 0.02 °C, for RF-based models $R^2$ reach up to 0.71 and the best MSE values amount to 0.03 °C (not shown). Thus, model performance is altogether worse for RF- than for the MLR-based approaches, but still robust models for different situations could be obtained (Fig. 2).

For MLR, the differences between $R^2$ from the fitting procedure and $R^2$ from leave-one-out cross validation are presented in Fig. 3. Mean values of these performance measures of different model setups and seasons confirm that the best models are obtained for situations in the morning and evening with low to moderate wind speed and varying cloudiness. $R^2$ from LOOCV is altogether

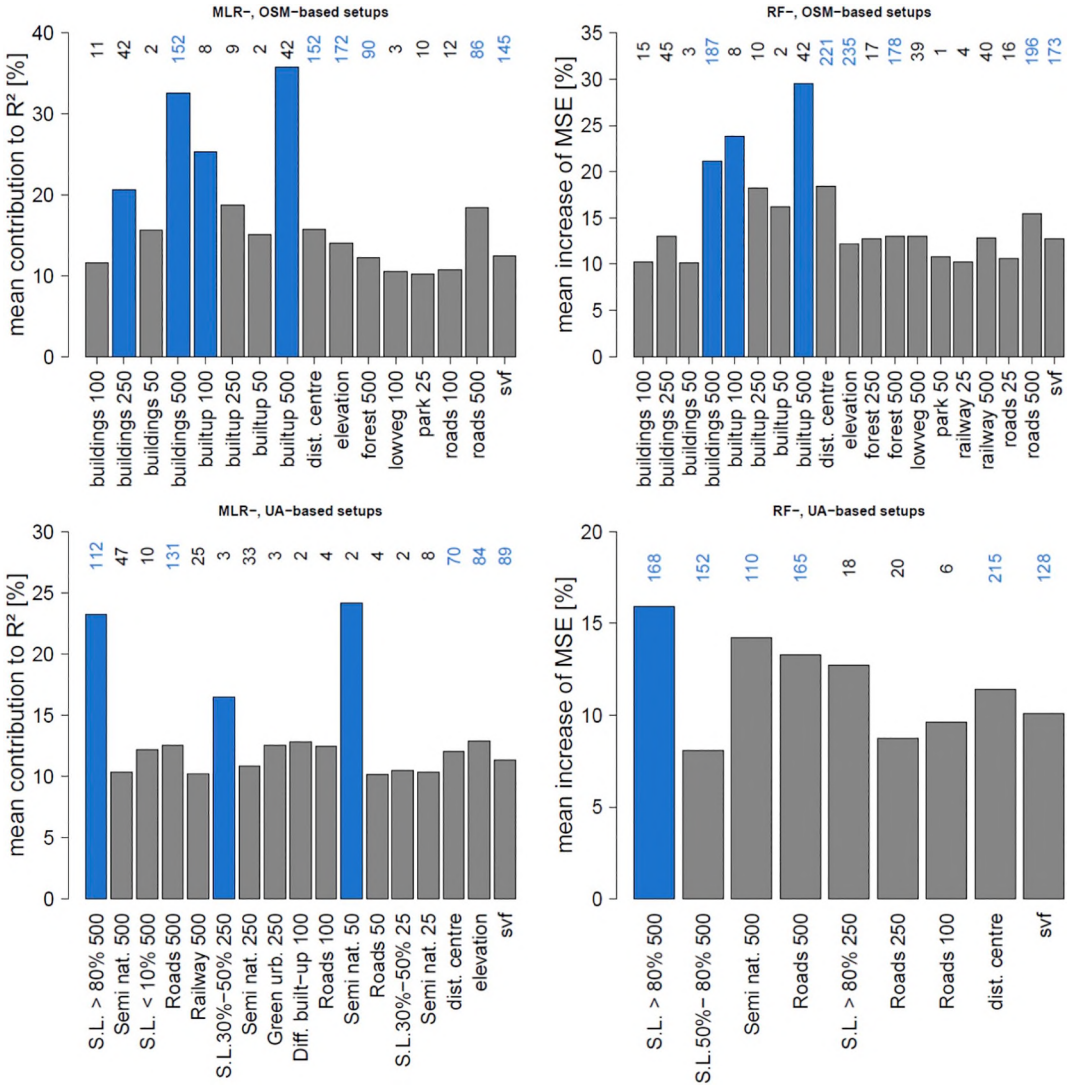


**Fig. 5.** Predictor importance averaged over all model variants and setups for variables with a considerable contribution to the model quality (contribution to $R^2$ of ≥10% for MLR and increase of MSE ≥ 10% for RF). For MLR-based approaches, the mean contribution to the overall explained variance of selected predictors is presented. For RF-based approaches, the predictor importance is expressed as the mean increase of the MSE if the respective predictor was excluded. Numbers above the bars indicate the number of model setups in which the respective predictor is included. Blue bars and numbers indicate the most important and the most frequently included explanatory variables, respectively. X-axis labels indicate predictor variable and – if available – size of included buffer zone [m]. dist. Centre = distance to the city centre, lowveg = low vegetation, S.L. = urban fabric, sealed layer x%, semi nat. = agricultural areas, semi-natural areas and wetlands, diff. Built-up = Industrial, commercial, public, military and private units, green urb. = green urban areas. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between 0.1 and 0.2 lower than $R^2$ from the fitted model. Most of the models for windy conditions at different times of the day as well as models for varying weather situations at noon show $R^2$ below 0.5 and $R^2$ from LOOCV below 0.3 (Fig. 3). Maxima of $R^2$ from LOOCV of up to 0.8 are reached for individual models (not shown), with reductions of < 0.1 compared to $R^2$ of the fitted models (Appendix D). Besides it should be mentioned, that MSSS and $R^2$ are identical or differ only slightly in most cases (not shown). When these two scores are identical, the models have no bias (Déqué, 2003).

### *3.1.2. Comparison of different model setups*

Model variants with different combinations of predictors each result in large variations in the performance of models for different seasonal, daily or meteorological situations (Fig. 2 and Fig. 4). Box-whisker plots of the performance measures for different model variants show that in most cases, highest MSSS exceed 0.8 for MLR-based models and 0.7 for RF-based models, respectively. The worst MLR-based models have MSSS below 0.2, for RF-based models, the MSSS becomes negative in several cases (Fig. 4 top). Considerable variations are also present in the MSE of the variants, being close to 0 °C for the best models but exceeding 1 °C for the

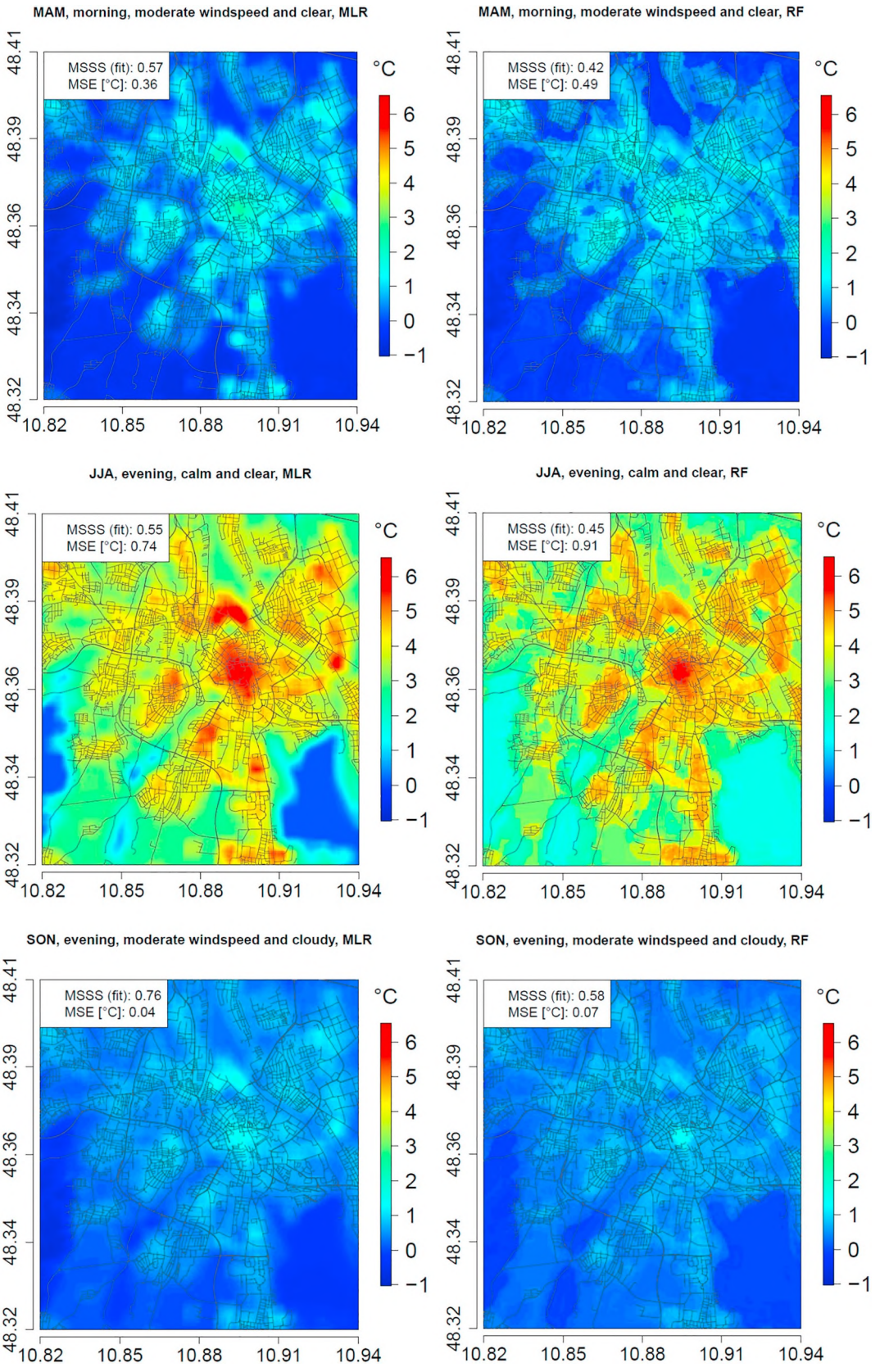
MAM, morning, moderate windspeed and clear, MLR
MSSS (fit): 0.57
MSE [°C]: 0.36
MAM, morning, moderate windspeed and clear, RF
MSSS (fit): 0.42
MSE [°C]: 0.49
JJA, evening, calm and clear, MLR
MSSS (fit): 0.55
MSE [°C]: 0.74
JJA, evening, calm and clear, RF
MSSS (fit): 0.45
MSE [°C]: 0.91
SON, evening, moderate windspeed and cloudy, MLR
MSSS (fit): 0.76
MSE [°C]: 0.04
SON, evening, moderate windspeed and cloudy, RF
MSSS (fit): 0.58
MSE [°C]: 0.07
°C
6
5
4
3
2
1
0
−1
48.41
48.39
48.36
48.34
48.32
10.82
10.85
10.88
10.91
10.94

**Fig. 6.** Exemplary results of estimated spatial distributions of air temperature deviations from the rural reference station including model performance measures for three selected situations as well as for both statistical model approaches: spring, morning, moderate wind speed and clear (top), summer, evening, calm and clear (middle) and autumn, evening, moderate wind speed and cloudy (bottom) as well as MLR (left), RF (right). Displayed are the results for the sub-daily OSM-based model setup without SVF ("OSM w/o SVF") for the study area as defined in Fig. 1.

worst (Fig. 4 bottom). Differences in the overall performance of individual setups are visible in the medians of the performance measures. For MLR, OSM- as well as UA-based model setups including seasonal variation of the times of day ("adj. time") reach highest medians of the MSSS of 0.6, the medians of the MSE values lie below 0.2 °C (Fig. 4 left). For RF, the OSM-based setups with all predictors ("OSM"), adjusted times of day ("adj. time") and summarized categories ("built-up") reach equally high performance with medians of the MSSS around 0.45 and medians of the MSE around 0.25 °C (Fig. 4 right). Models excluding the SVF ("w/o SVF"), especially the UA-based variants, are among those with lowest MSSS and highest MSE. The OSM-based setup using the land use/land cover information of source regions of advective airflow ("source") also results in relatively poor MSSS, but its MSE is comparably low. For both approaches, daily models have considerably higher MSSS with medians around 0.7 and 0.5 for MLR- and RF-based models, respectively, and OSM-based models performing better than UA-based ones. Negative values do not occur. The medians of the MSE for daily models are close to 0 °C, the maxima lie below 0.5 °C for both approaches and consequently are lower than for the sub-daily variants (Fig. 4).

As also visible in Fig. 2, the boxplots show, too, that the overall performance of RF is slightly worse than for MLR, with considerably lower median values of the MSSS for all setups and slightly higher MSE for many setups (Fig. 4).

For MLR, the differences between the explained variances from the fitted models and the leave-one-out cross validation are small for some setups. The medians of the differences for the UA-based models reach about 0.2 and accordingly are higher compared to the OSM-based setups, whose medians lie distinctly below 0.2. Small differences between $R^2$ and $R^2$ from LOOCV are reached especially for the "OSM built-up" setup, with a median of about 0.1 (cf. Appendix D).

### 3.2. Predictor importance

Fig. 5 presents the average predictor importance of selected explanatory variables for MLR and RF as well as both land use/land cover data sets over all model variants and setups. The most important predictors for both model approaches include variables describing artificial surfaces, thus the categories "buildings", "built-up" and "roads" for OSM-based approaches as well as "urban fabric" with different percentages of sealed layer for UA-based approaches. These predictors are important in the buffer zone sizes larger than 100 m. Furthermore, the distance from the city centre, the SVF as well as the surface elevation contribute substantially to the explained variance of the models. They have lower values of the respective predictor importance measures in many cases but they are included frequently (Fig. 5). Predictors describing natural surfaces are less important. For OSM-based approaches, "forest" is the most important of the natural types, classes describing low vegetation and water are less important. Here, land use/land cover characteristics within larger buffer zones are also more important than in the closer surrounding of the respective site. In the UA-based setups, the only important natural type is "agricultural areas, semi-natural areas and wetlands" for RF-based models (Fig. 5).

### 3.3. Spatial patterns of air temperature

Spatial patterns of the air temperature deviations from the rural reference station in the urban area of Augsburg were derived from the OSM-based model setup with fixed daytimes but without SVF, as this predictor was only available for selected sites. The patterns show distinct spatial differences of the air temperature deviations from the rural reference station. Exemplary situations are displayed in Fig. 6. In general, MLR- and RF-based models show similar patterns. In both approaches, highest air temperatures occur in densely built-up areas in the city centre and the central parts of the city districts. Large vegetated areas within the city are cooler, but the effect of small parks is not visible in the models. The lowest air temperatures occur on the unsealed areas surrounding the city, especially in the forest in the south-eastern part of the study area (10.91°E-10.94°E, 48.33°N-48.35°N). Sometimes these parts are even cooler than the rural reference station (Fig. 6). The temperature differences within the urban area of Augsburg are highest in calm and clear situations, especially in summer in the evening, when they exceed 6.0 °C in the MLR-based model and amount up to 5.5 °C in the RF-based setup (Fig. 6, middle). Intermediate temperature differences of up to 2.5 °C for MLR and 2.0 °C for RF occur e.g. in clear situations with moderate wind speeds in the morning in spring (Fig. 6, top). In addition to wind speed, cloudiness also contributes to a reduction of the air temperature differences within the study region. In cloudy situations, temperature elevation compared to the rural reference station amounts to e.g. 1.5 °C for autumn, evening and moderate wind speed for both, MLR and RF (Fig. 6, bottom), or 2.0 °C for winter, morning and calm conditions (not shown). Comparatively low temperature differences within the study area are also modelled at noon. In these situations, temperatures higher than those of the reference station occur only in the areas with highest density of buildings (not shown). Differences between the situations appear also in the spatial extent of the warmest parts of the city, with larger areas affected in situations with large temperature deviations within the study area (e.g. Fig. 6 middle and bottom for RF-based results). The similarity between the resulting air temperature patterns of MLR and RF also becomes apparent in the close correlation between them, with values of 0.9 and higher for the exemplary situations (cf. Appendix E).

The absolute differences between the patterns are small for most parts of the study region for the two situations with moderate wind speed, but reach magnitudes of up to 1.5 °C at some locations for the calm and clear situation in the evening in summer (cf. Appendix E). Major differences between the two methods concern an area north of the city centre (around 10.89°E, 48.39°N), which

appears to be as warm as the city centre in the MLR-based models but is notably cooler in the RF-based approaches. As there are no measurements available for this specific area, a direct comparison of the statistical models with observations is not possible. Besides, the forested areas in the southeast (10.91°E-10.94°E, 48.33°N-48.35°N) and southwest (10.82°E-10.83°E, 48.34°N-48.36°N) of the model region are cooler in the MLR-based model (Fig. 6, middle). Thus, differences between the model approaches concern as well the estimated range of air temperature differences within the study area, which appear to be slightly smaller in the RF-based models compared to the MLR-based models. This is also visible in the maps showing the differences between MLR- and RF-based patterns, where positive and negative deviations are present in every example (cf. Appendix E). Besides, the RF-based models show sharper air temperature gradients between different surface structures, which are homogeneous in themselves, as for example visible at the transition between the aforementioned forests and the surrounding built-up areas, especially in the examples for spring and autumn (Fig. 6, top and bottom).

The MSE values for the exemplary situations are smaller for MLR than for RF and the MSSS values are smaller for RF than MLR, indicating, as stated before, that the MLR-based models fit slightly better (Fig. 6).

## 4. Discussion

### *4.1. Model performance*

The MSSS and MSE of the MLR-based models are altogether better than that of the RF-based models (Fig. 2 and Fig. 4). Opposed to this, the studies by Ho et al. (2014) and Makido et al. (2016) result in higher performance of RF compared to other model approaches, but this could also result from the utilization of different kinds of input data in these studies. After Hjort et al. (2011), a large amount of observations is necessary for an optimal prediction with RF. However, in this study the formation of classes regarding season, time of day and meteorological conditions leads to a reduced sample size. Furthermore, for individual classes data was only available from a subset of the HOBO logger stations. Appendix F shows exemplarily for the sub-daily model setups with fixed times of day and source regions that for some situations no observations are available for a considerable percentage of the logger stations.

The explained variances of MLR-based models reach up to 85%, which is in line with the results of other studies ranging from 81% (Unger et al., 2001) to 92% (Alcoforado and Andrade, 2006). The medians of the differences between $R^2$ and $R^2$ from LOOCV for MLR-based models lie below 0.2 for most of the model setups. Consequently, they are comparable with the results from other studies, e.g. Buttstädt and Schneider (2014), where the difference amounts to a maximum of 0.17. The model performance exhibits considerable variations in relation to time of day and weather conditions with best models obtained for situations in the morning and evening with calm and moderately windy conditions. Models for situations around noon as well as for windy conditions perform quite poor in comparison. These findings are compatible with the results of e.g. Eliasson and Svensson (2003). A possible explanation could be the overall small spatial air temperature differences in the stated situations (e.g. Oke, 1993; Schatz and Kucharik, 2014), whereas under conditions of higher stability of the planetary boundary layer in situations with low or negative net radiation as well as in anticyclonic weather conditions air temperatures are influenced stronger by the underlying surface characteristics (Morris et al., 2001; Eliasson and Svensson, 2003). However, improvements of the model performance for situations around noon, when heat stress is most pronounced, are of particular interest for the various applications of such models, e.g. the assessment of human health in relation to outdoor thermal comfort (Matzarakis et al., 2008). A clear seasonal differentiation of model performance does not appear.

The performance of the daily model variants exceeds that of the sub-daily setups. A possible explanation is the categorization of the synoptic situation based on the current conditions only while not accounting for the meteorological conditions in advance of the modelled situation. Consequently, the situations included in one class can be heterogeneous. For example, a situation classified as “calm and clear” may follow on “windy and cloudy” conditions and thus be not an ideal representative for its category. For sub-daily models, this can have a negative impact on the model building process, whereas for the daily setups this effect is presumably smaller. Taking into account such time lagged effects is proposed by Unger et al. (2001). This aspect could be addressed in further studies, e.g. by taking into account the persistence of the meteorological boundary conditions over a certain time interval. As the daily cycle is an important feature when considering air temperatures in urban environments as well as thermal stress and associated health issues, for example because night-time cooling is essential for the recovery after hot days (Nicholls et al., 2008; Steeneveld et al., 2014), increasing the performance of sub-daily models is of great relevance for their various applications.

For many situations, better model performance is reached in OSM based models compared to UA based variants. The difference is visible especially regarding the MSSS and MSE from daily model setups, the MSSS from RF-based sub-daily models as well as the differences between $R^2$ and $R^2$ from LOOCV from MLR (Fig. 4 and Appendix D). A possible reason could be the heterogeneity of the land surface characteristics within several of the classes available in UA. While some of the UA land use/land cover classes are composed of one kind of surface, e.g. “water”, and thus can be assumed to have a well defined impact on energy and water balances of the respective surfaces, others, e.g. “agricultural areas, semi-natural areas and wetlands” contain elements whose properties with regard to albedo, heat and water capacity can be expected to be variable. This complicates the detection of correlations between air temperature and land surface characteristics and may contribute to a reduced robustness of the models.

Models without SVF, both OSM and UA based variants, are among those with worst performance, confirming that SVF is an important predictor for air temperature. This is in line with the results from other studies, e.g. Svensson (2004) and Konarska et al. (2015). For further improvements of the model performance seasonal variations of the SVF due to the phenological cycle of the vegetation cover should be taken into account (Svensson, 2004).

The OSM based variant using summaries of selected land cover types to form the new categories “built-up” and “low vegetation” is among the worse models regarding the medians of MSSS and MSE, but the difference between $R^2$ and $R^2$ from LOOCV is

comparably small. This indicates that including a smaller number of predictor variables, some of which contain different land use/land cover types with similar effects on air temperature, produces models that are more robust when transferred to test data, presumably due to a reduction of the multicollinearity.

The OSM based variant using the land cover characteristics of source regions of advective airflow seems to be a promising setup for weather conditions with moderate to higher wind speeds. Nevertheless, it does not perform better than the other model variants. The MSE is in fact comparably small; however, the MSSS is quite poor compared to the other model setups. A possible reason is the definition of the source areas as quadratic areas southwest of each logger station. A more sophisticated determination of source areas would be more appropriate, because urban wind patterns are complex and greatly influenced by the orientation of street canyons and the building geometry (Nakamura and Oke, 1988; Eliasson et al., 2006). Including situation specific wind speed and wind direction in combination with the site-specific urban geometry would allow a more accurate determination of position and size of the source areas for individual situations. Besides, as high wind speeds favour relatively homogenous spatial air temperature distributions, the investigation of the setup with source regions under various wind conditions, including low wind speeds, could be subject of further studies.

Other possibilities for the improvement of the model performances could include the utilization of smaller seasonal and/or meteorological subsets, as proposed by Eliasson and Svensson (2003). Their explanation is the relatively high variation of observations within large groups, which could also apply to this study. Choosing a larger number of classes, e.g. by utilizing monthly instead of seasonal subsets, could lead to an improvement of model performance. On the other hand, models taking into account seasonal variations of the hours of sunrise, apex and sunset use smaller temporal subsets than the other variants. However, for MLR-based models their overall performance is not clearly improved compared to the other variants and for RF-based models it is worse. This indicates that the negative impact of smaller sample sizes for the individual situations predominates the assumed positive effects of reduced variation within the groups. A stratification into smaller temporal subsets leads to an increase of the percentage of logger stations without any observations for individual classes, thus resulting in models based on a reduced number of stations. This is shown in Appendix F where for the model setup with source regions, which includes only situations with specific wind directions, the sample size and numbers of HOBO loggers with at least one observation is considerably lower than for the model setup with fixed times of the day. Consequently, a stratification into a larger number of smaller subsets would reduce the representativity of the affected models for the study area. Besides, as stated before, the meteorological conditions in advance of the modelled periods may also vary, inducing another source of inhomogeneity within the groups.

### 4.2. Predictor importance

Important predictors in both approaches are those describing the alteration of the land surface by urban elements, e.g. percentage of buildings, built-up area, types describing high level of sealed surface, the SVF, the distance to the city centre, but in some model variants also the elevation. These or similar explanatory variables have been found to be important predictors of air temperature in other cities, too, e.g. in Lisboa (Alcoforado and Andrade, 2006), Szeged (Unger et al., 2001) or Göteborg (Eliasson and Svensson, 2003). In cities with highly variable elevation, this parameter can be an important predictor due to phenomena related to the topographic influence on air temperature, as is the case e.g. in Aachen (Buttstädt and Schneider, 2014) with a variation of the elevation of 285 m within the study area. In this study, these variations are much smaller. They amount to only 89 m for the region of Augsburg and thus, the topographic influence can be expected to be less important. On the other hand, after Hjort et al. (2011) even small spatial differences in elevation can have an effect on air temperature. Variables describing land surface features with presumable cooling effect, e.g. “forest”, “urban green spaces” or “water”, are – except “forest” – less important for the models in this study. This is in contrast to other studies, where e.g. water is an important predictor of air temperature in coastal cities (e.g. Hart and Sailor, 2009; Alcoforado and Andrade, 2006) or cities in the vicinity of large water bodies (e.g. Hjort et al., 2011). It could be a result of the absence of large water bodies in the region of Augsburg. Here, it should be mentioned, that the cooling effect of water is discussed controversial. While on the one hand urban water is discussed in the context of UHI mitigation strategies (e.g. Oláh, 2012; Coutts et al., 2012; Hathway and Sharples, 2012), Steeneveld et al. (2014) state, that standing water can increase night-time UHI due to the thermal properties of water. Spatial predictors available as area percentages were most important in larger buffer zones, which is in line with the results from Buttstädt and Schneider (2014).

In other studies, a large number of other explanatory variables has been found to model air temperature well, e.g. the normalized difference vegetation index (NDVI, e.g. Szymanowski and Kryza, 2012), the normalized difference water index (NDWI, e.g. Ho et al., 2014), the albedo (e.g. Makido et al., 2016), land surface temperatures (e.g. Ho et al., 2014) or the building height (e.g. Bottyán and Unger, 2003). Further improvement of the models could be obtained by incorporating such additional parameters.

### 4.3. Spatial patterns of air temperature

The spatial patterns of air temperature for the different situations derived with each of the model approaches show distinct variations dependent on the surface characteristics. Highest air temperatures occur on the areas with high percentage of artificial surfaces, as would be expected due to the higher thermal conductivity and heat capacity of building materials (Oke, 1988) and the reduced SVF within street canyons (Oke, 1981; Svensson, 2004). In contrast, areas with high percentage of vegetation and water are cooler because of the higher evaporation rate (e.g. Jáuregui, 1975; Oláh, 2012). These findings are in line with the results from other cities (e.g. van Hove et al., 2015; Shojaei et al., 2017; Tong et al., 2018). The patterns form quasi-concentric circles around the city centre and the centres of the city districts, which in this case is due to the utilization of the predictor variable “distance from the city centre”. Concentric patterns were also obtained in other studies, e.g. by Bottyán et al. (2005). Differences in the spatial patterns of different meteorological situations concern mainly the magnitude of the air temperature deviations from the reference station and the

spatial extent of the warmest areas (Fig. 6). Intra-urban air temperature differences are most pronounced under clear and calm conditions, as was also observed by e.g. Yokobori and Ohta (2009) and van Hove et al. (2015). Wind speed notably reduces the intra-urban temperature differences due to the more effective mixing (e.g. Wienert et al., 2013; He, 2018; Shi et al., 2018). However, differences also result from the specific subset of predictor variables included in each model.

### *4.4. Meteorological and land surface data*

Limitations of the models result from uncertainties in the input data. Although several quality tests were conducted on the temperature data from the HOBO logger network, important aspects like the internal/physical consistency (e.g. Estévez et al., 2011) or the homogeneity (e.g. Aguilar et al., 2003) were not verified. On the other hand, comparisons between HOBO logger stations and automatic weather stations situated at the University and the University of Applied Sciences of Augsburg suggest that the HOBO logger network provides a solid data basis for this application. An exemplary figure is presented in the appendix. Monthly boxplots of the hourly air temperature show that the HOBO logger slightly underestimates maximum temperature values. However, the overall correspondence between the two devices is good (Appendix C). Besides, in order to derive valid models, the temperature measurements at the HOBO logger sites have to be representative for the study area. As can be seen in Appendix B and Beck et al. (2018), the logger stations are distributed over the various local climate zones (LCZ), which are present in the region of Augsburg and they cover transition areas between different LCZs as well. Furthermore, they are situated at different heights above sea level. Thus, the HOBO loggers represent differently characterized sites in terms of land cover and topography.

The land use/land cover data from Open Street Map also contain uncertainties because it is a crowd-sourcing project, which means that everybody can contribute to the data base (Arsanjani et al., 2013). Consequently, the data base constitutes of input from persons with different amount of expertise and experience on the subject of mapping which leads to large differences in the quality of individual contributions. However, according to Arsanjani et al. (2013), a large part of the input comes from persons that are able to generate map data of high accuracy. Besides, Zielstra and Zipf (2010) show that the positional accuracy as well as the completeness of the OSM data have been increasing steadily. Even so, for some areas in the city of Augsburg and its surroundings, the land use/land cover information from OSM is missing. Although the model area was selected to contain a minimum of missing land use/land cover data, this is a problem for the estimation of the continuous spatial patterns of air temperature because the modelled thermal conditions on the affected areas cannot be linked to surface parameters.

The Urban Atlas data covers the whole study area. However, as stated in section 4.1, a drawback of the Urban Atlas data for this application might be the combination of different surface cover features in some of the available land use/land cover categories. While some of them appear to be homogenous (e.g. “Forest”), others contain land use/land cover classes quite dissimilar in terms of their radiation and water balance. Thus, they can hardly be interpreted regarding their influence on air temperature (e.g. the category “Agricultural areas, semi-natural areas and wetlands” which contains different kinds of vegetation as well as bare soil and swamps) (EEA, 2011). This makes the interpretation of the UA based results difficult for the affected areas. Another drawback of the UA data is that they are updated only once in several years. For individual areas this can be a problem.

A possibility to combine the advantages of OSM and UA would be a model setup utilizing both data sets together. In this way, a larger range of predictors would be available and the strongest explanatory variables could be combined. On the other hand, most of the land surface information should be similar in both data sets. Still, the problem of missing values in the OSM data could be avoided by filling the respective areas with information from UA. If a combination of OSM and UA based predictors leads to a model improvement, could be subject of further studies.

## 5. Conclusion

In this study, sub-daily as well as daily spatial air temperature patterns within a heterogeneous urban environment were modelled for a variety of different seasonal, daily as well as meteorological situations using the statistical approaches of multiple linear regression and random forests. The aim was to investigate the skill of various model setups including different land use/land cover characteristics as predictors. These predictors were derived from different land surface data sets. Large differences in model performance between the model variants were obtained with the best MLR-based results reaching high MSSS of up to 0.85 as well as explained variance of up to 85% in the fitting procedure and 80% in leave-one-out cross-validation. Models utilizing the Open Street Map land use/land cover data perform slightly better than those based on Urban Atlas data. Thus, current land use/land cover information with categories, which are homogeneous in themselves and interpretable regarding their influence on the water and energy budget of the respective surface, is favourable for such studies. The estimated spatial patterns of the thermal conditions are realistic for different situations and are similar for both statistical modelling approaches. Consequently, these statistical models are an appropriate tool to derive the spatial distribution of meteorological variables as a function of land surface parameters for varying synoptic boundary conditions. However, there is still potential for improvement of the model results. Options for this task include the inclusion of further explanatory variables, the calculation of the sky view factor for the whole study area and a more realistic definition of source regions of advection. The maps resulting from such models can be useful for different applications, e.g. for urban planning, for the development of concepts for the climate change adaption of cities or for human biometeorological studies.

## Declaration of interest

None.

## Acknowledgements

The authors would like to thank the anonymous reviewers for the valuable comments and suggestions on the manuscript, which helped to improve this paper.

## Appendix A. Land use/land cover characteristics of the study region from the European Urban Atlas

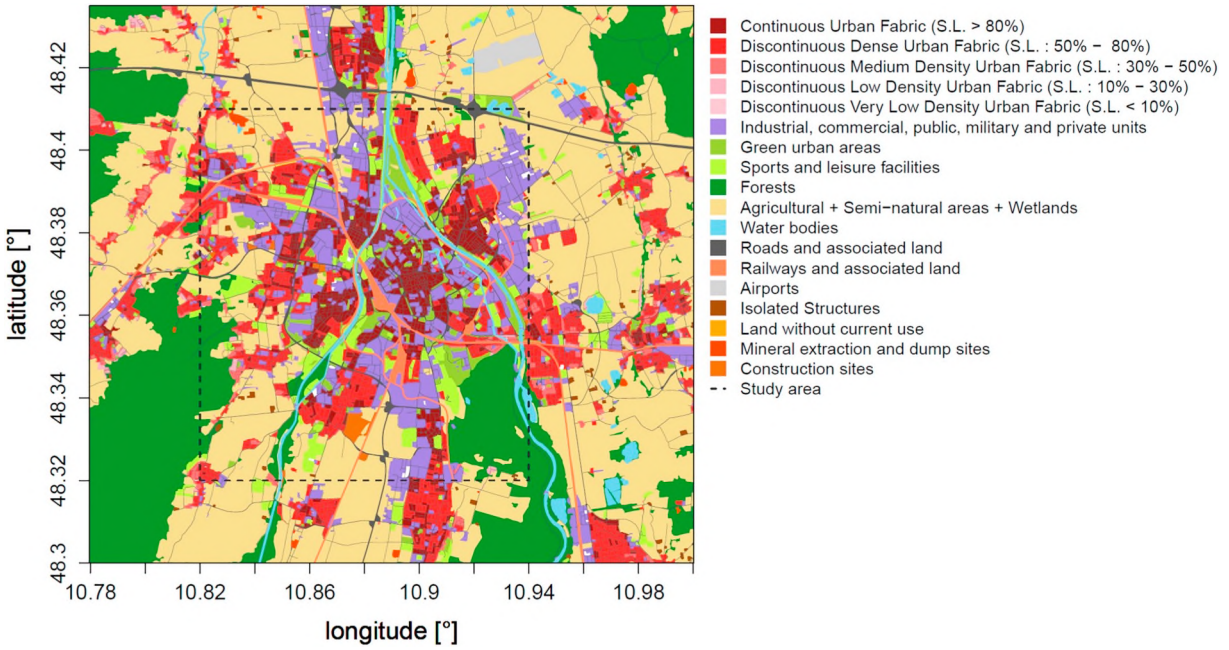


**Fig. 7.** Land use/land cover in the region of Augsburg from the Urban Atlas data set ("S.L." = sealed layer)

## Appendix B. Site characteristics of the HOBO logger stations

Table 3
Site characteristics of the HOBO logger stations: name and ID of the station, geographic coordinates and elevation of the measurement site [m above sea level], local climate zone (LCZ, after Stewart and Oke, 2012) at the site and area percentages of the LCZs within a 500 m buffer zone around the site. For a map of the local climate zones in Augsburg and its surroundings, including also the logger stations, cf. Beck et al. (2018). Abbreviations of the LCZs: 2 = Compact Mid-Rise, 5 = Open Mid-Rise, 6 = Open Low-Rise, 8 = Large Low-Rise, 101 = dense trees, 102 = scattered trees, 104 = low plants.

| Station name | ID | Lat [°] | Lon [°] | Height [m] | LCZ-Station | LCZ-Neighborhood (500 m) |
|---|---|---|---|---|---|---|
| Moritzplatz | 1 | 48.36707 | 10.89799 | 488.4 | 2 | 2 (64%); 5 (36%) |
| Marketplace | 2 | 48.36808 | 10.89561 | 490.7 | 2 | 2 (62%); 5 (38%) |
| Volunteer fire brigade, Pfersee | 3 | 48.36171 | 10.87526 | 478.4 | 5 | 5 (90%); 6 (2%); 8 (2%); 101 (1%); 102 (5%) |
| Police headquarters Augsburg | 4 | 48.35556 | 10.88699 | 494.5 | 5 | 5 (69%); 8 (18%); 102 (13%) |
| Measurement site of the state office for environment at Karlstrasse | 5 | 48.37036 | 10.89679 | 487.8 | 2 | 2 (42%); 5 (58%) |
| School at Rotes Tor | 6 | 48.35788 | 10.90291 | 484.3 | 5 | 2 (3%); 5 (97%) |
| Measurement site at the University of Applied Sciences Augsburg | 7 | 48.35794 | 10.90705 | 481.4 | 5 | 2 (3%); 5 (97%) |
| Train station forecourt | 8 | 48.36633 | 10.88583 | 486.7 | 8 | 2 (10%); 5 (59%); 6 (13%); 8 (18%) |
| Measurement site of the state office for environment at Koenigsplatz | 9 | 48.36459 | 10.89503 | 490.7 | 5 | 2 (50%); 5 (45%); 6 (1%); 8 (4%) |
| Fire brigade Augsburg | 10 | 48.3741 | 10.91424 | 474.9 | 5 | 2 (8%); 5 (79%); 6 (5%); 8 (8%) |
| Kindergarten St. Moritz | 11 | 48.36563 | 10.90106 | 491.5 | 2 | 2 (58%); 5 (42%) |
| Gaertnerstrasse | 12 | 48.3603 | 10.90738 | 480.1 | 5 | 2 (8%); 5 (91%); 8 (1%) |
| Gollwitzerstrasse Pfersee | 13 | 48.36051 | 10.87592 | 476.5 | 5 | 5 (84%); 6 (5%); 101 (2%); 102 (9%) |
| Measurement site of the state office for environment at Bourges-Platz | 14 | 48.37659 | 10.88848 | 473.8 | 5 | 2 (20%); 5 (69%); 6 (3%); 8 (7%); 104 (1%) |

Table 3 (*continued*)

| Station name | ID | Lat [°] | Lon [°] | Height [m] | LCZ-Station | LCZ-Neighborhood (500 m) |
|---|---|---|---|---|---|---|
| Weather station St. Stephan | 15 | 48.37428 | 10.90013 | 486.8 | 5 | 2 (21%); 5 (77%); 8 (1%); 104 (1%) |
| Wittelsbacher Park | 16 | 48.35683 | 10.88118 | 494.3 | 102 | 5 (67%); 6 (9%); 8 (6%); 101 (2%); 102 (16%) |
| Cemetery, Hermanstrasse | 17 | 48.36401 | 10.89271 | 494 | 5 | 2 (39%); 5 (44%); 6 (1%); 8 (16%) |
| Botanic Garden | 18 | 48.3503 | 10.91506 | 482 | 102 | 5 (69%); 6 (1%); 102 (29%); 104 (1%) |
| Kulperhuette | 19 | 48.35141 | 10.86784 | 479.3 | 101 | 5 (22%); 6 (56%); 101 (14%); 102 (2%); 104 (6%) |
| Siemens Technopark | 20 | 48.3417 | 10.90532 | 490.7 | 8 | 5 (28%); 8 (53%); 102 (12%); 104 (7%) |
| Roof station at the University of Augsburg | 21 | 48.33466 | 10.89713 | 490.7 | 5 | 5 (63%); 6 (1%); 8 (19%); 104 (17%) |
| Municipal construction yard, Neusaess | 22 | 48.39322 | 10.82416 | 479.3 | 8 | 5 (38%); 6 (26%); 8 (24%); 101 (1%); 102 (11%) |
| Church at Goegginger bridge | 23 | 48.33781 | 10.85684 | 487.7 | 6 | 6 (48%); 101 (26%); 102 (5%); 104 (21%) |
| Measurement site at the State office for environment | 24 | 48.32602 | 10.90308 | 494.2 | 5 | 5 (71%); 6 (9%); 8 (5%); 104 (3%); 106 (12%) |
| Gasworks | 25 | 48.38771 | 10.86526 | 473.2 | 8 | 6 (60%); 8 (40%) |
| Krautgartenweg Leitershofen | 26 | 48.34864 | 10.83809 | 487.5 | 6 | 6 (60%); 102 (25%); 104 (15%) |
| Private garden site, Neusaess | 27 | 48.38133 | 10.82 | 507.8 | 102 | 5 (1%); 6 (51%); 101 (33%); 102 (8%); 104 (7%) |
| Vegetable field, Baerenkeller | 28 | 48.39232 | 10.86423 | 472.2 | 6 | 6 (87%); 8 (7%); 104 (6%) |
| Water supply works, Lochbach | 29 | 48.3332 | 10.91299 | 487.6 | 101 | 5 (25%); 8 (1%); 101 (29%); 102 (44%); 104 (1%) |
| Hat factory Augsburg | 30 | 48.35203 | 10.90268 | 486.5 | 5 | 5 (96%); 6 (1%); 102 (2%); 104 (1%) |
| Ulmer Strasse, company building | 31 | 48.3792 | 10.84645 | 487.3 | 6 | 6 (77%); 8 (11%); 104 (12%) |
| Friesenstrasse | 32 | 48.38225 | 10.91751 | 472.8 | 5 | 2 (16%); 5 (42%); 6 (42%) |
| Neuburger Strasse | 33 | 48.38993 | 10.91457 | 469.7 | 6 | 5 (6%); 6 (59%); 8 (5%); 104 (30%) |
| Muehlstrasse/equestrian farm | 34 | 48.33034 | 10.85532 | 489.8 | 102 | 6 (8%); 101 (22%); 102 (37%); 104 (33%) |
| Eiskanal/Friedberger Strasse-Spickelstrasse | 35 | 48.35496 | 10.92768 | 482.3 | 5 | 5 (88%); 6 (3%); 101 (2%); 102 (7%) |
| Diedorf (Church of Herz Maria) | 36 | 48.35469 | 10.78167 | 484.9 | 6 | 5 (8%); 6 (67%); 8 (5%); 102 (3%); 104 (17%) |
| Municipal construction yard, Gessertshausen | 37 | 48.32821 | 10.72022 | 500.3 | 104 | 6 (3%); 101 (1%); 102 (25%); 104 (71%) |
| Volunteer fire brigade, Strassberg | 38 | 48.27558 | 10.78446 | 558.5 | 101 | 6 (15%); 101 (75%); 102 (4%); 104 (6%) |
| Bobingen, Church of St. Felizitas | 39 | 48.2609 | 10.83076 | 525.8 | 6 | 6 (53%); 104 (47%) |
| Grossaitingen, Church of St. Georg | 40 | 48.22859 | 10.78223 | 537.9 | 6 | 6 (84%); 102 (2%); 104 (14%) |
| Kleinaitingen, cemetery | 41 | 48.2206 | 10.8397 | 541.9 | 104 | 5 (4%); 6 (25%); 104 (71%) |
| Kissing, Church of St. Bernhard | 42 | 48.30366 | 10.96725 | 499.1 | 6 | 6 (76%); 8 (24%) |
| Friedberg, fire brigade | 43 | 48.3582 | 10.98659 | 512.3 | 6 | 6 (92%); 104 (8%) |
| Adelsried, parish office St. Johann Baptist | 44 | 48.42505 | 10.71755 | 487.6 | 6 | 5 (6%); 6 (35%); 101 (1%); 102 (39%); 104 (19%) |
| Heretsried, Church of St. Martin | 45 | 48.45951 | 10.7375 | 487.6 | 6 | 5 (4%); 6 (11%); 102 (63%); 104 (22%) |
| Schwabmuenchen | 46 | 48.1768 | 10.7578 | 533.4 | 6 | 6 (100%) |
| Mering, private premises | 47 | 48.2638 | 10.985 | 510.1 | 5 | 5 (43%); 6 (51%); 104 (6%) |
| Altenmuenster, day nursery | 48 | 48.46744 | 10.59051 | 487.6 | 6 | no data |
| Rehling, company building | 49 | 48.47405 | 10.89649 | 465.9 | 104 | 104 (100%) |
| Poettmes, company building | 50 | 48.58476 | 11.08857 | 487.6 | 6 | no data |
| Aichach, old fire station | 51 | 48.45771 | 11.13174 | 487.6 | 6 | no data |
| Mering, private premises | 52 | 48.27729 | 10.97062 | 506.8 | 6 | 6 (34%); 8 (4%); 104 (62%) |
| Poettmes, hall | 53 | 48.58207 | 11.0861 | 487.6 | 6 | no data |
| Horgau, church | 54 | 48.3939 | 10.68422 | 525.7 | 6 | no data |
| Zusmarshausen, welfare centre | 55 | 48.3991 | 10.59974 | 525.7 | 6 | no data |
| Gasworks | 56 | 48.38741 | 10.86638 | 473.5 | 8 | 2 (3%); 6 (53%); 8 (44%) |
| Car service station | 57 | 48.38547 | 10.85329 | 485.2 | 8 | 6 (57%); 8 (32%); 104 (11%) |
| Kobelgraben | 58 | 48.37917 | 10.80824 | 479.9 | 6 | 6 (27%); 101 (4%); 102 (11%); 104 (58%) |
| Burger King | 59 | 48.34708 | 10.87108 | 485 | 5 | 5 (34%); 6 (40%); 101 (23%); 102 (3%) |
| Tattenbachstrasse | 60 | 48.3078 | 10.90966 | 501.5 | 5 | 5 (82%); 6 (2%); 102 (15%); 104 (1%) |
| Siemens parking lot | 61 | 48.34079 | 10.90543 | 490.7 | 8 | 5 (32%); 8 (49%); 102 (12%); 104 (7%) |
| Aulzhausen | 63 | 48.44711 | 10.96793 | 468.8 | 104 | 5 (6%); 6 (2%); 104 (92%) |
| Derching | 64 | 48.40845 | 10.97907 | 502 | 104 | 5 (10%); 101 (2%); 102 (9%); 104 (79%) |
| Bobingen | 65 | 48.27243 | 10.82205 | 508.5 | 104 | 6 (18%); 101 (3%); 102 (16%); 104 (63%) |
| Wellenburg | 66 | 48.33599 | 10.82553 | 505.4 | 101 | 101 (64%); 102 (25%); 104 (11%) |
| Haunstetten, horse enclosure | 67 | 48.27298 | 10.90399 | 511.9 | 102 | 6 (15%); 101 (15%); 102 (20%); 104 (50%) |
| Ellensindstrasse | 68 | 48.31888 | 10.91557 | 495.8 | 102 | 5 (7%); 6 (15%); 8 (21%); 102 (53%); 104 (4%) |
| Stadtbergen, golf club | 69 | 48.36562 | 10.83696 | 490.4 | 102 | 6 (26%); 102 (53%); 104 (21%) |
| Diedorf, Kapellenweg | 70 | 48.34373 | 10.76339 | 487.1 | 104 | 101 (1%); 102 (30%); 104 (69%) |
| Oberottmarshausen | 71 | 48.23847 | 10.85647 | 533.6 | 6 | 6 (38%); 104 (62%) |
| Herbertshofen | 73 | 48.5298 | 10.85112 | 480.2 | 6 | 6 (44%); 104 (56%) |
| Measurement station at the University of Augsburg | 74 | 48.33676 | 10.89541 | 493.7 | 5 | 5 (42%); 6 (1%); 8 (44%); 104 (13%) |

(*continued on next page*)

Table 3 (*continued*)

| Station name | ID | Lat [°] | Lon [°] | Height [m] | LCZ-Station | LCZ-Neighborhood (500 m) |
|---|---|---|---|---|---|---|
| Meadow at the University of Augsburg | 75 | 48.33071 | 10.89843 | 493.9 | 5 | 5 (68%); 6 (2%); 104 (29%); 106 (1%) |
| Pond at the University of Augsburg | 76 | 48.33603 | 10.89607 | 494.3 | 5 | 5 (47%); 8 (36%); 104 (17%) |
| Parking lot at the exhibition centre | 77 | 48.33859 | 10.89185 | 498.9 | 8 | 5 (35%); 6 (1%); 8 (46%); 104 (18%) |
| Europaplatz | 78 | 48.33382 | 10.90139 | 493.9 | 5 | 5 (82%); 6 (2%); 8 (11%); 104 (5%) |
| Moritzplatz 2 | 79 | 48.36709 | 10.89811 | 485.6 | 2 | 2 (64%); 5 (36%) |
| Institute of Geography, University of Augsburg | 80 | 48.33599 | 10.89747 | 491.9 | 5 | 5 (53%); 8 (39%); 104 (8%) |
| Tauroggener Strasse | 81 | 48.38642 | 10.92158 | 472 | 6 | 5 (19%); 6 (77%); 8 (4%) |

## Appendix C. Comparison between HOBO logger and automatic weather station

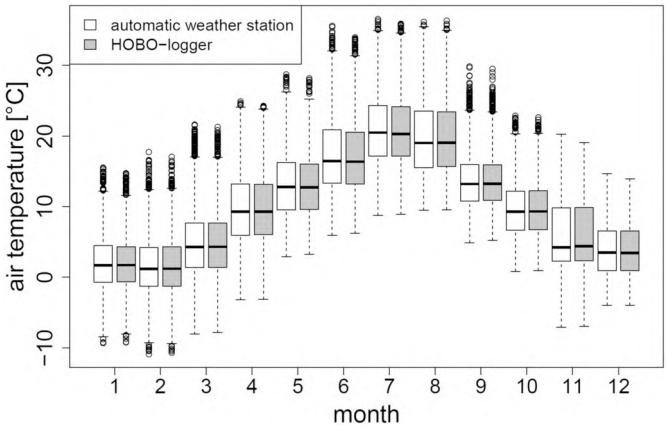


**Fig. 8.** Comparison of hourly air temperature measurements of an automatic weather station and a HOBO logger at the University of Applied Sciences in Augsburg on a monthly basis (comparison based on the period from December 2012 to June 2016). The boxes show the median as well as the 1st and 3rd quartile (lower/upper boundary of the box, respectively), whiskers indicate the 1.5-fold interquartile range from the upper/lower boundary of the box.

## Appendix D. Differences between $R^2$ from the model fitting and $R^2$ from LOOCV for MLR-based model setups

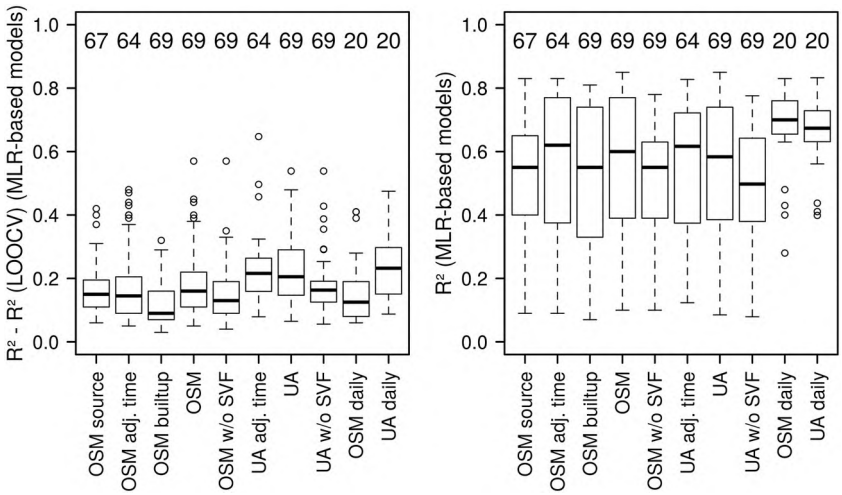


**Fig. 9.** Boxplots of the differences between $R^2$ from the model fitting and $R^2$ from LOOCV (left) and $R^2$ from the model fitting (right) for MLR-based model setups. The boxes show the median as well as the 1st and 3rd quartile (lower/upper boundary of the box, respectively), whiskers indicate the 1.5-fold interquartile range from the upper/lower boundary of the box. Each box contains the model variants for different seasons (MAM, JJA, SON, DJF) as well as meteorological situations (Met 1 – Met 6) and in the case of sub-daily models different times of day (morning, noon, evening), too. Numbers above the boxes indicate the sample size (number of models) contained in each box.

## Appendix E. Differences between MLR- and RF-based modelled air temperature deviations

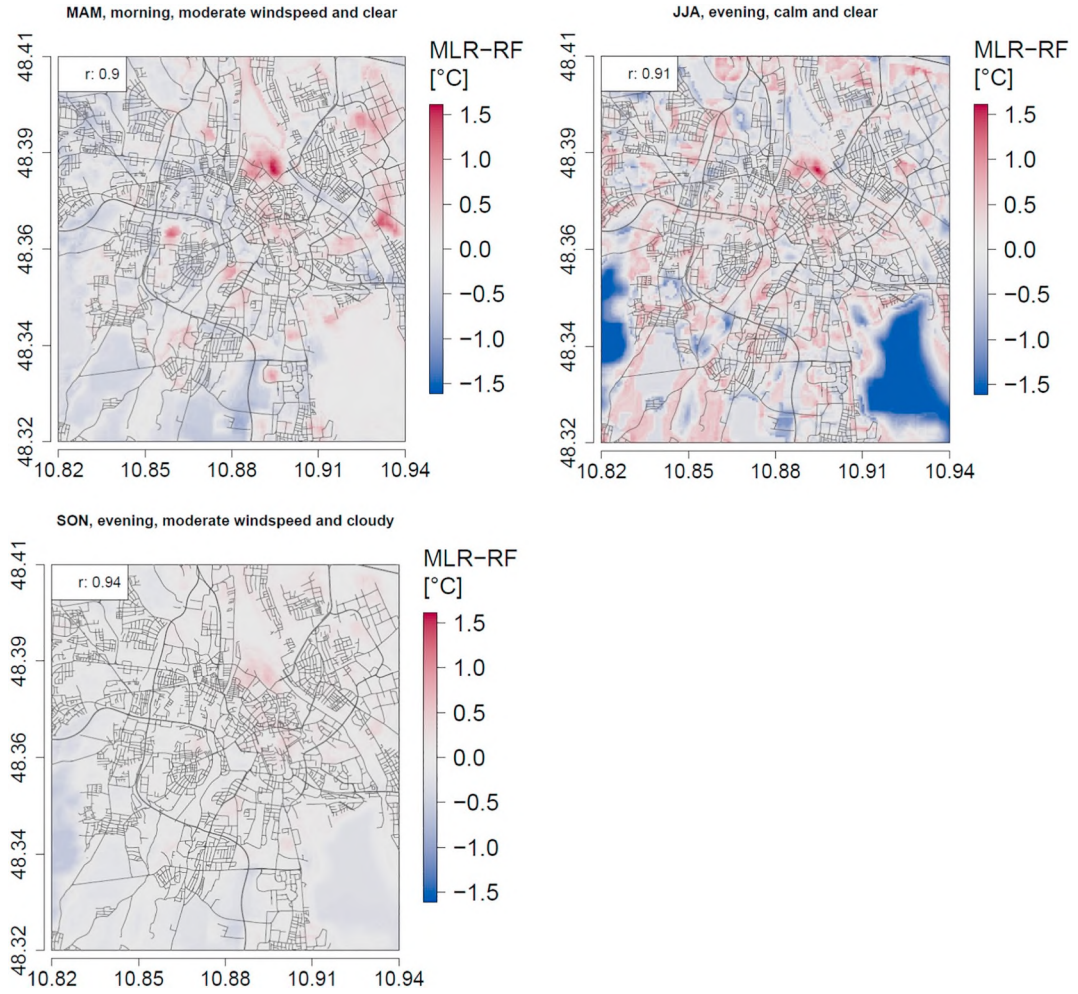


**Fig. 10.** Differences between modelled MLR- and RF-based temperature deviations from the rural reference station for exemplary situations as well as the spatial correlation between the patterns (Pearson correlation coefficient r). Displayed are the results for the sub-daily OSM based model setup without SVF ("OSM w/o SVF") for the study area as defined in Fig. 1.

## Appendix F. Data availability in the meteorological data sets (DWD station and HOBO logger network)

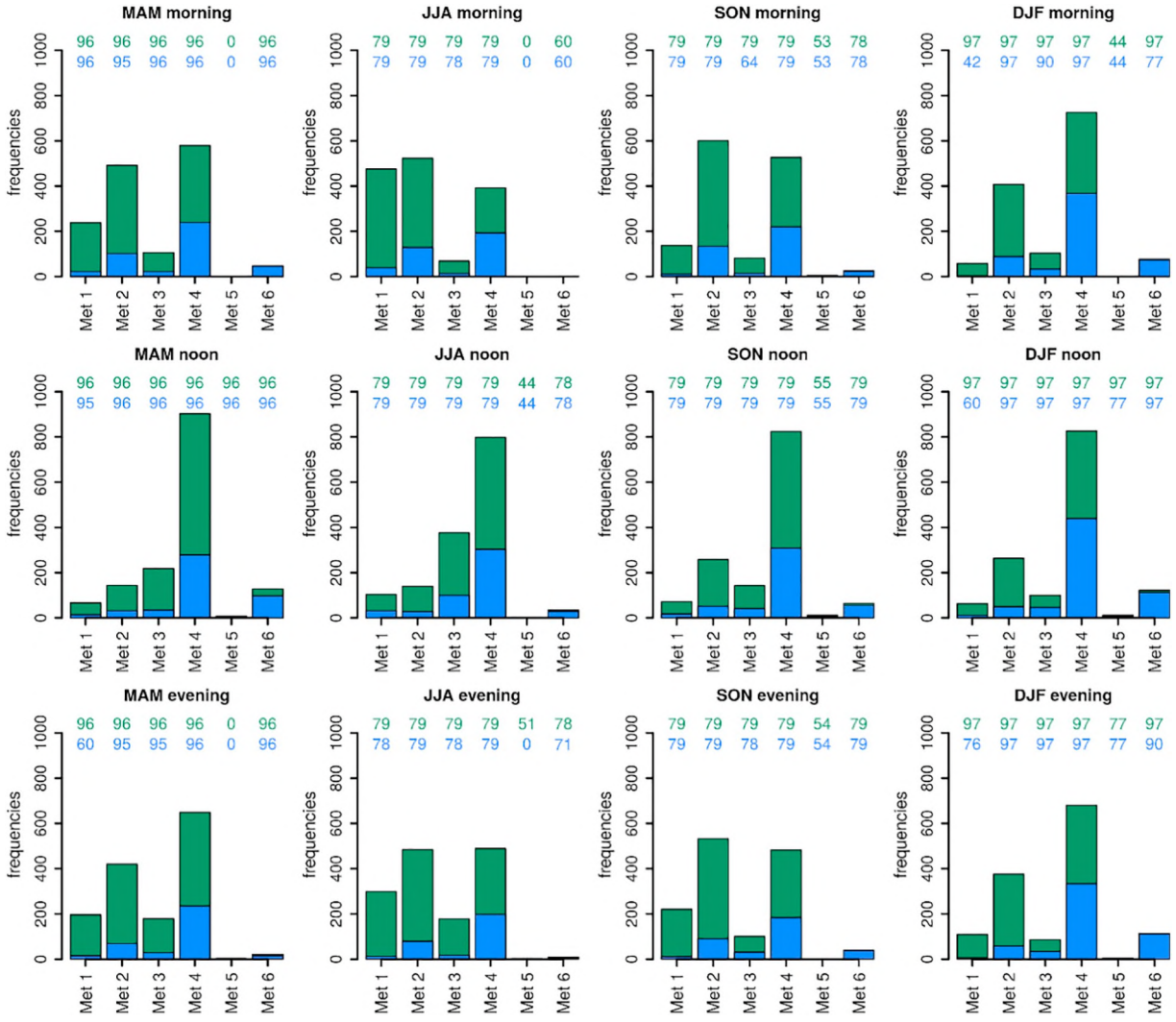


**Fig. 11.** Overview of data availability of the HOBO logger data for different seasons, times of day and meteorological conditions, exemplarily for the sub daily model setup with fixed time spans (green) compared to the model setup with source regions (blue) (cf. Table 1). The bars indicate the number of available observations in the DWD data for the respective situation. The numbers above the bars indicate the percentage of HOBO loggers with at least one observation for the respective situation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## References


Aguilar, E., Auer, I., Brunet, M., Peterson, T.C., Wieringa, J., 2003. Guidelines on Climate Metadata and Homogenization. WMO/TD No. 1186. World Meteorological Organization (53 pp).

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723.

Alcoforado, M.-J., Andrade, H., 2006. Nocturnal urban heat island in Lisbon (Portugal): main features and modelling attempts. Theor. Appl. Climatol. 84, 151–159.

Alonso, M.S., Fidalgo, M.R., Labajo, J.L., 2007. The urban heat island in Salamanca (Spain) and its relationship to meteorological parameters. Clim. Res. 34, 39–46.

Arnds, D., Böhner, J., Bechtel, B., 2017. Spatio-temporal variance and meteorological drivers of the urban heat island in a European city. Theor. Appl. Climatol. 128, 43–61.

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int. J. Climatol. 23, 1–26.

Arsanjani, J.J., Barron, C., Bakillah, M., Helbich, M., 2013. Assessing the Quality of OpenStreetMap Contributors Together with their Contributions. AGILE 2013, Leuven, Belgium.

Augsburg, Stadt, 2015. Hrsg.. Statistisches Jahrbuch 2014. (Augsburg, 187 pp).

Augsburg, Stadt, 2017. Hrsg.. Strukturatlas der Stadt Augsburg 2016. (Augsburg, 148 pp).

Basu, R., 2009. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. Environ. Health 8, 40.

Beck, C., Breitner, S., Cyrys, J., Hald, C., Hartz, U., Jacobeit, J., Richter, K., Schneider, A., Wolf, K., 2015. Statistical modeling of urban air temperature distributions under different synoptic conditions. Geophys. Res. Abstr. 17 (EGU2015-15292).
Beck, C., Straub, A., Breitner, S., Cyrys, J., Philipp, A., Rathmann, J., Schneider, A., Wolf, K., Jacobeit, J., 2018. Air temperature characteristics of local climate zones in the Augsburg urban area (Bavaria, Southern Germany) under varying synoptic conditions. Urban Clim. 25, 152–166.
Bernard, J., Musy, M., Calmet, I., Bocher, E., Kéravec, P., 2017. Urban heat island temporal and spatial variations: empirical modelling from geographical and meteorological data. Build. Environ. 125, 423–438.
Bottyán, Z., Unger, J., 2003. A multiple linear statistical model for estimating the mean maximum urban heat island. Theor. Appl. Climatol. 75, 233–243.
Bottyán, Z., Kircsi, A., Szegedi, S., Unger, J., 2005. The relationship between built-up areas and the spatial development of the mean maximum urban heat island in Debrecen, Hungary. Int. J. Climatol. 25, 405–418.
Bowler, D.E., Buyung-Ali, L., Knight, T.M., Pullin, A.S., 2010. Urban greening to cool towns and cities: A systematic review of the empirical evidence. Landsc. Urban Plan. 97, 147–155.
Breiman, L., 2001. Random Forests. In: Machine Learning. 45. pp. 5–32.
Buttstädt, M., Schneider, C., 2014. Thermal load in a medium-sized european city using the example of Aachen, Germany. Erdkunde 68 (2), 71–83.
Chapman, S., Watson, J.E.M., Salazar, A., Thatcher, M., McAlpine, C.A., 2017. The impact of urbanization and climate change on urban temperatures: a systematic review. Landsc. Ecol. 32 (10), 1921–1935.
Chow, W.T.L., Roth, M., 2006. Temporal dynamics of the urban heat island of Singapore. Int. J. Climatol. 26, 2243–2260.
Coutts, A.M., Tapper, N.J., Beringer, J., Loughnan, M., Demuzere, M., 2012. Watering our cities: The capacity for Water Sensitive Urban Design to support urban cooling and improve human thermal comfort in the Australian context. Prog. Phys. Geogr. 37 (1), 2–28.
Deilami, K., Kamruzzaman, M., Liu, Y., 2018. Urban heat island effect: a systematic review of spatio-temporal factors, data, methods, and mitigation measures. Int. J. Appl. Earth Obs. Geoinf. 67 (1), 30–42.
Déqué, M., 2003. Continuous Variables. In: Jolliffe, I.T., Stephenson, D.B. (Eds.), Forecast Verification. A Practitioner's Guide in Atmospheric Science (Chichester, 240 pp).
DWD Climate Data Centre, 2017. ftp://ftp-cdc.dwd.de/pub/CDC/ (January 2017).
EEA (European Environment Agency), 2010. Urban Atlas. (Koppenhagen, 24 pp).
EEA (European Environment Agency), 2011. Mapping Guide for a European Urban Atlas. (30 pp.).
EEA (European Environment Agency), 2016. Urban Atlas – LUZ Delivery Report – Augsburg. http://www.eea.europa.eu/data-and-maps/data/urban-atlas (24.02.2016).
Eliasson, I., Svensson, M.K., 2003. Spatial air temperature variations and urban land use – a statistical approach. Meteorol. Appl. 10, 135–149.
Eliasson, I., Offerle, B., Grimmond, C.S.B., Lindqvist, S., 2006. Wind fields and turbulence statistics in an urban street canyon. Atmos. Environ. 40, 1–16.
Erell, E., Williamson, T., 2007. Intra-urban differences in canopy layer air temperature at a mid-latitude city. Int. J. Climatol. 27, 1243–1255.
Estévez, J., Gavilán, P., Giráldez, J.V., 2011. Guidelines on validation procedures for meteorological data from automatic weather station. J. Hydrol. 402, 144–154.
Fenner, D., Meier, F., Bechtel, B., Otto, M., Scherer, D., 2017. Intra and inter 'local climate zone' variability of air temperature as observed by crowdsourced citizen weather stations in Berlin, Germany. Meteorol. Z. 26 (5), 525–547.
Friedberg, Stadt, 2018. Daten und Fakten. https://www.friedberg.de/staticsite/staticsite.php?menuid=95&topmenu=1485&keepmenu=inactive (22.02.2018).
Gandin, L.S., 1988. Complex quality control of meteorological observations. Mon. Weather Rev. 116, 1137–1156.
Gedzelman, S.D., Austin, S., Cermak, R., Stefano, N., Partridge, S., Quesenberry, S., Robinson, D.A., 2003. Mesoscale aspects of the urban Heat Island around new York City. Theor. Appl. Climatol. 75, 29–42.
Geofabrik, 2016. OpenStreetMap Data Extracts. www.geofabrik.de (20.05.2016).
Gersthofen, Stadt, 2018. Daten & Zahlen. http://www.gersthofen.de/seite/stadt-gersthofen/daten-zahlen.php (22.02.2018).
Goosen, H., de Groot-Reichwein, M.A.M., Masselink, L., Koekoek, A., Swart, R., Bessembinder, J., Witte, J.M.P., Stuyt, L., Blom-Zandstra, G., Immerzeel, W., 2014. Climate adaptation Services for the Netherlands: an operational approach to support spatial adaptation planning. Reg. Environ. Chang. 14, 1035–1048.
Grömping, U., 2006. Relative importance for linear regression in R: the package relaimpo. J. Stat. Softw. 17 (1), 1–27.
Grömping, U., 2007. Estimators of relative importance in linear regression based on variance decomposition. Am. Stat. 61 (2), 139–147.
Guo, Y., Gasparrini, A., Armstrong, B., Li, S., Tawatsupa, B., Tobias, A., Lavigne, E., De Sousa Zanotti Stagliorio Coelho, M., Leone, M., Pan, X., Tong, S., Tian, L., Kim, H., Hashizume, M., Honda, Y., Guo, Y.-L.L., Wu, C.-F., Punnasiri, K., Yi, S.-M., Michelozzi, P., Nascimento Saldiva, P.H., Williams, G., 2014. Global variation in the effects of ambient temperature on mortality: a systematic evaluation. Epidemiology 25 (6), 781–789.
Haklay, M., Weber, P., 2008. OpenStreetMap: user-generated street maps. IEEE Pervasive Computing 7 (4), 12–18.
Hart, M.A., Sailor, D.J., 2009. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. Theor. Appl. Climatol. 95, 397–406.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning – Data Mining, Inference, and Prediction, 2nd edition. (New York, 739 pp).
Hathway, E.A., Sharples, S., 2012. The interaction of rivers and urban form in mitigating the urban Heat Island effect: a UK case study. Build. Environ. 58, 14–22.
He, B.-J., 2018. Potentials of meteorological characteristics and synoptic conditions to mitigate urban heat island effects. Urban Clim. 24, 26–33.
He, J.F., Liu, J.Y., Zhuang, D.F., Zhang, W., Liu, M.L., 2007. Assessing the effect of land use/land cover change on the change of urban heat island intensity. Theor. Appl. Climatol. 90, 217–226.
Heusinkveld, B.G., Steeneveld, G.J., van Hove, L.W.A., Jacobs, C.M.J., Holtslag, A.A.M., 2014. Spatial variability of the Rotterdam urban heat island as influenced by urban land use. Journal of Geophysical Research: Atmospheres 119, 677–692.
Hinkel, K.M., Nelson, F.E., Klene, A.E., Bell, J.H., 2003. The urban heat island in winter at Barrow, Alaska. Int. J. Climatol. 23, 1889–1905.
Hjort, J., Suomi, J., Käyhkö, J., 2011. Spatial prediction of urban-rural temperatures using statistical methods. Theor. Appl. Climatol. 106, 139–152.
Ho, H.C., Knudby, A., Sirovya, P., Xu, Y., Hodul, M., Henderson, S.B., 2014. Mapping maximum urban air temperature on hot summer days. Remote Sens. Environ. 154, 38–45.
Hoffmann, P., Krueger, O., Schlünzen, K.H., 2012. A statistical model for the urban heat island and its application to a climate change scenario. Int. J. Climatol. 32, 1238–1248.
Jáuregui, E., 1975. Microclima del bosque de Chapultepec, Bulletin No.6. Instituto de Geografía, University of Mexico.
Jongtanom, Y., Kositanont, C., Baulert, S., 2011. Temporal variations of urban Heat Island intensity in three major cities, Thailand. Mod. Appl. Sci. 5 (5), 105–110.
Ketterer, C., Matzarakis, A., 2015. Comparison of different methods for the assessment of the urban heat island in Stuttgart, Germany. Int. J. Biometeorol. 59, 1299–1309.
Kim, Y.-H., Baik, J.-J., 2002. Maximum urban Heat Island intensity in Seoul. J. Appl. Meteorol. 41, 651–659.
Kim, Y.-H., Baik, J.-J., 2004. Daily maximum urban heat island intensity in large cities of Korea. Theor. Appl. Climatol. 79, 151–164.
Kim, H., Gu, D., Kim, H.Y., 2018. Effects of urban Heat Island mitigation in various climate zones in the United States. Sustain. Cities Soc. 41, 841–852.
Konarska, J., Holmer, B., Lindberg, F., Thorsson, S., 2015. Influence of vegetation and building geometry on the spatial variations of air temperature and cooling rates in a high-latitude city. Int. J. Climatol. 36 (5), 2379–2395.
Königsbrunn, Stadt, 2018. Daten & Fakten. https://www.koenigsbrunn.de/meine-stadt/wissenswertes/zahlen-daten/ (22.02.2018).
Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15 (3), 259–263.
Kuttler, W., Barlag, A.-B., Roßmann, F., 1994. Study of the thermal structure of a town in a narrow valley. Atmos. Environ. 30 (3), 365–378.
LDBV (Landesamt für Digitalisierung, Breitband und Vermessung), 2016. Geländemodell. Ganz Bayern in 3D - vom Main bis zur Zugspitze. http://www.ldbv.bayern.de/produkte/3dprodukte/gelaende.html (19.11.2016).
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. vol. 2/3. R News, pp. 18–21.
Makido, Y., Shandas, V., Ferwati, S., Sailor, D., 2016. Daytime variation of urban Heat Islands: the case study of Doha, Qatar. Climate 4 (2), 32.
Matzarakis, A., Rutz, F., Mayer, H., 2007. Modelling radiation fluxes in simple and complex environments – application of the RayMan model. Int. J. Biometeorol. 51,

323–334.
Matzarakis, A., Röckle, R., Richter, C.-J., Höfl, H.-C., Steinicke, W., Streifeneder, M., Mayer, H., 2008. Planungsrelevante Bewertung des Stadtklimas - Am Beispiel von Freiburg im Breisgau. Gefahrstoffe – Reinhaltung der Luft 68 (7–8), 334–340.
Morris, C.J.G., Simmonds, I., Plummer, N., 2001. Quantification of the influences of wind and cloud on the nocturnal urban Heat Island of a large City. J. Appl. Meteorol. 40, 169–182.
Murphy, A.H., 1988. Skill scores based on the mean square error and their relationship to the correlation coefficient. Mon. Weather Rev. 116, 2417–2424.
Nakamura, Y., Oke, T.R., 1988. Wind, temperature and stability conditions in an east-west oriented urban canyon. Atmos. Environ. 22 (12), 2691–2700.
Nicholls, N., Skinner, C., Loughnan, M., Tapper, N., 2008. A simple heat alert system for Melbourne, Australia. Int. J. Biometeorol. 52, 375–384.
Oke, T.R., 1981. Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. J. Climatol. 1, 237–254.
Oke, T.R., 1982. The energetic basis of the urban heat island. Q. J. R. Meteorol. Soc. 108 (455), 1–24.
Oke, T.R., 1988. The urban energy balance. Prog. Phys. Geogr. 12, 471–508.
Oke, T.R., 1993. Boundary Layer Climates, 2. ed. (London, 435 pp).
Oláh, A.B., 2012. The possibilities of decreasing the urban heat island. Appl. Ecol. Environ. Res. 10 (2), 173–183.
Onset, 2010. HOBO Pro v2 (U23-00x) Manual. http://www.onsetcomp.com/files/manual_pdfs/10694-P%20MAN-U23.pdf (09.01.2018).
Pauleit, S., Ennos, R., Golding, Y., 2005. Modeling the environmental impacts of urban land use and land cover change—a study in Merseyside, UK. Landsc. Urban Plan. 71, 295–310.
Roth, M., 2007. Review of urban climate research in (sub)tropical regions. Int. J. Climatol. 27, 1859–1873.
Schatz, J., Kucharik, C.J., 2014. Seasonality of the urban Heat Island effect in Madison, Wisconsin. J. Appl. Meteorol. Climatol. 53 (10), 2371–2386.
Shafer, M.A., Fiebrich, C.A., Arndt, D.S., Fredrickson, S.E., Hughes, T.W., 2000. Quality assurance procedures in the Oklahoma Mesonetwork. J. Atmos. Ocean. Technol. 17, 474–494.
Shi, Y., Katzschner, L., Ng, E., 2018. Modelling the fine-scale spatiotemporal pattern of urban heat island effect using land use regression approach in a megacity. Sci. Total Environ. 618, 891–904.
Shojaei, P., Gheysari, M., Myers, B., Eslamian, S., Shafieiyoun, E., Esmaeili, H., 2017. Effect of different land cover/use types on canopy layer air temperature in an urban area with a dry climate. Build. Environ. 125, 451–463.
Skarbit, N., Stewart, I.D., Unger, J., Gál, T., 2017. Employing an urban meteorological network to monitor air temperature conditions in the 'local climate zones' of Szeged, Hungary. Int. J. Climatol. 37, 582–596.
Spronken-Smith, R.A., Oke, T.R., 1998. The thermal regime of urban parks in two cities with different summer climates. Int. J. Remote Sens. 19 (11), 2085–2104.
Stadtbergen, Stadt, 2018. Stadtbergen in Zahlen. http://www.stadtbergen.de/index.php?id=295 (22.02.2018).
Steeneveld, G.J., Koopmans, S., Heusinkveld, B.G., Theeuwes, N.E., 2014. Refreshing the role of open water surfaces on mitigating the maximum urban heat island effect. Landsc. Urban Plan. 121, 92–96.
Stewart, I.D., 2011. A systematic review and scientific critique of methodology in modern urban heat island literature. Int. J. Climatol. 31, 200–217.
Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93, 1879–1900.
Suomi, J., Käyhkö, J., 2012. The impact of environmental factors on urban temperature variability in the coastal city of Turku, SW Finland. Int. J. Climatol. 32, 451–463.
Svensson, M.K., 2004. Sky view factor analysis – implications for urban air temperature differences. Meteorol. Appl. 11, 201–211.
Szymanowski, M., Kryza, M., 2012. Local regression models for spatial interpolation of urban heat island—an example from Wrocław, SW Poland. Theor. Appl. Climatol. 108, 53–71.
Tong, S., Wong, N.H., Jusuf, S.K., Tan, C.L., Wong, H.F., Ignatius, M., Tan, E., 2018. Study on correlation between air temperature and urban morphology parameters in built environment in northern China. Build. Environ. 127, 239–249.
Tyson, P.D., Du Toit, W.J.F., Fuggle, R.F., 1972. Temperature structure above cities: review and preliminary findings from the Johannesburg urban heat island project. Atmos. Environ. 6, 533–542.
Unger, J., 2004. Intra-urban relationship between surface geometry and urban heat island: review and new approach. Clim. Res. 27, 253–264.
Unger, J., Sümeghy, Z., Gulyás, Á., Bottyán, Z., Mucsi, L., 2001. Land-use and meteorological aspects of the urban heat island. Meteorol. Appl. 8, 189–194.
Upmanis, H., Eliasson, I., Lindqvist, S., 1998. The influence of green areas on nocturnal temperatures in a high latitude city (Göteborg, Sweden). Int. J. Climatol. 18, 681–700.
van Hove, L.W.A., Jacobs, C.M.J., Heusinkveld, B.G., Elbers, J.A., van Driel, B.L., Holtslag, A.A.M., 2015. Temporal and spatial variability of urban heat island and thermal comfort within the Rotterdam agglomeration. Build. Environ. 83, 91–103.
Vicente-Serrano, S.M., Cuadrat-Prats, J.M., Saz-Sánchez, M.A., 2005. Spatial patterns of the urban heat island in Zaragoza (Spain). Clim. Res. 30, 61–69.
Wang, X., Cheng, H., Xi, J., Yang, G., Zhao, Y., 2018. Relationship between park composition, vegetation characteristics and cool island effect. Sustainability 10 (3), 587.
Wienert, U., Kreienkamp, F., Spekat, A., Enke, W., 2013. A simple method to estimate the urban heat island intensity in data sets used for the simulation of the thermal behaviour of buildings. Meteorol. Z. 22 (2), 179–185.
Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, 2nd edition. (Amsterdam, 627 pp).
Wouters, H., De Ridder, K., Poelmans, L., Willems, P., Brouwers, J., Hosseinzadehtalaei, P., Tabari, H., Vanden Broucke, S., van Lipzig, N.P.M., Demuzere, M., 2017. Heat stress increase under climate change twice as large in cities as in rural areas: a study for a densely populated midlatitude maritime region. Geophys. Res. Lett. 44, 8997–9007.
Yang, X., Li, Y., Luo, Z., Chan, P.W., 2017. The urban cool island phenomenon in a high-rise high-density city and its mechanisms. Int. J. Climatol. 37, 890–904.
Yokobori, T., Ohta, S., 2009. Effect of land cover on air temperatures involved in the development of an intra-urban heat island. Clim. Res. 39, 61–73.
Zielstra, D., Zipf, A., 2010. Quantitative studies on the data quality of OpenStreetMap in Germany. In: Sixth International Conference GIScience 2010. Zuerich, Switzerland.