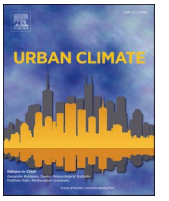
Urban Climate 61 (2025) 102400

Contents lists available at ScienceDirect

Urban Climate

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



Climate sensitive designs for policy makers: How LES model resolution affects accuracy in capturing urban micro-scale weather during heatwaves☆

Leyla Sungur[a,*], Wolfgang Babel[a,b], Eva Späte[a], Johann Schneider[a,c], Christoph K. Thomas[a,b]

[a] University of Bayreuth, Micrometeorology Group, Bayreuth, Germany
[b] University of Bayreuth, BayCEER, Bayreuth Center for Ecology and Environmental Research, Bayreuth, Germany
[c] University of Bayreuth, Plant Ecology, Bayreuth, Germany

ARTICLE INFO

Keywords:
LES simulation
Model validation
Multiresolution decomposition
PALM-4 U
Urban heat modeling
Urban health
Recommendation for policy makers

ABSTRACT

Climate sensitive designs have been implemented recently in science to fill the niche of developing scientific tools to help mitigating urban heat island effects. A model capable of identifying hot and cool spots and testing adaptation and mitigation strategies to form recommendations for policy makers is in high demand. We present a novel two-step validation approach using 1) absolute comparison and 2) space-time evaluation of model performance across resolutions and against observations ensemble-averaged for representative urban microclimate types. Two Large Eddy Simulation (LES) models with 5 m and 20 m resolution are evaluated against a 14-point measurement station network during an extreme heatwave in Germany. To show space time behavior relationship between model and measurements, multiresolution decomposition (MRD) was used to investigate air temperature, specific humidity, and wind speed across time scales ranging from 11.25 to 180 min. The MRD analysis revealed strong correlations between stations with similar microclimatic conditions. Increasing model resolution did not uniformly improve accuracy across all climate elements; wind speed showed the greatest benefit from higher resolution. Absolute comparisons between simulations and observations demonstrated well-represented diurnal cycles for 5 m wind speed, whereas scalar variables remained sensitive to land surface parameterization. The simulated specific humidity exhibited deviations from measurements and nighttime air temperature tended to be overestimated. Space-time behavior instead was generally well captured except for nighttime specific humidity. This study contributes to correctly identifying and quantifying urban heat effects during extremes and supports urban planners and decision-makers in evaluating modeling tools for spatial planning.

1. Introduction

Temperature extremes pose a greater risk of heat stress and increased mortality rates for urban (Clarke, 1972; Hayhoe et al., 2010) than for rural residents. Although urban areas cover only 3 % of the global land surface (Liu and Morawska, 2020), more than 50 % of

☆ This article is part of a Special issue entitled: 'Urban Wind Environment' published in Urban Climate.

* Corresponding author.
E-mail address: leyla.sungur@uni-bayreuth.de (L. Sungur).

https://doi.org/10.1016/j.uclim.2025.102400
Received 7 January 2025; Received in revised form 19 March 2025; Accepted 25 March 2025
Available online 16 April 2025

the global population lives in cities. Until 2050 this share is expected to increase to 68 % (UN, 2018). The bioclimate and therefore the human thermal comfort are severely threatened by the urban heat island effect (UHI), global climate change, increasing heat waves, further surface sealing, and redensification of urban areas (Nowak and Greenfield, 2020; Revi et al., 2014; Santamouris, 2023). Under very hot and humid conditions, the ability of the human body to regulate the core temperature and dissipate heat via sweat evaporation is reduced, provoking heat stress (Parsons, 2002). This represents a health risk especially for people with respiratory or cardiovascular diseases. An 1 K increase of indoor temperature doubles the risk of death for elderly people (Buchin et al., 2016)). Additionally, there is an increased risk of pre-term birth during hot days (Kloog, 2019). Meteorological conditions associated with high radiation loads and/or low wind speed can exacerbate heat stress. To protect human well-being today and in the future, it is crucial to better understand how to mitigate those temperature extremes in cities. Out of this urgency, the importance of moderate urban climate is recognized by political authorities and state-owned institutions cooperating with research institutions. Since cities are constantly growing and transforming in response to the residents' needs, a planned climate friendly infrastructure is not easy to establish. This demand forms a niche for research to assess local meteorological effects that affect the lower atmosphere ground layer in which human activity takes place. Research on micro-meteorological conditions during hot days is crucial to formulate recommendations for mitigation strategies by means of urban planning measures to counter heat stress in urban areas. In the past, research strongly focused on the UHI effects in larger cities. However, not only large cities suffer from a significant UHIs. The magnitude depends on a series of factors such as geographical position, population size, topography, building geometry, building material and anthropogenic heat release (Mehrotra et al., 2019). Cities that share the same climate zone and do have equal numbers of population do not necessarily show a similar magnitude of the UHI effect. Cities can achieve extraordinary high or low effects due to their spatial and topographical setting or because they produce extreme anthropogenic heat release. For moderate city climate in European summers, park cool islands (PCI) as recreational areas are of high importance during the day. Nighttime hot spots serve as key areas affecting urban health. To map and counter the urban heat island effect, urban climate modeling has become a frequently employed tool to investigate the impact and demonstrate applied research recently.

Numerical high-resolution simulations of meteorological processes can be used to simulate urban development strategies aiming at mitigating the adverse effects of hot spells in cities. Large Eddy Simulations (LES) models directly resolve large-scale turbulence and only parameterize the smaller eddies. LES models are therefore known to require larger computing power but also deliver more accurate results (Martilli and Santiago, 2009). A great asset of LES is their capability of mapping changes of individual climate elements and energy and mass fluxes in the urban body. Complex LES turbulence models are found to be most accurate for local scale implementations. LES is particularly advantageous in scenarios with complex geometries and unsteady flows, where time-dependent and three-dimensional characteristics significantly influence the results (Ciofalo, 2022). Their implementation of complex boundary conditions is still not widely used correctly (Kubilay et al., 2020). To date, most climate models do not include explicit modeling of urban areas as it is at too high cost to investigate regional climate trends at very high spatial resolution (Daniel et al., 2018). Yet, explicit modeling of time-dependent, three-dimensional simulations using the LES technique surpass the performance of idealized LES models. This is because fully resolved LES can capture the intricate dynamics and transient behaviors of turbulent flows more accurately than idealized versions, which may rely on greater approximations and simplifications. This distinction is crucial for applications requiring high-fidelity simulations of complex fluid behaviors. Despite their identified use as advanced tools for city planning, LES models often lack meaningful validation approaches. In his review, (Toparlar et al., 2017) found that the majority of modeling studies of real urban climate are not validated, despite validation being essential to demonstrate the meaningfulness of modeled results. Validation constitutes a crucial step in the application of a numerical model, involving the process of ascertaining how accurately simulations portray the real world, using empirical observations as a benchmark. Observation measurements are considered to be the most independent and complete validation approaches for LES applications (Kempf, 2008). In practice, insufficient statistical moments (mean and variance), not very sensitive to small amplitudes typical for the smallest scales, serve as a basis for comparison (Kempf, 2008). Hertwig et al. (2017) also identified the lack in detailed validation and questioned whether such LES provide a sufficiently accurate picture of the time-dependent structure of the turbulent flow. After Brozovsky et al. (2021) a majority of publications on thermal conditions in real urban areas provide a validation process, even though the degree of detail and number of climate variables in the validation varied significantly. In their study, Brozovsky et al. found that most study validate with air temperature as single climate element only, whereas few consider also wind speed and direction, but no quantitative analysis agreement was considered. Kellnerová et al. (2018) extended common validation technique by not only comparing the mean values but also implementing a multi-point time-series analysis by means of advanced statistical tools, which allowed to detect transient structures, their shape and also frequency distribution. An evaluation of the PALM 6.0 LES model by Resler et al. (2021) served as a pilot study for explicit LES validation and compared model results of surface temperature, air temperature, wind speed as well as wall heat flux and concentrations of NO_x and PM10 with measurement campaigns as time series comparisons.

Our study aims at evaluating to what extent high resolution LES modeling can be used as a planning aid to assist cities' authorities and policy makers in their decision making to adapt to extremely high temperatures. We use the July 2019 heatwave as a proxy for future events, exemplifying extreme weather conditions projected to increase in frequency and intensity due to climate change. Heatwaves frequently arise as a result of overarching weather patterns. Established mesoscale and macroscale circulation patterns in the atmosphere over Europe have long been systematically identified for the European region. Extreme events serve as critical indicators for testing the effectiveness of models in accurately capturing the tail distributions of researched data. By analysing the model's performance in simulating high-impact occurrences, we can assess its robustness and reliability.

We evaluate urban LES model simulations against measurements from a multi-station observational network reference. The measurement locations represent typical types of urban land surface cover and thus microclimatic hot spots and cool islands. We address the following research questions:

1 How different is the space-time correlation of the relevant climate elements between observations from the city-scale station network compared to the modeling using the urban LES for the three land surface cover subcategories?

2 Does a four-fold increase in model resolution improve model-observation agreement?

3 What are the mechanisms explaining the deviations between model resolutions and stations?

2. Data and methods

2.1. Study area and measurement network

Our simulated period is a 48 h period (2019-07-25 12:00 to 2019-07-27 12:00 CEST) covering extremely hot days of a heat wave lasting from 24 to 26 July 2019 in Germany (Bissolli et al., 2019). With an average air temperature of 10.3 ^{o}C $^{\circ}$C, 2019 was rated as the second warmest year in Germany since 1881 (Kaspar and Friedrich, 2019). Overall, the year was characterized by rather dry conditions, more sunshine than usual, and extreme heat waves during summer which caused many record breaking temperatures throughout Germany of which several persist so far as all-time heat records per city up to today (Kaspar and Friedrich, 2019). The all-time heat record measured in Germany was also part of the same heat event we analyzed: 42.6C was measured in Lingen on July 25th 2019 (Bissolli et al., 2019). The estimated death rate for heat periods in 2019 was estimated to 6900 which is a significantly high mortality compared to 1991–2021. As the years 2018 and 2020 also showed significant high mortality rates due to heat, showing consecutive high mortality rates has lead to an increase in estimated mortality rates of the decade 2012–2022 (Winklmayr et al., 2022) The city of Bayreuth, Germany (see Fig. 1) also reached its peak air temperature of 38.2 $^{\circ}$C for 2019 during the same heat event. It is a medium-sized town with nearly 75,000 citizens and 67 km^2. The average annual air temperature is 8.9C (1991–2020) (Wetterdienst, 2024). A network of 14 micro weather stations at structurally different urban sites (see Fig. 2a) was installed in the years 2018 and 2019 within the scope of the MiSKOR (Mitigation of urban climate and ozone risks) project to measure meteorological variables, and is operational to date. MiSKOR aims at analysing the microscale meteorological conditions in the city of Bayreuth to expand the

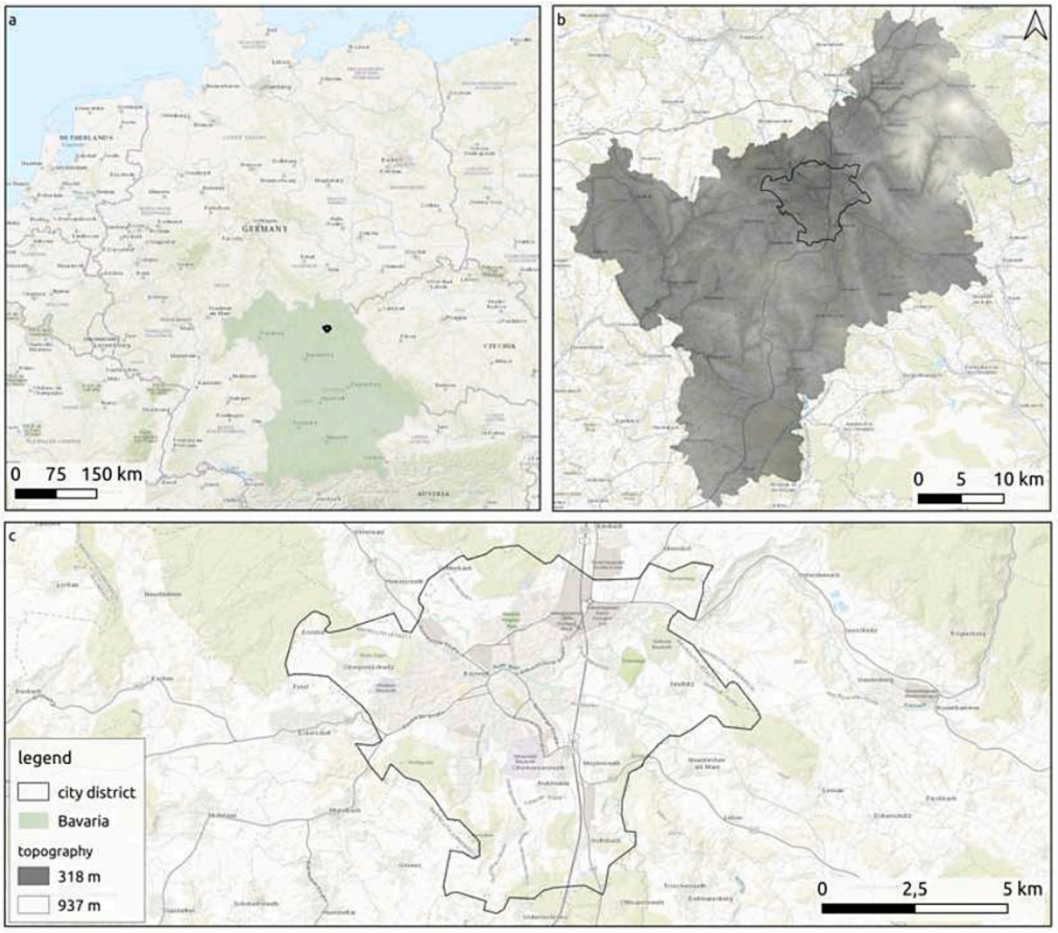


Fig. 1. Study area and topographical setting a) City district located in Bavaria, Germany, b) Topographical setting surrounding the city district, c) City district. (OpenStreetMap contributors, 2017).

knowledge on micrometeorological and microclimatic processes urban effects in mid-sized cities. An overview of the results from overall analysis of the measurements is given in Section 3 whereas the two-step validation is based on results from the 72 h heatwave period only. The high density of continuous weather observations makes the city of Bayreuth an ideal location for thorough validation of meteorological simulations against observed data. Micro-weather stations (ATMOS-41, Meter Group USA, Germany with data logger EM60G) were mounted at 3.5 m height above the surface to existing urban structures such as street lights or traffic signs. The following meteorological variables were of interest to this study and have been measured several times per minute and recorded as 5-min averaging or sums: absolute humidity, barometric pressure, air temperature, incoming shortwave radiation, precipitation, wind direction, and wind speed. Before installation, a side-by-side comparison of all instruments was conducted over 17 days, detecting a mean average deviation of $\leq < 0.1$ K for air temperature, and ≤ 0.06 ms^{-1} for wind speed across all stations. The measurement sites were chosen to represent the spatial variability of typical urban microclimates. They represent urban areas that vary in terms of sealing density, building height, proximity to main traffic axes, as well as to parks and water areas. For 2018 to 2020 the analysis of observations showed that the exact locations of urban hot and cool spots vary temporally within the study area across the course of the day and across weather patterns. For this reason, modeling approaches offer a good alternative to have spatially continuous data of climate elements.

The PALM-4 U LES model was set up to run a realistic replica of the cityscape of the study site (see Fig. 2 and Fig. 3). The research period chosen for simulation and model validation was characterized by an anticyclonic Fennoscandian high leading to clear-sky conditions at high air temperatures and low wind speeds. Three important climate elements were considered for the validation of the model: Air temperature, specific humidity, and airflow (hereby referred to wind direction and wind speed). These variables, together with solar radiation, are considered critical factors influencing human thermal comfort. The perceived temperature (PT) is an equivalent temperature derived from a comprehensive heat budget model of the human body. It is specifically designed for outdoor conditions and is defined as the air temperature in a reference environment where the thermal sensation would match that of the actual environment (Staiger et al., 1997). Validation of the measured incoming shortwave radiation was not conducted, as the model directly used shortwave radiation data from the reference measurement station as input.

2.2. PALM-4 U model setup

The simulations were done using the PALM model (23.10) (Maronga et al., 2015, 2019). This LES model solves the Navier-Stokes equations while separating large-scale turbulence which is directly resolved from the subgrid small-scale turbulence which is fully parameterized. Besides the core components the urban application module PALM-4 U is used to research atmospheric flows, temperature and humidity in the urban setting of this study (Maronga et al., 2019). The PALM-4 U setup offers realistic topographical integration of spatial information and time-dependent steering of synoptic forcing conditions. The initialization and runtime module steer basic model operations and employ settings for input and output LES. Here a spinup time of 24 h is embedded to initialize adjustments between the surface, soil- and wall-layer temperatures to prevailing atmospheric conditions before simulation start (Maronga et al., 2019). A Land Surface Module (LSM) summarizes information about land-use data for which cartographic basis was provided by the State Institute for Geodesy of Bavaria. A static driver covers the information of the LSM and consists of a multi-layer soil model, predicting soil temperature and moisture content, and a solver for the energy balance at the surface. A nesting module includes one root domain in 20 m resolution and one nested inner domain in 5 m resolution that covers the city core area of 2 km $\times$ 2 km in a one way nesting mode. An urban surface module simulates building surface interactions using a tile approach. It consists of a

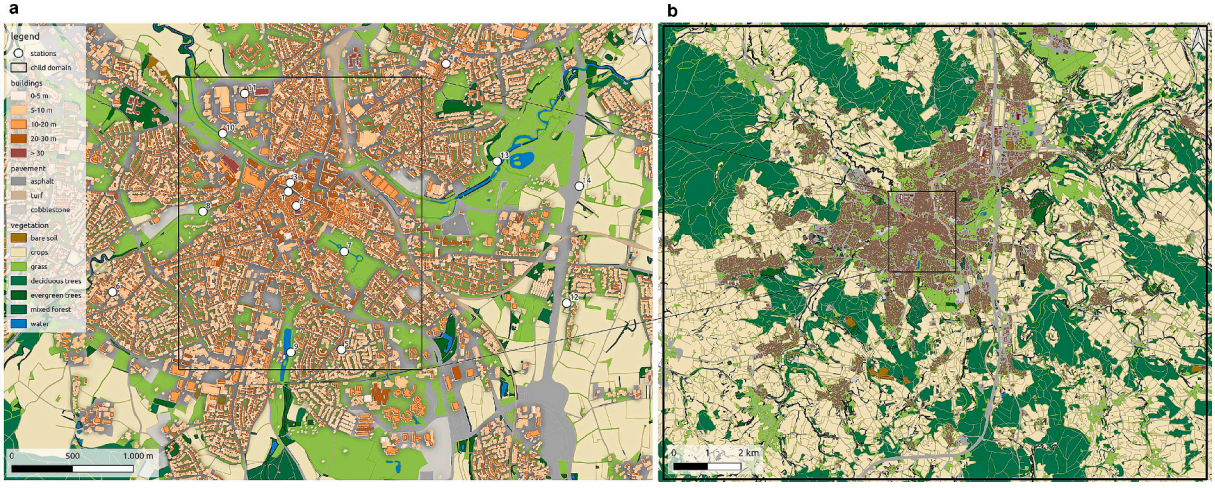


Fig. 2. Spatial distribution map of inner domain (a) embedded in its parent domain (b). Modeled domain boundaries are presented by black border. The parent domain spans the surrounding mountain ridges, which influence the topographical setting of the city located in the valley. The inner domain maps the 14 stations of the measurement network used for validation of both, parent and inner domain. The station labels indicate the corresponding station number. (Vermessungsverwaltung, 2024).

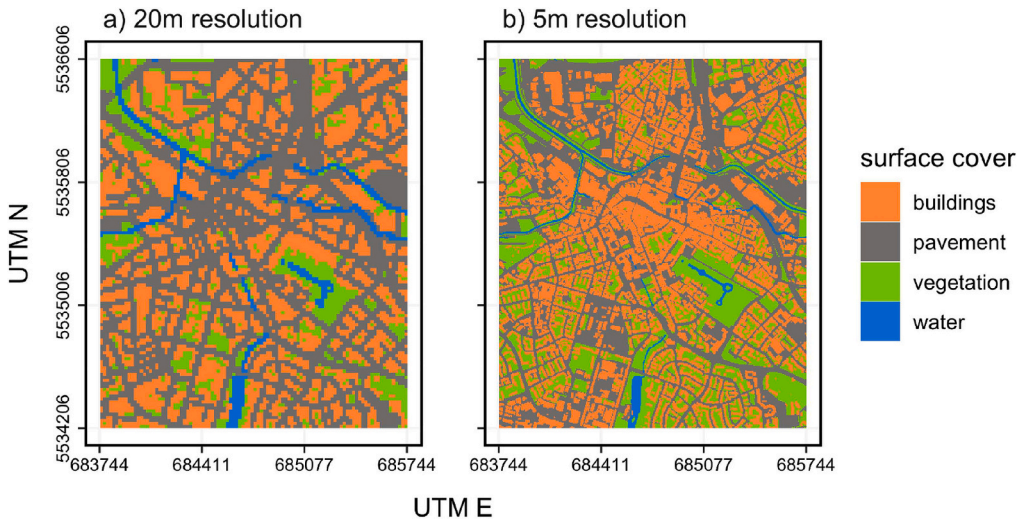


Fig. 3. Example of depicted differences in modeled resolution of 20 m (a) and 5(m) for the center of the city.

multi layer wall and soil model, predicting wall and soil temperature and moisture content (Maronga et al., 2019). A solver for the energy balance is applied to calculate the temperature of the surface and the skin layer for each urban surface tile individually. The synoptic forcing conditions at start are provided externally. Nesting and radiation modules are used including the process tool Mesoscale interface for initializing and forcing PALM (INIFOR). INIFOR allows to link the output data from the high resolution regional model Consortium for small-scale modeling (COSMO-DE2), designed by the German Meteorological Service, to PALM. The hourly COSMO-DE2 model output included velocity components, specific humidity, potential temperature, perturbation pressure, soil moisture and soil temperature. The incoming shortwave radiation was taken from a measurement station situated in the ecological botanical garden, Bayreuth within the simulation domain. Together with the COSMO-DE2 data, this time-dependent boundary data was mapped onto the lateral and top boundaries of the PALM model domain. These boundaries are non-cyclic, allowing simultaneous in- and outflow based on input data. This way, it is possible to account for changing synoptic conditions during the simulation without further tendency terms (Maronga et al., 2019). The root domain of 20 m resolution spans an area of 810 × 675 × 3500 pixels. For z-values greater than 500, the grid was stretched by a factor of dz.(k + 1) = dz.(k)1:08. The inner domain is shifted by 4420 cells from the lower left corner in x-direction and by 2980 from the lower left in y-direction of the root domain. It is a quadratic shape of 400 × 400 × 60 pixels (see Table 1 and Fig. 2). Each pixel from each domain covers information about topography, coordinates, and land-use type according to PALM-4 U classifications or building. For each surface element of the model grid, the equations for sensible heat, aerodynamic resistance, ground heat, thermal conductivity and and latent heat flux are solved locally and respectively from the energy balance. The chosen resolutions were found to be representative values providing an optimal trade-off between accuracy, covering city districts and computational feasibility (Vogel et al., 2022). (See Figs. 3 and 5.)

Model output for the virtual stations was retrieved in 2D cross sections in the two model domains. The cross sections were specified according to the locations of the observational network. The output for cross-sections is subject to minimal imprecision, as scalar variables are sampled at the center of a pixel cell, while u,v and w are taken at the edges. The cross-section closest to the measurement point was selected, and spatial deviations at pixel edges were averaged with the neighboring cell to enhance accuracy. For the 20 m resolution, the data output height is calculated at 10 m above the surface, which is as close as possible to the ground. For the 5 m resolution, the averages between the data output height levels of 2.5 and 7.5 m cross sections were applied. Model output data is given as 10 min statistics each consisting of 432 time steps. Observational data recorded at 5 min resolution were aggregated to create data matching the time stamps of the simulations.

Table 1
Model domain setup detailing the dimensions and parameters of root and inner domain.

	parent domain	inner domain	unit
Resolution	20	5	m
x-direction	810	1620	pixel
y-direction	675	1620	pixel
z-direction	64	72	pixel
z-height	3500	400	m
Origin	677,024	681,444	*E*-UTM
Origin	681,444	5,531,006	N-UTM

2.3. Categorization of the observed microclimates

The surface layer is influenced by roughness, thermal mixing, and moisture conditions at the surface. Thus, important micrometeorological processes influenced by present heterogeneous surface properties and building geometry need to be taken into consideration when analysing space time behavior. In order to map the spatial variability of microclimate within the city, we clustered the stations from the network into phenomenological microclimate groups. For this, we considered the surrounding area influencing each station's local weather. As a first approximation, an area considered for describing each measurement stations' microclimate was chosen based on the internal boundary layer height

$$z_{IBL} = 0.3 \cdot \sqrt{x}(m). \quad (1)$$

An internal boundary layer (IBL) forms when air moves across a surface change which leads to forming a new equilibrium layer in which the air adjusts to the changed surface properties such as brightness temperature and surface momentum roughness. The wind profile gradually adjusts as it moves downwind by the mean flow creating a layer of discontinuity reflecting the before- and post-change conditions called the internal boundary layer (Foken, 2017). This means that the exchange of energy and matter from the ground up to a up to z_{IBL} is affected by the new surface properties of the downwind area. The internal boundary layer height at the measurement location was equivalent to the installation height of the measurement devices with $z_{IBL} = 3.5$ m. Therefore, the distance x within which the atmospheric processes at the measurement location were influenced by the surface cover was estimated using Eq. 1 to 136 m. A square with a side length of 260 m centered at station's location was chosen for this approximation (see Fig. 4).

Secondly, we established a metric grading system that clusters the measurement stations based on land surface and infrastructural properties within the considered influencing area. The metric covers the proportion of different land cover types, the extent and arrangement of building cover, the average building height, and the median distance from each station to the nearest building in at least two different directions (see Fig. 6). The outcome per category and per station was subsequently clustered into three groups 'urban stations', 'semi-urban stations' and 'open-spaced' (see Fig. 6). The clustering was performed using the statistical k-means clustering, with each station being clustered after its nearest mean per each category. Cluster analysis revealed a clear fit for urban and open-spaced stations whereas the semi-urban stations were characterized by ambiguous results. We attribute this finding to the fact that highest variance per category is found for the semi-urban stations, resulting in an increased range of plausible results for this specific station group. The k-means clustering does not consider variability in variance and can be misleading. Incorporating the statistical two-point correlation coefficient into the metric grading system (Fig. 6 g) extends the analysis to a more comprehensive evaluation of meteorological state variables between stations. The magnitude of such processes steered by the energy flux components during the course of a day can vary greatly with the surface cover type, the available energy, and available moisture content. Like this, the correlation in between the stations reveals also strong similarity for stations within the semi-urban group. The fit between the metric grading clustering and the two-point correlation clustering is in good agreement except for stations 4, 10 and 11.

As a final result, the three categories urban, semi-open and open-spaced consist of the stations depicted after color in Fig. 6 after the comprehensive evaluation in Fig. 6 of all included categories a) - g).

2.4. Statistics used for model evaluation

Validation of LES against measurements shall be done utilizing state-of-the-art, time-resolved single-point measurements (Hertwig et al., 2017) whenever possible. Following (Hertwig et al., 2017) we pursued a multi-level validation concept. Initially, we conducted data analysis by computing absolute and relative differences between the modeled and observed climate elements for the two different model resolutions. Air temperature, specific humidity and wind were selected as the most relevant elements for human thermal comfort (Djongyang et al., 2010). Common validation metrics do not offer direct physical insight, making it difficult to infer the reasons for accuracy limitations. As a next level in the validation process we applied the stochastic multiresolution decomposition (MRD) technique and computed two-point correlation coefficients as described by (Howell and Mahrt, 1997) to characterize the space-

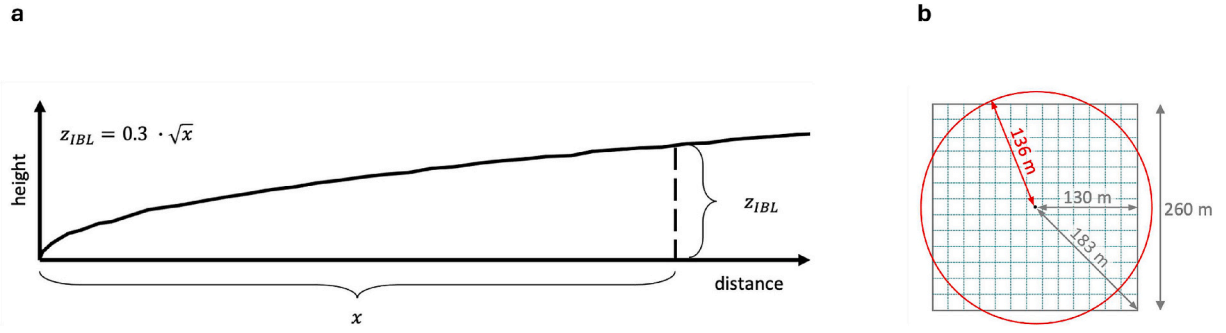


Fig. 4. a) sketch of the internal boundary layer and the relation between internal boundary layer height z_{IBL}(m) and distance x (m) after (Raabe and Foken, 2003) and b) sketch of the square with a side length of 260 m centered at a measurement station (black dot) at which the atmospheric processes at the measurement location are approximately influenced by the surrounding surface cover.

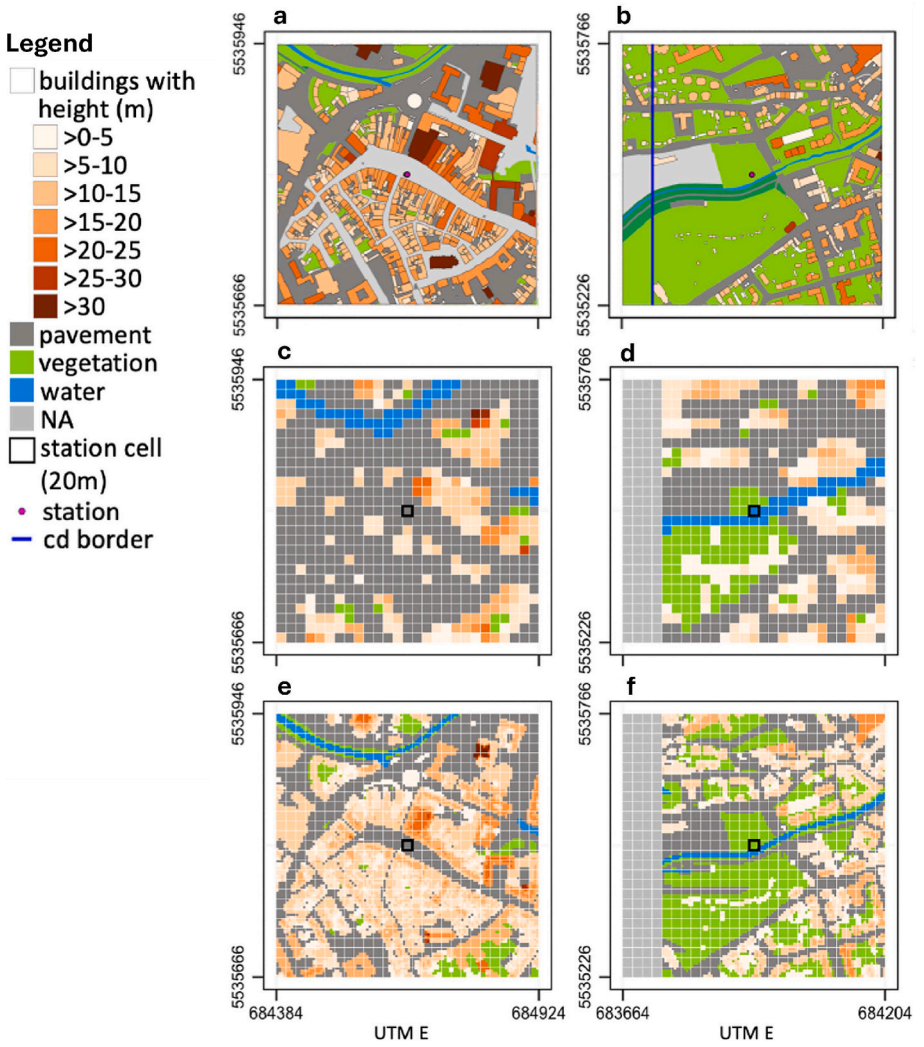


Fig. 5. Example representation of the area considered for each station's microclimate picturing station 1 (a, c, e) and station 8 (b, d, f). Differences between mapped data (a, b) modeled 20 m resolution (c, d) and modeled 5 m resolution (e, f) are presented.

	Stations	1	2	3	4	5	6	7	8	9	10	11	12	13	14
b	Land Cover Type in %														
	Pavement	95	95	95	50	15	30	<5	15	5	45	90	5	0	7
	Vegetation	5	5	5	50	85	70	95	85	70	45	10	85	95	93
	Water	0	0	0	0	0	0	10	0	25	10	0	0	15	0
c	Building Cover in %	85	75	80	50	15	20	10	15	15	10	40	5	1	1
d	Building Arrangement	in row	in row	in row	in row	spread	spread	spread	spread	spread	isolated	spread	isolated	isolated	isolated
e	mean building height	12	12	12	9	7	6	8	8	6	9	10	5	10	4
	distance to building	1.8, 9, 35	0	0	1.5, 3	11, 20, 16	4, 5, 22, 19	60, 67, 64, 69	58, 80, 82, 68, 138	35, 61, 94	21, 56, 91, 86	15, 13, 36	57, 49, 53, 64	113,111, 132	119, 180
f	median distance to building	9	0	0	2	16	12	68	80	61	71	15	55	113	150
g	data analysis														

urban-stations semi-urban stations open-spaced stations

Fig. 6. Classification metric for measurement stations. The 14 stations were allocated after the metrics on the left side to the three microclimatic categories 'urban', 'semi-urban' and 'open-spaced' based on their phenomenological similarities.

time behavior across the urban microclimates. In MRD, cospectra derived from the orthogonal Haar decomposition are related to variations in the Reynolds averaging length (Howell and Mahrt, 1997). We selected the MRD technique to offer direct physical insight into space-time behavior replication, rather than show the percentage of a fit. This approach aligns with the recommendation by Hertwig et al. (2017), who advocated for the use of such measures in combination with detailed point-by-point analyses to focus on space and time patterns or trends in the variables of interest reproduced by the LES. In spite of this recommendation, the approach was

rarely followed. Both the measurements and the modeled output at the measurement locations were analyzed using the MRD technique. By calculating the spatial two-point correlation coefficients at different temporal scales, coherences between all stations should be revealed. The datasets were segmented into dyadic increment periods of 2^n width, where n ranged from 0 to M. Here, $n = 0$ corresponds to the signal at its original recorded resolution of 10 min, and M represents the longest time interval, which in this study was 48 h. The 5 specific time scales were selected under consideration of sufficient data points starting at 180 min to 11.25 min cutting of still above $n = 0$. Each of the MRD modes M contains 2^{M-n} data points per interval. Therefore five scales of averaging lengths were used: 11.25 min, 22.5 min, 45 min, 90 min, 180 min. The MRD is calculated individually for the observational data, and the 5 m and 20 m resolution simulations. After the MRD was performed, the two-point correlation coefficients between any two stations were calculated for every MRD time scale across the entire dataset. The two-point correlation coefficients for the u and v-components are combined as $R_{uv}(r) = \left[\left(R_u^2 + R_v^2\right)/2\right]^{0.5}$ after (Mahrt et al., 2009). Information on the sign of the correlation is lost, providing only the magnitude of the overall correlation.

3. Results and discussion

For day and night time, stark mean differences in micro-scale weather within the city were found. We determined the average UHI for our study area to be on average 0.5 K warmer than the rural reference station located in the Ecological Botanical garden during daytime, and 2.0 K warmer at night. The largest mean spatial amplitudes show the UHI to increase to 1.0 K during daytime and 4.8 K at night. In general, vegetated sites were cooler than the highly sealed sites. This result is interpreted to be mostly an effect of increased heating by absoprtion of solar radiation at the sealed sites, and to a lower degree the result of increased cooling by evapotranspiration in the parks. In addition to the vegetation, effective ventilation was most relevant for cooling in these cases. To express measured heat variability within the city, the daily median was calculated for a 24 h interval. Within the city, large green areas were on average 0.8 K cooler than the other urban areas during daytime and 2.0 K cooler at night. On hot days this park cool island reached values of −2.0 K (day) and − 4.2 K (night). The core of the heat island changed its spatial location with the time of the day. In the morning station 4 was warmest, while in the afternoon highest air temperatures were observed for station 7. From early evening on the city center showed on average the highest temperatures: represented by station 1 and 2, which stayed warmest over night. The intensity of the inner-city air

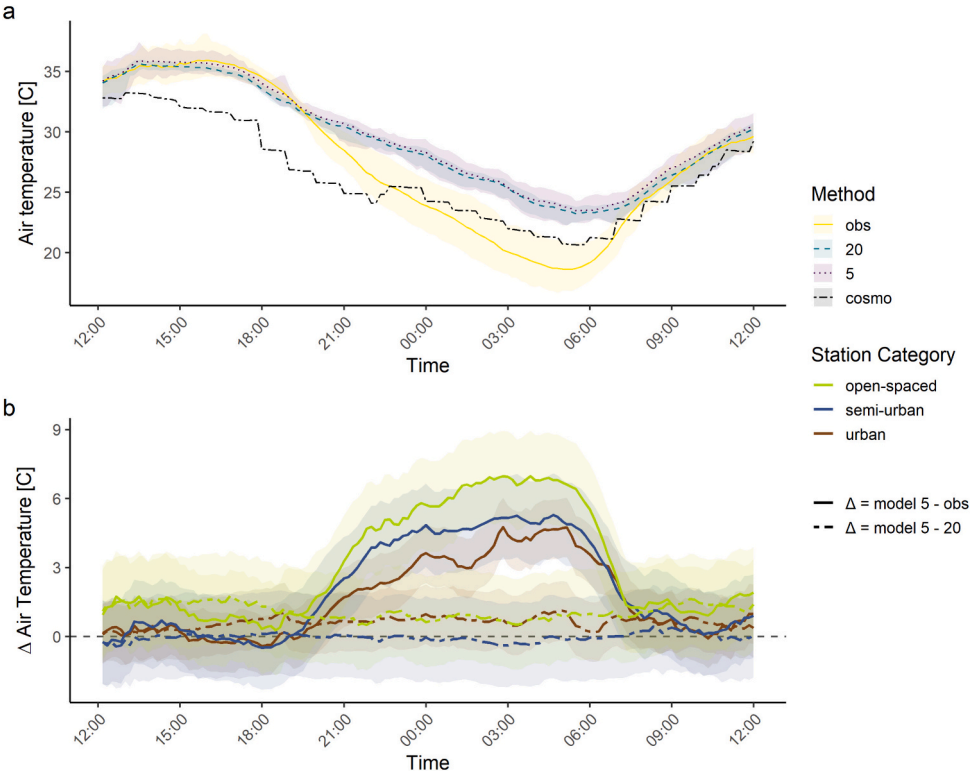


Fig. 7. Mean diurnal air temperature from day 1 to 3 (2019-07-25 - 2019-07-28) for a) observed (obs) and simulated specific humidity in 5 m (5) and 20 m (20) resolution and the modeled synoptic forcing boundary condition from COSMO-DE2 (cosmo) and b) mean differences between simulated 5 m resolution and observations (5 m - obs) and simulated 5 m resolution and simulated 20 m resolution (5 m - 20 m) for respective microclimatic subcategories from mrd analysis. The shaded area in the corresponding color indicates measured and modeled variability between the minimum and maximum values across all 14 stations over time, illustrating the spread of the data. The x-axis shows time in CEST, the y-axis shows a) air temperature and b) δ air temperature in celsius degree.

temperature differences depended on large-scale weather type. Highest network internal thermal differences were related to weather types characterized by anticyclones above Central Europe (weather types HM and BM), as well as westerly (WA) and north-westerly (NWA) weather conditions. The analyzed absolute air temperature differences within the network reached 4.2 K for these weather types.

3.1. *Validating the LES results for absolute differences and diurnal course*

The results of our analysis revealed differences between the measured and simulated values specific to each climate element. For specific humidity and air temperature our data revealed stark influence of simulated results by the coupled large scale model COSMO-DE2 (Fig. 7, Fig. 8 used to force the model. In contrast, the simulated horizontal wind was independent from that in the large-scale nesting model (Fig. 9. The modeled air temperature and wind were in good agreement with observations in terms of their temporal evolution during the day, as simulated values from both resolutions captured the observed diurnal pattern. A daytime temperature maximum is observed at 16:00 CEST. Most variability between observed stations appears during night when the radiative cooling processes of the surface are active. However, modeled air temperature overestimated the nighttime minimum at 5:30 by up to 4.9 K and thus largely lacked nighttime cooling (Fig. 7). Among the modeled station categories, urban stations show the smallest overestimation compared to semiurban and openspaced stations (Fig. 7). The decorrelation between simulated and observed values becomes apparent after sunset, when solar radiation ceases, and longwave radiative transfer prevails. Previous studies have experienced problems with nighttime cooling under comparison with real urban environments (Resler et al., 2021). We attribute neglected nighttime cooling to uncertainties in the radiation module and high heat diffusivity in the soil. This is further explained in Section 3.3. In contrast, the model-observation difference for specific humidity did not follow a distinctive diurnal course. The modeled and observed mean diurnal cycles showed little temporal coherence, but was rather characterized by time-independent discrepancies across the diurnal cycle. Overall, the model underestimated the specific humidity (Fig. 8) by average 1.2 gkg^{-1} for 20 m resolved data and 1.1 gkg^{-1} for 5 m resolved data. Temporal resolved differences were largest on the first day of the analyzed period between 6 PM and 2 PM the second day (Fig. 16). Specific humidity primarily followed its synoptic mesoscale model COSMO-DE2 forcing leading to a disagreement compared to the observations. This implies that the mesoscale model lacked capturing the regional-scale moisture and fine-scale variability in the city. Differences were smallest for open-spaced stations for both, observations and 5 m modeled simulated data (Fig. 8). If the coupling does not correctly account for the complex interaction between urban surfaces and the air it is reasonable

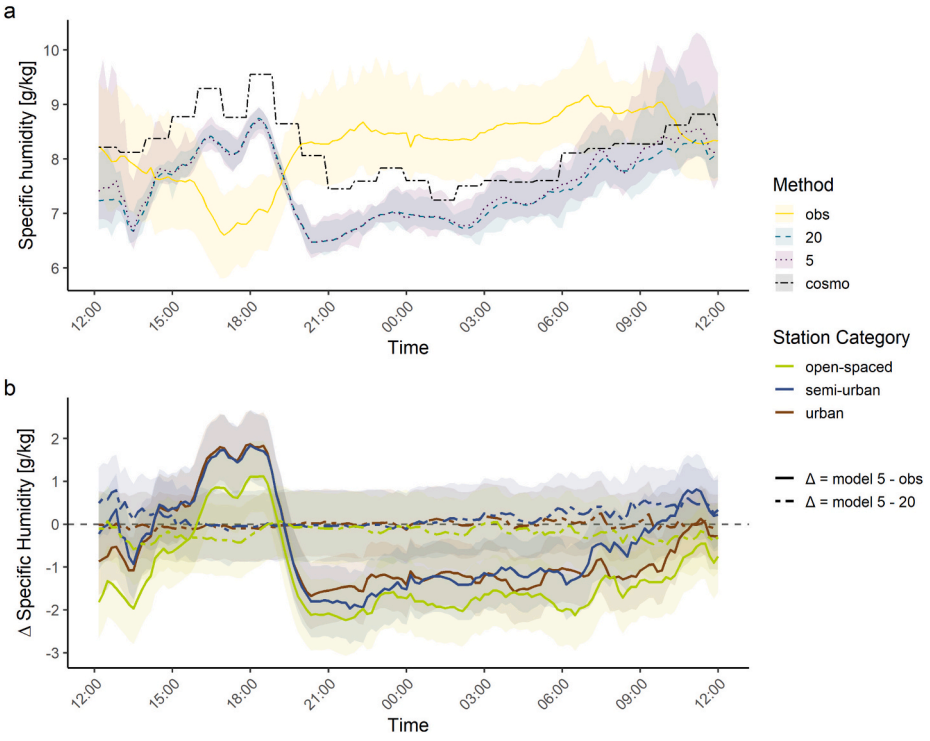


Fig. 8. Mean diurnal air temperature from July 25 to July 28, 2019, for: (a) Observed (obs) and simulated specific humidity at 5 m (5) and 20 m (20) resolution, along with the modeled synoptic forcing boundary condition from COSMO-DE2 (cosmo), (b) mean differences between simulated 5 m resolution and observations (5 m - obs) and between simulated 5 m and 20 m resolutions (5 m - 20 m) for the respective microclimatic subcategories derived from MRD analysis. Shaded areas in the respective colors represent the range of measured and modeled variability (minimum to maximum values) across all 14 stations over time, illustrating data spread. The x-axis represents time in CEST, while the y-axis represents (a) specific humidity and (b) specific humidity differences in g/kg.

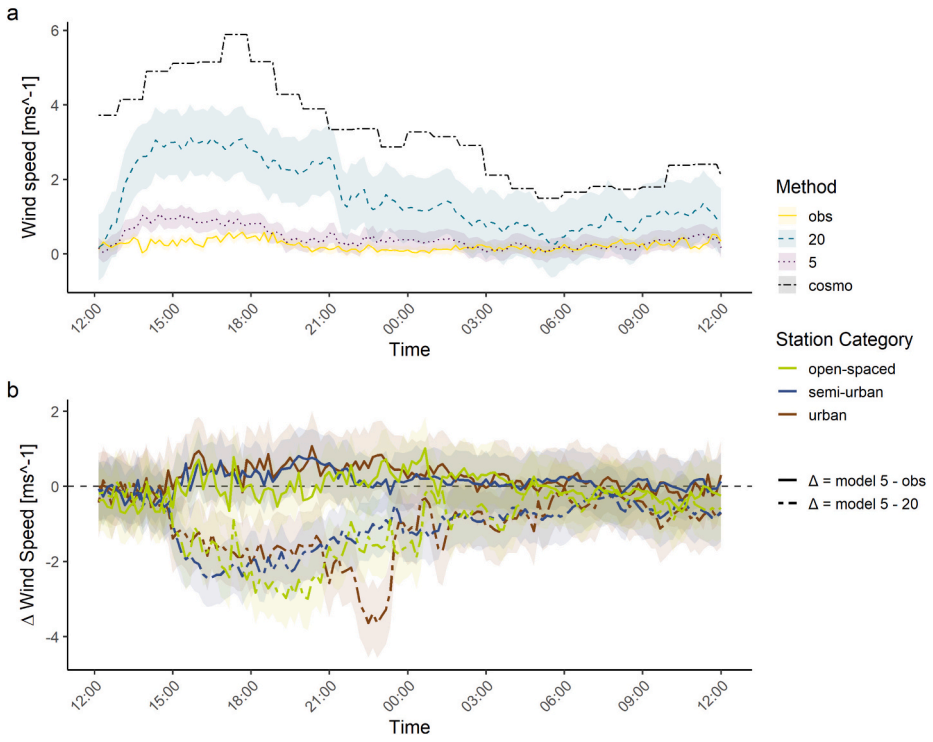


Fig. 9. Mean vectorial diurnal wind speed from July 25 to July 28, 2019, for: (a) Observed (obs) and simulated specific humidity at 5 m (5) and 20 m (20) resolution, along with the modeled synoptic forcing boundary condition from COSMO-DE2 (cosmo), (b) mean differences between simulated 5 m resolution and observations (5 m - obs) and between simulated 5 m and 20 m resolutions (5 m - 20 m) for the respective microclimatic subcategories derived from MRD analysis. Shaded areas in the respective colors represent the range of measured and modeled variability (standard deviation). The x-axis represents time in CEST, while the y-axis represents (a) the vectorial mean wind speed and (b) vectorial mean wind speed differences in $ms^{-}1$.

to lead to such inconsistencies in humidity fields. Our findings indicate no significant improvement in model-observation agreement for an increase from 20 m to 5 m resolution for simulated air temperature and specific humidity, while the agreement for the horizontal wind does improve(Fig. 7, Fig. 8, Fig. 9). Observed and modeled surface winds compare much better for the 5 m model resolution for all station categories Fig. 9. Accurate representation of building geometries is crucial for realistic wind flow modeling. An larger, well-resolved building layer around the target model area effectively captures the obstruction effects of buildings on wind flow leading to agreement between simulated and observed wind velocities. We thus attribute the improved model agreement to a well resolved urban flow field forming in the building body. Note that evaluated winds include the combined effects of speed and direction, which are difficult to capture and highly variable in both observations and simulation due to micro-scale changes in urban environment interacting with the flow.

3.2. Validating the LES model for spatiotemporal variability

We investigated the two-point correlation coefficient across all 14 stations against the separation distance between the measurement stations using the results from MRD analysis for observed values. As expected for structurally diverse urban geometries comprised of a quasi-randomized patchwork of different land cover in close proximity, no correlation between the stations' separation distances and the two-point correlation metric was found for any of the tested climate elements Fig. 10 (Mahrt et al., 2009). Notably, only one station representative of the urban core (station 2) showed a modest decreasing trend in correlation of specific humidity with increasing distance (Fig. 10 d). The classification Fig. 6 used in 2 based on land surface types in the immediate surrounding of a station, approximated by the depth of the IBL, already indicated a strong influence of microclimate on station-station correlation. An analysis of two-point correlation across all 14 stations with varying station groups revealed that the spatial patterns of the state variables were more strongly influenced by local microclimatic conditions Fig. 12, Fig. 15 than by separation distance. Thus, we expect the stations to primarily reflect their immediate environmental settings within the urban landscape. In result, the mrd analysis of observed and simulated data was classified into 3 station groups of 'urban', semi-urban' and 'open-spaced' stations.

The MRD results of both, observed and simulated data, reveal expected variations in two-point correlation analysis across the shortest (11.25 min) and longest (180 min) time scales (Fig. 11). Notably, there are distinct differences in the behavior of scalars and wind: both scalars exhibit a concave trend, with lower correlations at the shortest scales and higher correlations at the longest, while wind displays a convex trend with an increasing median correlation coefficient RR at larger time scales (Fig. 11). These varying time

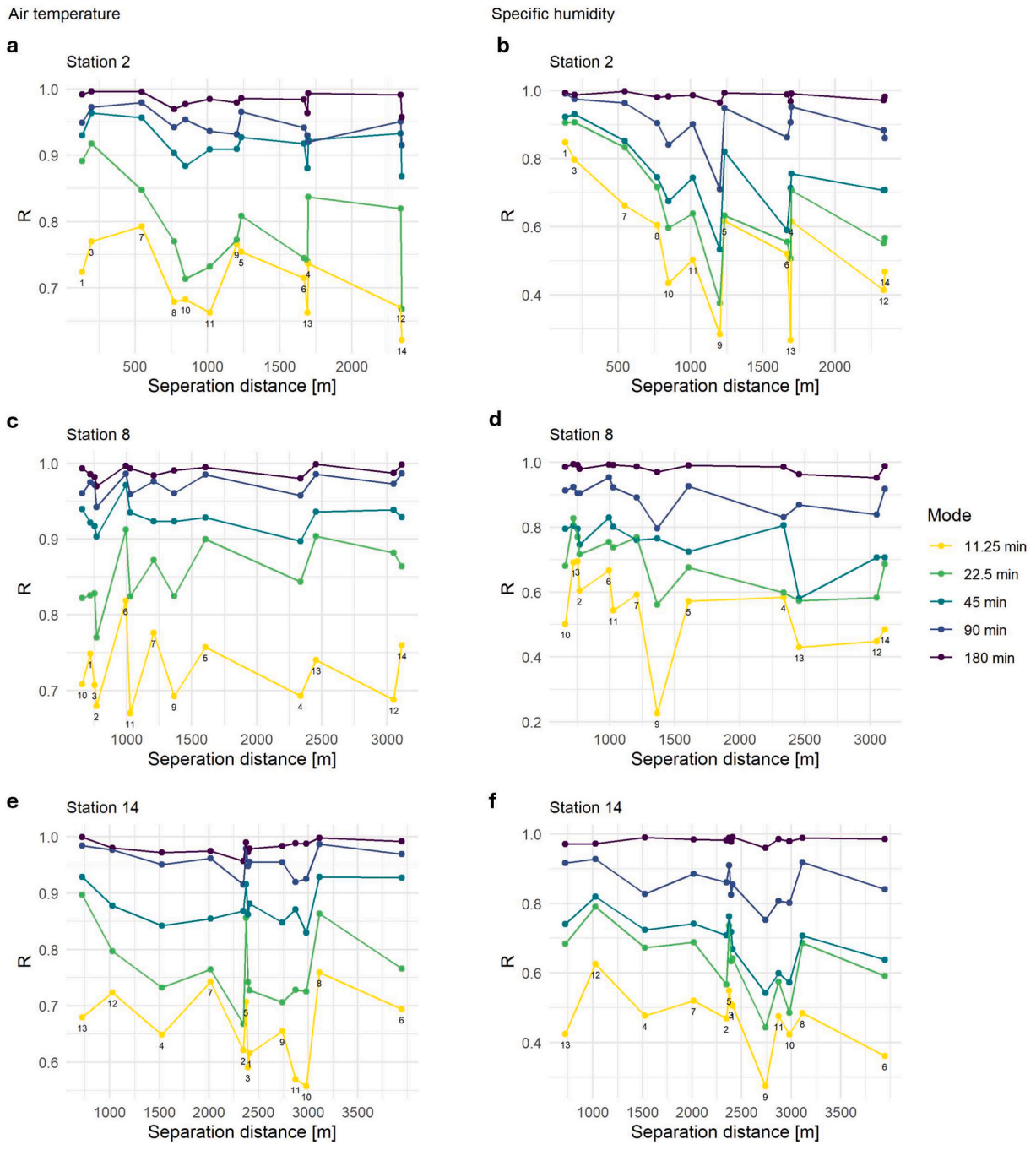


Fig. 10. Two-point correlation coefficient R for individual stations representatives of the station classes representing urban locations (a,b), semi-urban locations (c,d) open-spaced locations (e,f). The correlation is calculated between the station and its separation distance to all other stations of the measurement network. The separation distance is shown along the y-axis.

scales reflect distinct physical processes, with shorter time scales capturing turbulent transport and flow dynamics, and longer time scales representing mean flow characteristics, such as radiative forcing and advection. We infer a scale independence at the longest time scales, driven by the diurnal cycle of air temperature governed by solar heating during the day and radiative cooling at night. A correlation R near 1 indicates consistent diurnal oscillations in which all stations participate driven by the availability of solar radiation driven by the Earth's rotation (Thomas, 2011). Overall, temporal decorrelation across all stations was greater for wind than for scalars. Median R values for air temperature and humidity approached nearly perfect correlation ($R \approx 1$), while the median R coefficient for wind peaked at 0.6, indicating that wind exhibits greater inherent variability (Fig. 11). High variance in two-point

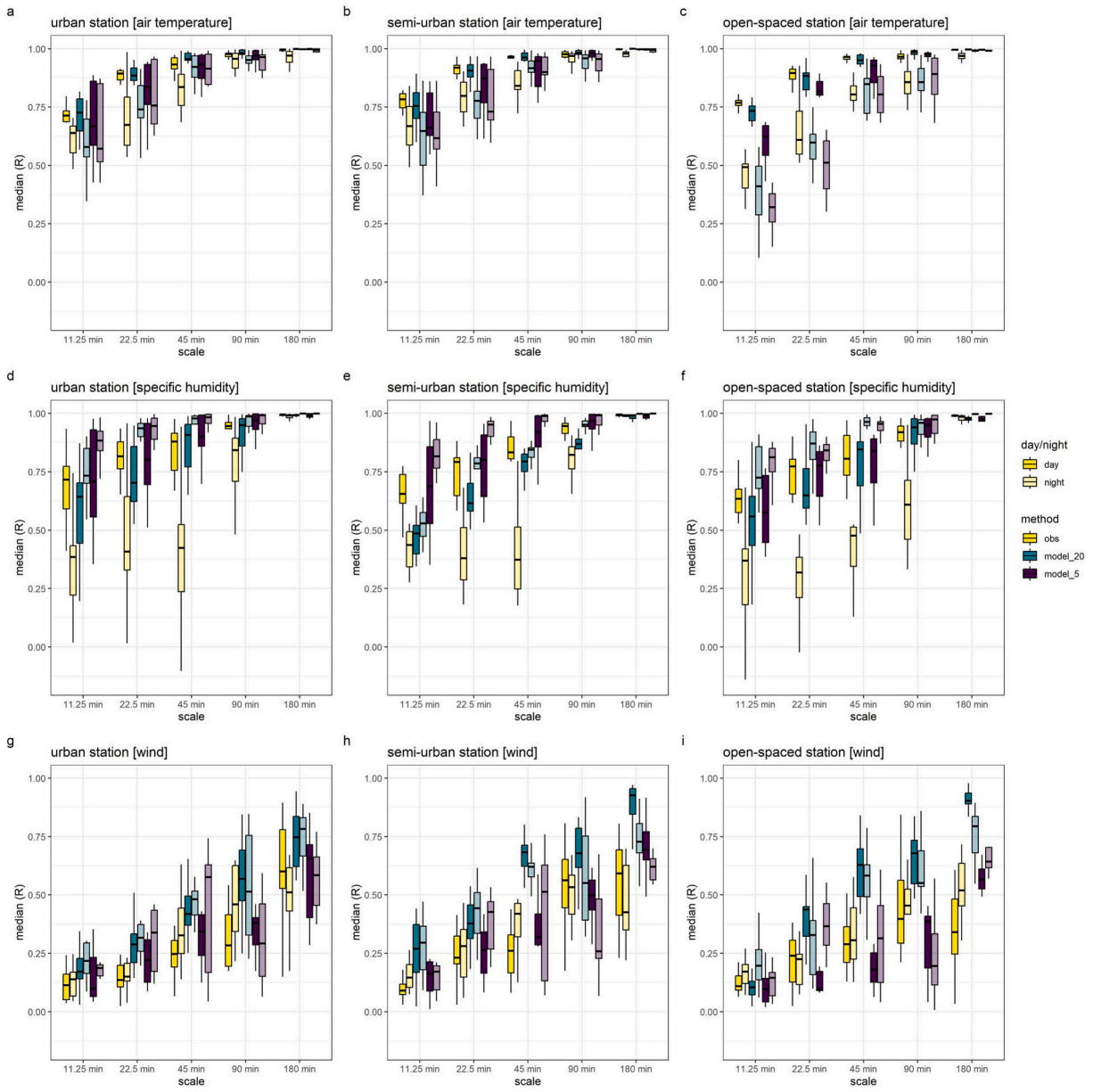


Fig. 11. Median of the two-point correlation validation of observation and simulation results including the 25 % and 75 % percentiles and whiskers. The columns represent the different station categories urban, semi-urban and open-spaced whereas the rows represent the climate element air temperature, specific humidity and vectorial mean wind. Every plot shows the median and interquartile range R for the 5 different temporal MRD scales. Results are colored after observations, 20 m simulation and 5 m simulation. Day and night are distinguished by the brightness of each color.

correlation among stations was especially prominent at short time scales for scalars and at large time scales for wind (Fig. 12), suggesting that these scales capture microclimatic variability in fluxes and airflow specific to each location. Strong diurnal differences in correlation and variance were observed for air temperature and specific humidity, but not for wind. Differences in the two-point correlation coefficient between day and night ranged from approximately 0.2 for air temperature to 0.4 for specific humidity (Fig. 12, Fig. 15), with greater day-to-night discrepancies at shorter MRD scales (11.25 min, 22.5 min) compared to longer scales (90 min, 180 min). Early Morning Transition (EMT) data closely resembled nighttime conditions, while Late Afternoon Transition (LAT) data aligned with daytime conditions, prompting us to combine EMT with nighttime and LAT with daytime for all subsequent analyses. Consequently, MRD results were interpreted across five time scales and grouped into urban, semi-urban, and open-space station clusters, with separate analyses conducted for summarized daytime and nighttime conditions.

In the following, we will assess how accurately the simulated values from the MRD analysis reflect the observed data. Results showed spatiotemporal agreement between modeled and observed data for air temperature, wind and daytime specific humidity.

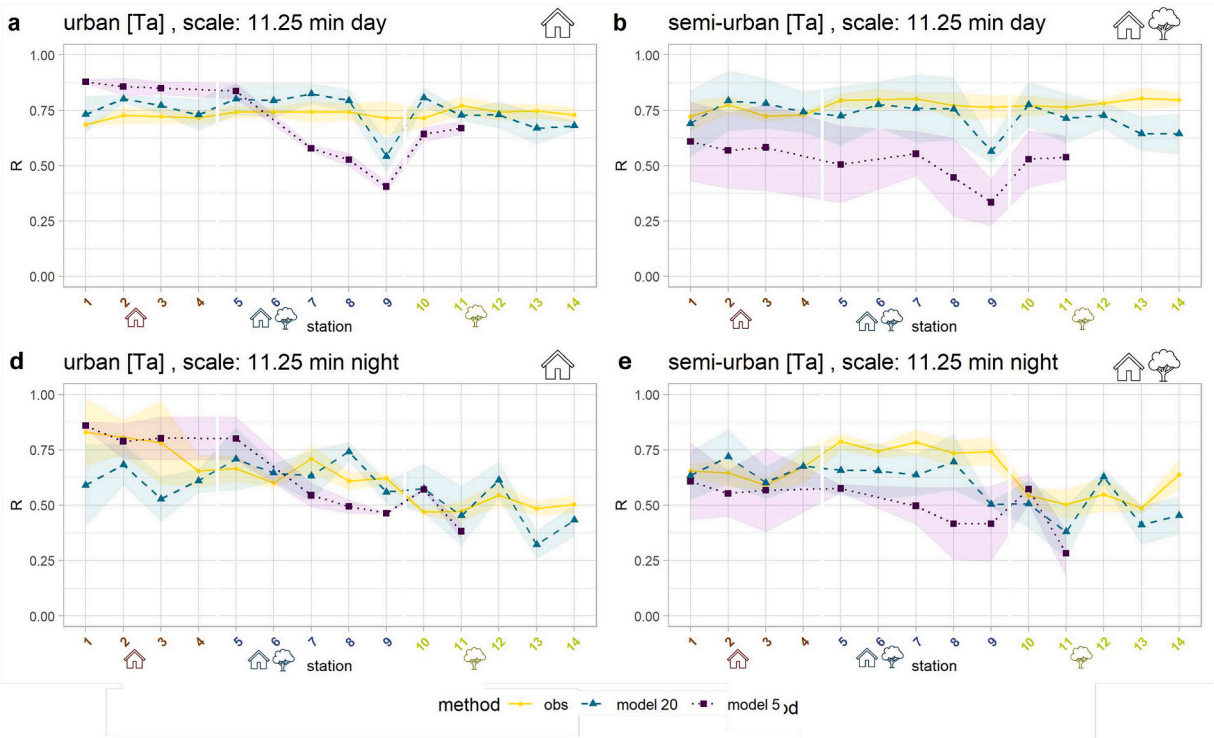


Fig. 12. Spatial two-point correlation coefficient from the MRD analysis of observation and simulation results for air temperature. Urban and semi-urban stations are compared at 11.25 min MRD scale. Correlation was computed in a and d for the mean of all urban stations and in b and e for the mean of all semi-urban stations. The stations on the x-axis were ordered numerically in the same sequence across each plot. Each color and colored pictogram depicts stations of the same classification. A and b depict the day time whereas d and e show night time correlations. Point symbols as well as colors depict the different methods of observation, 20 m and 5 m simulation. The envelope shows the variability between the station of the lowest and highest correlation within the station group.

Although air temperature displayed different spatial and temporal behavior during day and night, the model represented both conditions (Fig. 11) well. Air temperature differences in day and night capture the mixing processes of the urban roughness layer during daytime and the establishment of a weakly stable boundary layer during nighttime. Differences between stations become particularly enhanced during the nighttime when the longwave radiative transfer is tightly coupled to the heat diffusivity of the surface, which is defined as the ratio of thermal conductivity to heat capacity. Our findings for air temperature show agreement between stations of the same microclimate cluster for the night, but not during the day (see Fig. 12). During daytime, correlation among all stations was similar (see Fig. 12). This can be explained by the diurnal radiative and turbulent mixing processes. Spatial temperature differences are expected to become larger at nightime, the UHI intensifies before sunset (7 pm) and peaks around midnight. The establishment of the weakly stable nocturnal boundary layer and additionally, the differential surface thermal emission above the fine alternating urban surface structure do play an important role for the enhancement of night time differences. During daytime solar radiative fluxes produce local sources of heating or cooling within the UBL and greatly influence its turbulent structure and dynamics. The enhanced daytime turbulent kinetic energy (TKE) causes air temperature differences to be minimized through strong mixing. Largest differences between day and night for both, observed and simulated values, are observed for the stations located within the open-spaced cluster (Fig. 11). For specific humidity, the modeled results showed strong correlation with observed data during daytime only (see Fig. 15, Fig. 11d-f). At night, however, the model displayed pronounced decorrelation from measured conditions, consistent across all stations in the network (Fig. 15, Fig. 11). Fig. 11, shows decorrelation up to 0.3 median R. Fig. 16 revealed decorrelation of specific humidity occurring throughout the first half of the diurnal cycle and across the magnitude of values. These findings suggest that the modeled generation and distribution of humidity are largely disconnected from observed nighttime moisture field behavior. Several factors may account for this discrepancy. During daytime, specific humidity is primarily driven by evapotranspiration, which is governed by solar radiation, water availability, and turbulent mixing that disperses scalars vertically. Under extreme heat, evaporation rates are highly sensitive to surface characteristics and soil water availability, leading to spatial heterogeneity in humidity distribution across the urban boundary layer (UBL). At night, surface cooling brings additional processes like condensation and dew formation into play, though these are minimal during intense summer heat. Instead, the development of a weakly stable boundary layer at night becomes crucial, as it limits vertical mixing and often increases near-surface specific humidity. In some cases, nocturnal jets can also transport moisture horizontally, further affecting nighttime humidity distributions (Seo and Dirmeyer, 2022). Moreover, horizontal variability in specific humidity within the UBL arises largely from non-uniform surface properties, such as variations in soil and surface moisture content and

latent heat flux. For example, stagnant water bodies retain enough heat capacity to continue emitting fluxes into the atmosphere, while other surfaces see reduced humidity differences as fluxes become minimal or directed towards the surface at night. A likely source of decorrelation in the model, however, is the reliance on coarse-resolution COSMO-DE2 data, which may not accurately capture these local variations in surface moisture or soil water content. As shown in Fig. 16, the model's specific humidity closely follows synoptic input, so any discrepancies in these inputs could alter the modeled evaporation and transpiration processes, leading to further decorrelation.

For horizontal winds, a high correlation within stations of the same cluster was absent. It was neither observed for day, nor for nighttime. Varying wind directions are common for urban environments as they feature complex geometry at high heterogeneity. At some places, the airflow moves in a direction opposite or different from that of the prevailing wind caused by the building and other flow obstacles leading to systematic deflection of the flow. Observed and simulated results indicated fairly low average wind speeds near the ground, with maximum wind speeds of 1.7 ms^{-1}. We assume that the high variability in wind direction at very low wind speeds both caused by the urban roughness layer combine such stark heterogeneity that no correlation within stations of the same land surface type cluster could be observed. High variability in Fig. 13 b, c e and f of observed results and large variance in Fig. 11 g, h and i for the larger time scales show the high variability of the wind field in the lower urban boundary layer.

The four-fold increase in simulated resolution from 20 m to 5 m did not improve the spatiotemporal agreement between model and

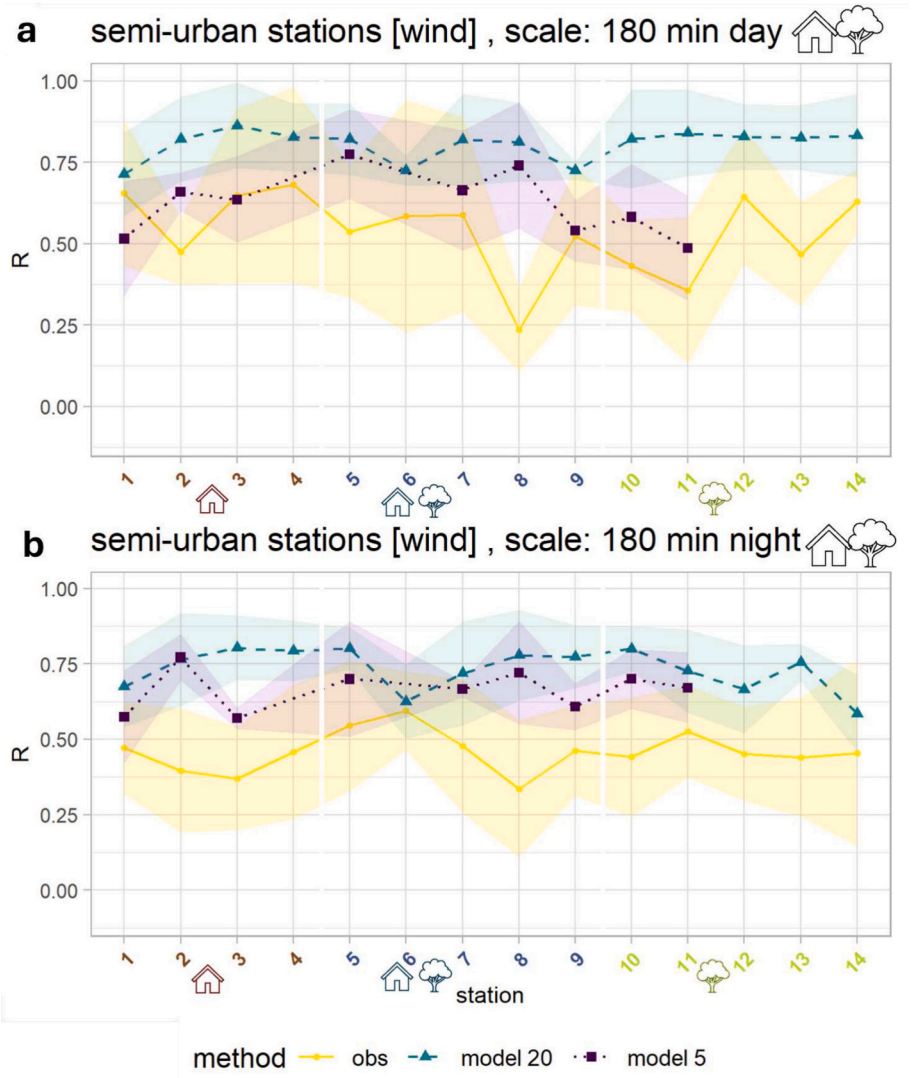


Fig. 13. Spatial two-point correlation of MRD analysis of observation and simulation results for vectorial mean wind among semi-urban stations at 180 min MRD scale. Correlation was computed in a and d for the mean of all semi-urban stations. Stations along the x-axis were sorted numerically classified to urban, semi-urban and open-spaced stations. A depicts the day time whereas b shows nighttime correlations. Point symbols as well as colors depict the different methods of observation, 20 m and 5 m simulation. The envelope shows the variability between the station of the lowest and highest correlation within the station group.

observation for air temperature. Instead, in particular the semi-urban and open-spaced stations in our analysis revealed a better agreement for air temperature using 20 m resolution (Fig. 11). The four-fold increase in simulated resolution for specific humidity showed improvement in spatiotemporal agreement in particular for the semi-urban and open-spaced station cluster. Little improvement can be seen for the urban station cluster (see Fig. 11). Please note that this finding only includes daytime specific humidity only, since nighttime specific humidity showed no correlation for both resolutions. We attribute little improvement in agreement between simulation and observation for increasing grid resolution to uncertainties associated with the land surface model. We further discuss expected underlying mechanisms that lead to such lack in agreement in the subsequent section 3.3. For both of the scalars, modeled 5 m spatiotemporal variability is larger than observed. Air temperature differences in variability between observed and modeled values reach 0.2R, expressed by the large interquartile range (IQR), are shown. Large variability indicated by a an IQR of 0.3R for specific humidity modeled short time scales for all three station cluster is found in contrary to variability of 0.1 R IQR for observed values. In contrast to the scalars, the simulated space and time precision of the wind has dramatically increased for the four-fold increase in resolution. Both, median and averaged correlation support this finding for all three station cluster (Fig. 13, Fig. 11). Fig. 11 shows an overestimated wind of the 20 m simulation for both, day and night. The increase in correlation within stations across larger time scales comes with higher variance per time scale. This suggests location-specific wind microclimate at the larger time scales capturing the turbulence of flow rather than the mean flow. The wind microclimate is likely to depend on the presence of individual flow obstacles such as buildings and local variation of topography. A systematic deviation in spatial variability is not observed (Fig. 11, Fig. 13).

The spatial variances depicted in Fig. 17 are defined as the temporal mean deviation from the spatial network mean on logarithmic scale. In spatiotemporal terms it depicts a deviation of the corresponding microclimate of the state variables. The spatial variability of the flow was computed for each station i as its deviation from the spatial network average marked as,

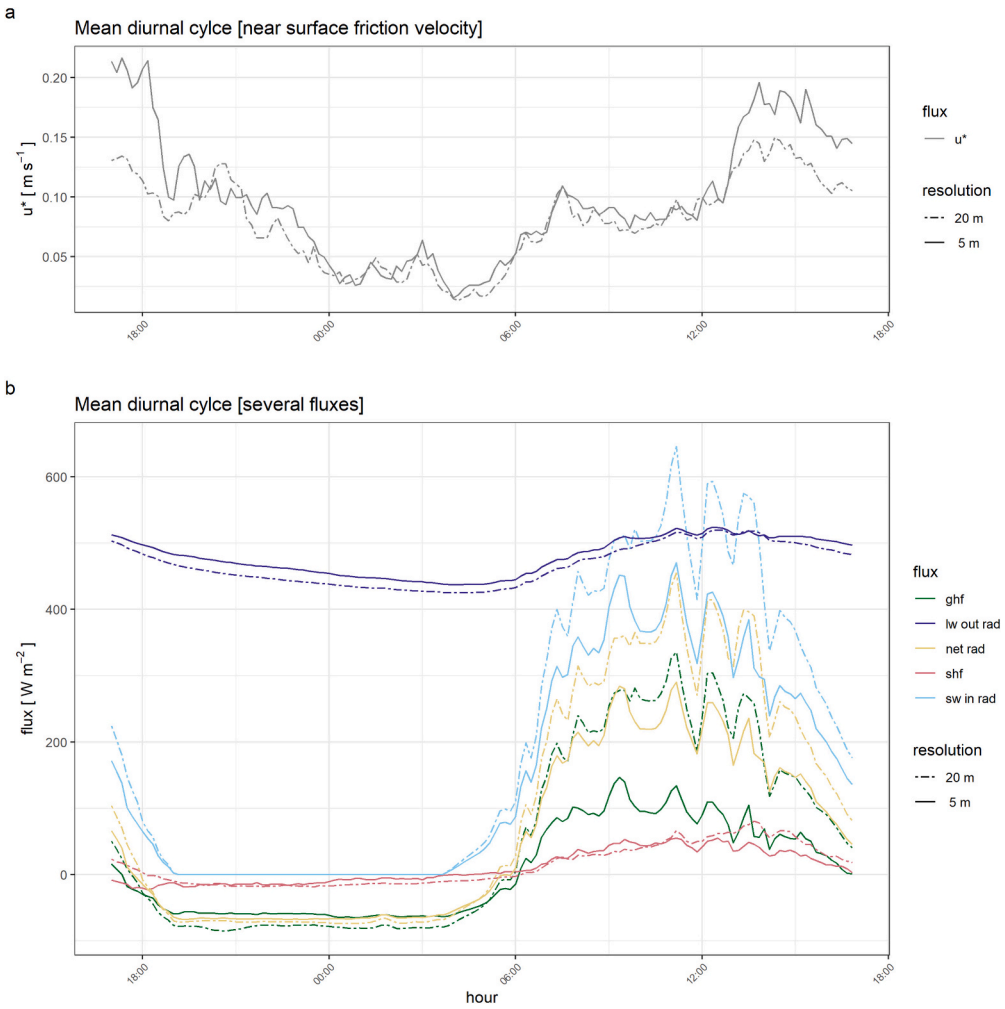


Fig. 14. Mean diurnal cycle of the modeled near surface friction velocity defined as the density-normalized momentum flux (u_* in ms^{-1}), ground heat flux (ghf, in Wm^{-2}), sensible heat flux (shf in Wm^{-2}), net radiation (net rad in Wm^{-2}), shortwave incoming radiation (sw in rad in Wm^{-2}) and longwave outgoing radiation (lw out rad in Wm^{-2}). The diurnal mean was computed between two consecutive days 2019-07-25 17:00–2019-07-27 17:00. A mean between all virtual stations was computed. Solid and dashed lines present values at the 20 m and 5 m spatial resolution, respectively.

$$\widetilde{\phi}_i = \phi_i - \langle \phi \rangle. \quad (2)$$

shows that air temperature exhibits the greatest deviation at larger scales, with absolute deviations exceeding those of specific humidity or wind. This suggests that motions at the largest time scales exhibit the strongest dependence on spatial scale, with only minor differences observed between day and night. The urban station cluster demonstrates the lowest spatial deviation, while the open-spaced station cluster displays the highest spatial variability. The spatial variances of the modeled air temperature are decorrelated with the observed values, showing reduced deviation, particularly at larger time scales. Deviations for specific humidity are generally lower than those for air temperature, and exhibit a scale dependency, with specific humidity being influenced by longwave radiation at larger time scales (Fig. 17 b). The most notable discrepancies between modeled and observed specific humidity variances occur at night, as previously discussed in 3. In contrast, wind spatial variability is independent of time scale, indicating that the transport of momentum is captured uniformly across all time scales (Fig. 17 c). However, momentum is weaker at nighttime.

3.3. Mechanisms explaining differences between model and observations across resolutions

This section delves into key findings reported in the preceding section and highlights the mechanisms that may explain unexpected deviations between model outcomes and observations. We focus on three main findings. First, while the two consecutive diurnal cycles tested showed good agreement for daytime air temperatures, the model overestimated nighttime temperatures by far and thus exaggerates the UHI effect. However, the space-time representation of air temperature variations across the domain remained robust. Second, the space-time validation indicated strong agreement between the model and observations across all variables, except for nighttime specific humidity, where notable discrepancies emerged compared to the well-represented daytime dynamics of specific humidity. Finally, contrary to our expectations, the enhanced model resolution of 20 m versus 5 m had a small impact on the accuracy of scalar variable simulations, but yielded notable improvement for winds. All key findings are relevant to assess the utility of the model to support policy making and the decision to curb and mitigate urban heat stress for humans.

3.3.1. Insufficient air temperature cooling during night time

A comparison of the COSMO-DE2 mesoscale driving model with our simulations shows higher simulated nocturnal temperatures than the larger-scale forcing. This suggests that the nighttime cooling discrepancies are likely due to mechanisms embedded in the PALM model. The ground heat fluxes at the stations were strikingly high. Urban areas are characterized by large ground heat fluxes. However, the ground heat flux exceeded the sensible heat flux by up to over 130 % which appears implausible. Such high ground heat flux means that during the day, the majority of the available energy was transported into the ground. During the night, the energy deficit at the surface generated by net radiative cooling was mostly compensated by the ground heat flux. Therefore, the nocturnal sensible heat flux, which transports heat from the near-surface air to the surface, was smaller due to the smaller temperature difference. An explanation for the overestimation of the ground heat flux of the model and a subsequent underestimation of the sensible heat flux could be an overestimated thermal diffusivity of the ground. Thermal diffusivity describes how easily heat propagates through the ground (Oke et al., 2017). An indication of an overestimated thermal diffusivity was the high responsiveness of the ground heat flux to the net energy flux, shown by the close mirroring of both fluxes, as seen in Fig. 14. However, the thermal diffusivity values of different types of surface cover used in the PALM model were in agreement with the values found in the common literature or were even lower. Despite the general agreement of the thermal diffusivity values of the model and the literature values, the high responsiveness of the ground heat flux hints at an overestimation of the thermal diffusivity nevertheless. The findings of insufficient cooling during the night are reflected in previous studies (Gehrke et al., 2020; Salim et al., 2020; Resler et al., 2021), indicating a link to the radiative module. The choice of radiation model is tested to be significant at night (Gehrke et al., 2020). Our simulation used external radiation prescribed by measured global, that is, shortwave downwelling radiation in combination with the clear sky module (Gehrke et al., 2020) report similar experiences for the clear sky module being accurate during the day but overestimating net radiation during night. A known problem among coupled approaches is the limitation in estimating atmospheric longwave radiation cooling of near-surface air during night (Maronga et al., 2019). Also Gehrke et al. (2020)) and Resler et al. (2021) reported an underestimated outgoing longwave radiation flux in relation to observations due to prescribed incoming radiation. A high value of received longwave radiative flux even at night as contribution to the surface radiation budget is considerable, and for some surfaces, especially those located in daytime shade, it is the main source of radiation. The reception of longwave irradiance from urban surfaces constitutes a major part of the longwave radiation budget received (Salim et al., 2020). Longwave radiative cooling is an important contributor to the temperature budget during calm nights (Steeneveld et al., 2010), which is also the case for our study. Surface net radiation studies of the radiation module of PALM have shown high values for received nighttime radiation flux for surfaces in densely built areas (Salim et al., 2020) and underestimated longwave radiative cooling, especially when combined with prescribed incoming radiation (Gehrke et al., 2020). An additional potential explanation for the increased modeled air temperatures is too much turbulent mixing under weak-wind nocturnal conditions at night. An amplified mixing of the air's moisture and temperature properties causes lower vertical and horizontal temperature and moisture gradients as the enthalpy is mixed downward into the urban canopy layer. Overestimated TKE is a common issue in the PALM-4 U model (Gronemeier et al., 2021). In the model domain obstacles are represented on a rectilinear grid turning smooth vertical and horizontal surfaces into sharp edges, increasing surface roughness which causes higher turbulence Fig. 14.

3.3.2. Decorrelation of specific humidity space time behavior during night time

We observed a significant decorrelation in the nighttime spatiotemporal humidity between the simulations of both 5 m and 20 m

resolutions and the observations. Specific humidity and water vapor mixing ratio are height-conserved meteorological elements invariant to changes in pressure or temperature. The humidity of an Eulerian air parcel can change through advection or turbulence flux, which add or subtract water vapor to or from the parcel. Evaporation, the primary transport mechanism for moistening the air, is typically larger and more pronounced during the day because of solar radiation that delivers the enthalpy necessary for the phase change. The high surface heterogeneity in urban areas leads to significant variability in evaporation and, consequently, spatially variable urban humidity, especially during daytime. While spatiotemporal differences within the urban boundary layer were well captured during the day, which indicates a good representation of turbulence and daytime evaporation, there were noticable spatial differences in specific humidity among urban, semi-urban and open-spaced stations during nighttime. This is shown by differences in median R values of approximately 0.3 across all stations (see Fig. 11 d-f) and also shown by Fig. 15, suggesting that the model struggled to reproduce the local surface conditions and transport processes. Strikingly, also absolute values for specific humidity compared poorly and closely followed the coupled synoptic COSMO-DE2 data rather than being driven by local sinks and sources. Weak nighttime evapotranspiration and dewfall are difficult to capture correctly in the model which are extremely sensitive to simulated air and surface temperatures and thus radiative forcing. However, in contrast to the poor representation of simulated absolute data distribution of nocturnal specific humidity, the spatiotemporal daytime differences were well resolved in the model when compared to

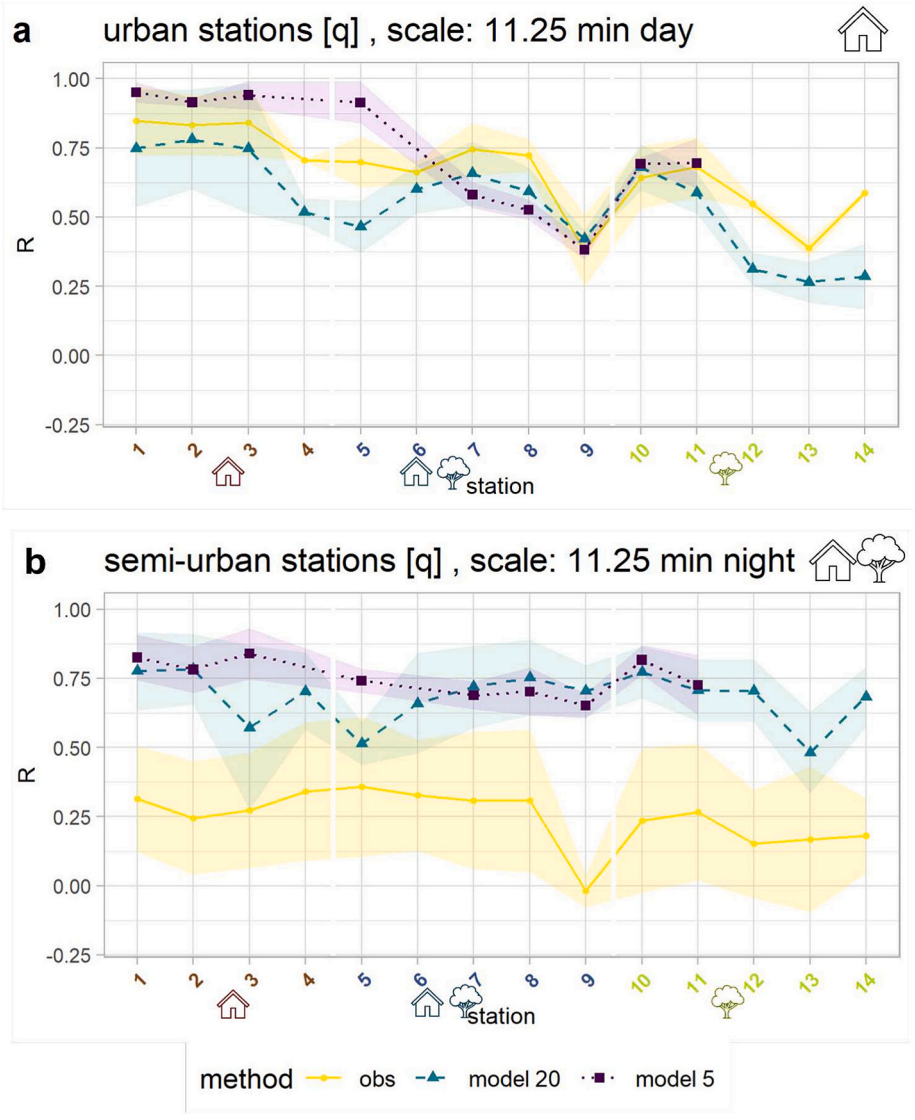


Fig. 15. Spatial two point correlation of MRD analysis of observation and simulation results for specific humidity among urban (a) and semi-urban (b) stations at 11.25 min MRD scale. Correlation was computed in a for the mean of all urban stations and in b for the mean of all semi-urban stations. The x-axis display all network stations sorted numerically and color and symbol-coded in urban, semi-urban and open-spaced stations. A depicts the day time whereas b shows nighttime correlations. Point symbols as well as colors depict the different methods of observation, 20 m and 5 m simulation. The envelope shows the variability between the station of the lowest and highest correlation within the station group.

the observed data. One driving force for large-scale humidity is horizontal moisture advection as the transport of humidity by the mean airflow. An increase in humidity of an air parcel occurs if more moisture enters the parcel via advection than leaves the parcel. Judging from Fig. 16, specific humidity increased over the simulated 72 h period indicating a positive advection, while we did not investigate this result further. We conclude that the strong surface heating by the absorbed solar radiation leading to differential heating resulting in enhanced free convection through buoyancy and increased wind speeds leading to stronger forced convection both contributing to enhanced TKE lead to eliminating to the moisture differences in the model. In contrast, the insufficiently captured and inherently difficult to model nighttime transport processes for moisture lead to poor model-observation agreement and the found significant decorrelations.

3.3.3. Small differences in model-observation agreement across model resolutions

Our data analysis revealed an insignificant improvement for the space-time agreement between simulated data at 20 m and 5 m resolution and the observations for air temperature across all stations. When examining the variances, the simulated variances of air temperature at 5 m and 20 m exhibited minimal differences, yet they showed significant variation compared to observations. In contrast, the variances for specific humidity remained consistent across different resolutions and model and observation Fig. 17. This finding is consistent with both the absolute and the spatiotemporal validation methods. In contrast, wind data demonstrated significant improvements with increased resolution. Note that neither the measurements nor the simulations were intentionally calibrated or nudged to enhance agreement prior to validation. Instead, we conducted an independent comparison, limiting data exchange to essential boundary condition information before executing the model simulations. We attribute the different sensitivity to resolution in processes steered by the land surface module. The LSM has a greater influence on temperature and humidity than turbulent exchange. We selected the two resolutions to strike a balance between coarser resolution suitable for simulating larger domains and finer resolution required for street canyon analysis of inner city areas. The choice of resolution significantly influences various processes within the model, including:

- differences in the x, y, and z coordinates and thus location between measurement stations and modeled output data
- the accuracy of land cover heterogeneity and type in the vicinity of the virtual stations
- the omission of overlapping objects in favor of higher-priority features

As model resolution increases, improving the accuracy of preprocessed data becomes essential. Given these considerations, we conclude that the observed discrepancies stem primarily from the LSM in PALM-4 U, which is sensitive to issues related to coordinate differences, land cover type representation, and object prioritization. A key trade-off involves the representation of vegetation within the simulated environment, which is entirely parameterized for both resolutions. While the plant canopy module in PALM-4 U allows for the inclusion of trees, vegetation is represented only as fully covered pixels. This choice results in unrealistically large tree widths and sizes, rather than more accurate representations. Such simplifications can influence critical mechanisms including shading, transpiration, evaporation, infiltration, which collectively modify the water balance for each pixel. Additionally, there are other surface parameters which we expect to contribute to the observed discrepancies outside of those directly linked to vegetation, including bulk parameter settings in the model's equations, such as orders of magnitude for heat conductivity, heat capacity, albedo, water temperature, roughness lengths for various surfaces, and diffusivity, among others. Based on the analysis above, we attribute the discrepancies in temperature and humidity to their regulation by the LSM, whereas wind speed is better captured due to its detachment

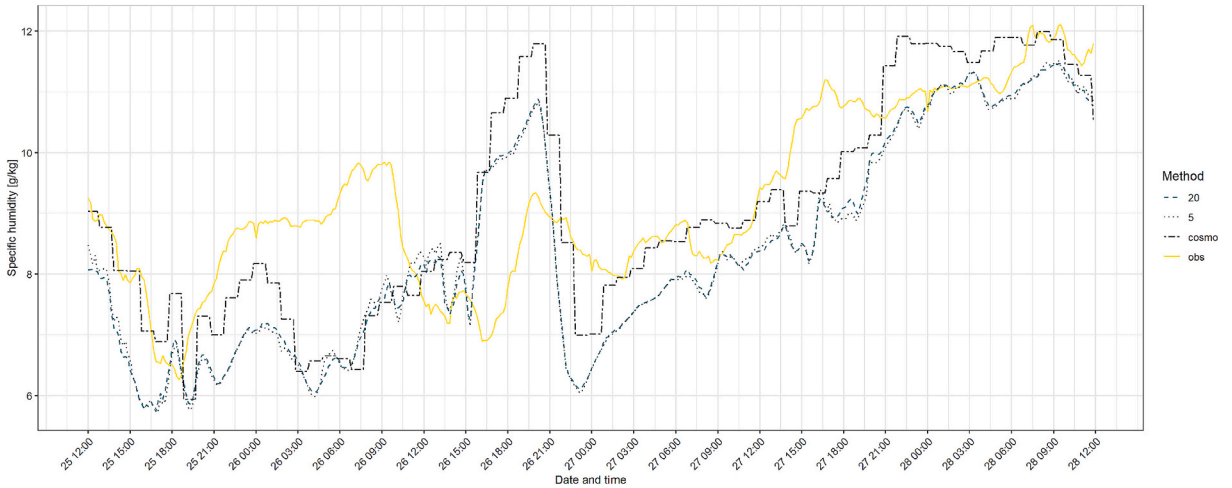


Fig. 16. Specific humidity spatial mean among all stationary locations for measured (mean obs) and simulated (mean 20 and mean 5) data in respectively 20 m and 5 m resolution. As a reference, the output of the mesoscale model COSMO-DE2 (cosmo) is included that serves synoptic forcing data as input into the coupled LES. The time series is taken from 2019 to 07-25 12:00 CEST until 2019-07-28 12:00 CEST.

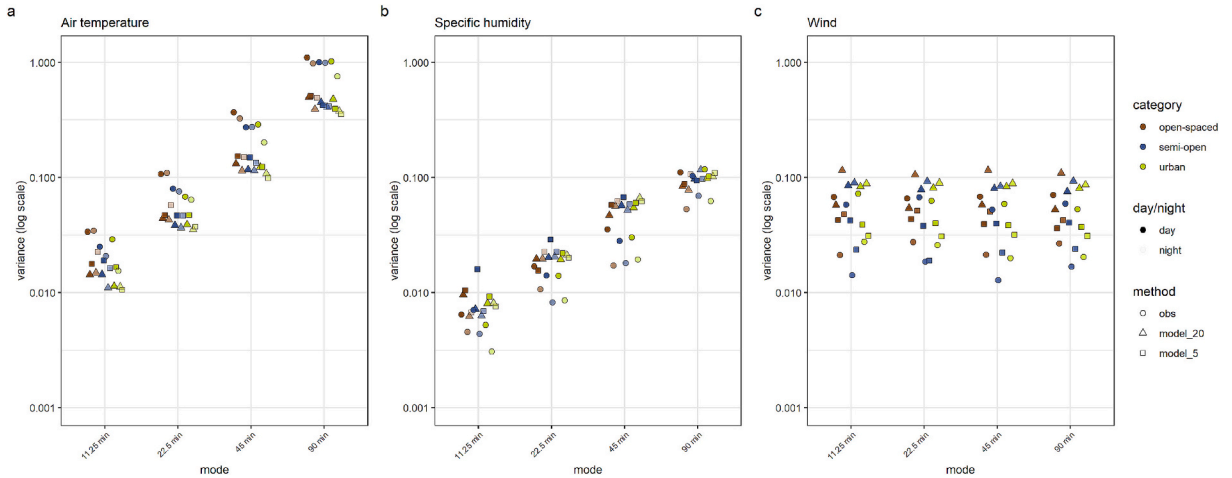


Fig. 17. Logarithmic variances from the MRD analysis of observation and simulation results of air temperature (a), specific humidity (b) and wind (c). The variances are grouped along the y-axis for the MRD scales from 11.25 min to 90 min. Differences in color transparency depict day or nighttime values.

from surface-driven processes. Two possible mechanisms within the LSM may contribute to these deviations: (1) an exaggerated differentiation of surface properties, leading to inconsistencies with actual conditions at observational stations, and (2) excessive coupling between TKE and the surface layer, which may enhance mixing and obscure finer-scale variations in surface properties.

4. Conclusions

Building-resolved modeling technologies are increasingly recognized as vital tools for policymakers in urban planning and climate adaptation, facilitating evidence-based decision making for sustainable development. Policymakers are leveraging these insights to design green infrastructure, optimize energy efficiency, and mitigate the effects of heat islands, thus aiming to improve urban resilience. This study underscores the complexity of urban microscale modeling and the importance of careful validation of modeling tools. We emphasize the importance of space-time variability alongside traditional evaluations of absolute differences between modeled outputs and observational data. Our approach used a densely deployed measurement network comprised of 14 stations, strategically positioned in structurally diverse urban environments.

A preliminary estimation of the spatiotemporal distribution of microclimates can be effectively derived through a detailed phenomenological approach based on land surface classification. This method is particularly valuable for policy makers who may lack the resources or expertise to conduct more complex analyzes. We show in our validation approach that differences in the agreement in absolute values and agreement in space-time patterns are independent of each other. In the following, we provide answers to the research questions posed in 1:

1. Difference in space-time correlation:

- Air Temperature: The differences between model simulations and observations are negligible in urban and semi-urban environments, regardless of the time of day. However, for open-space locations, the difference between modeled and observed median correlation coefficient deteriorates to 0.125 during both day and night for the smallest mrd scale, indicating a small discrepancy.
- Specific Humidity: While daytime differences remain negligible, nighttime comparison exceeds a median correlation coefficient difference of 0.3, suggesting significant deviations between modeled and observed specific humidity during nocturnal periods.
- Wind Speed: Across all station categories, MRD scales for both day and nighttime conditions, the difference between modeled and observed median correlation coefficient deteriorates to 0.1 and is thus negligible.

We thus conclude that the differences in space-time correlation between observations and modeling are negligible except for specific humidity at night.

2. A four-fold increase in model resolution leads to mixed improvements in model-observation match:

- Air Temperature: No significant improvements are observed in either absolute values or space-time correlation, suggesting that factors beyond resolution, such as land surface parameterization schemes or urban surface representation may have a stronger impact on air temperature.
- Specific Humidity: While improvements are observed in the representation of spatiotemporal variability during daytime, the increase in resolution does not eliminate the found nighttime discrepancies.

- Wind Speed: higher resolution improves both the representation of spatiotemporal patterns and absolute values, leading to a closer correlation match with observations mostly reducing the difference in correlation coefficient to less than 0.1 R.

These findings suggest that increasing model resolution improves representation of the wind and daytime moisture fields in the LES model, it has a small effect on other variables as they are more sensitive to the land-surface scheme.

3. Deviations between model simulations and observations occur primarily for scalar variables (air temperature and moisture), and not the airflow:

- Underestimated Longwave Radiative Flux: Deviations in absolute value comparison of air temperature can largely be attributed to an underestimation of longwave radiative flux within the radiative module. This contributes to the bias in nighttime air temperature.
- Horizontal Moisture Advection: The observed discrepancies in nighttime specific humidity, both in absolute values and space-time correlation, suggest insufficient representation of horizontal moisture advection.
- Influence of Land Surface Parameterization: A possible mechanism explaining the deviations, particularly for scalar variables, could be a limitation by model approaches to resolve fine-scale atmospheric variability. Given that scalar variables are highly sensitive to land surface properties and the parameterizations connecting it to the airflow, the discrepancies are closely tied to the performance of the land surface module. Improving it for an urban landscape with drastically different radiative and heat and moisture conductive properties this aspect remains a challenge. Land surface heterogeneity plays a significant role in determining scalar variable distributions and transport.

While an increase in model resolution from 20 m to 5 m has demonstrated improvements in space-time behavior, achieving a resolution finer than 5 m is expected to further enhance model accuracy. However, a 1 m resolution entails substantial computational costs and is only recommended when addressing highly localized adaptation measures.

Urban models that can accurately interpret spatial and temporal behavior are valuable tools for designing mitigation strategies against urban heat. LES simulation results should be carefully selected; scenarios where such models prove particularly useful include the identification of space-time differences including urban heat hot spots for targeted green infrastructure interventions, simulating the cooling effects of parks and vegetated areas, and assessing the impact of increased ventilation corridors on heat dissipation. For simulations aimed at forecasting heat exposure or assessing absolute cooling potential, we recommend using very fine resolutions of approximately 1 m. This level of detail is essential for accurately capturing small-scale processes within a restricted area, ensuring more precise and reliable model outcomes.

CRediT authorship contribution statement

Leyla Sungur: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Wolfgang Babel:** Supervision, Software. **Eva Späte:** Visualization, Validation, Data curation. **Johann Schneider:** Investigation, Data curation. **Christoph K. Thomas:** Writing – review & editing, Supervision, Software, Resources, Project administration, Funding acquisition, Conceptualization.

Funding

Funding: This study was conducted in the framework of the MiSKOR project (reference number K3-8503-PN 18-17) as part of the network initiative “Climate Change and Health” (Verbundprojekt Klimawandel und Gesundheit, https://www.vkg.bayern.de) funded by the Bavarian State Ministry of the Environment and Consumer Protection and the Bavarian State Ministry of Health, Care and Prevention. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre (www.lrz.de). **Funded by the Open Access Publishing Fund of the University of Bayreuth.**

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Christoph K. Thomas reports financial support was provided by Bavarian State Ministry of the Environment and Consumer Protection. Christoph K. Thomas reports financial support was provided by Bavarian State Ministry of Health and Care. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Andreas Tschuschke and Isabel Spieß for research data installation and collection. Furthermore we would like to thank Julian Zeidler for cooperation in geodata preparation specifically designed for PALM-4 U. The authors gratefully acknowledge the support from the PALM group of Leibniz University Hannover and the technical support from the Leibniz Supercomputing Centre (www.lrz.de). Funded by the Open Access Publishing Fund of the University of Bayreuth.

Data availability

Data will be made available on request.

References

Bissolli, P., Deutschländer, T., Imbery, F., Haeseler, S., Lefebvre, C., Blahak, J., Fleckenstein, R., Breyer, J., Rocek, M., Kreienkamp, F., Rösner, S., Schreiber, K.-J., 2019. Hitzewelle juli 2019 in westeuropa – neuer nationaler rekord in deutschland. techreport. Deutscher Wetterdienst.

Brozovsky, J., Gustavsen, A., Gaitani, N., 2021. Zero emission neighbourhoods and positive energy districts – a state-of-the-art review. Sustain. Cities Soc. 72. https://doi.org/10.1016/j.scs.2021.103013, 103013. ISSN 2210-6707. https://www.sciencedirect.com/science/article/pii/S2210670721002973.

Buchin, O., Hoelscher, M.-T., Meier, F., Nehls, T., Ziegler, F., 2016. Evaluation of the health-risk reduction potential of countermeasures to urban heat islands. Energy and Buildings 114, 27–37. https://doi.org/10.1016/j.enbuild.2015.06.038. ISSN 0378–7788. http://www.sciencedirect.com/science/article/pii/S0378778815300657. SI: Countermeasures to Urban Heat Island.

Ciofalo, M., 2022. Large Eddy Simulation. Springer International Publishing, pp. 47–63 (ISBN 978-3-030-81078-8.).

Clarke, J.F., 1972. Some effects of the urban structure on heat mortality. Environ. Res. 5 (1), 93–104. https://doi.org/10.1016/0013-9351(72)90023-0. URL. ISSN 0013-9351. http://www.sciencedirect.com/science/article/pii/0013935172900230.

M. Daniel, A. Lemonsu, and V. Viguié. Role of watering practices in large-scale urban planning strategies to face the heat-wave risk in future climate. *Urban Clim.*, 23: 287–308, 2018. ISSN 2212–0955. doi: https://doi.org/10.1016/j.uclim.2016.11.001. URL https://www.sciencedirect.com/science/article/pii/S2212095516300505. ICUC9: The 9th International Conference on Urban Climate.

Djongyang, N., Tchinda, R., Njomo, D., 2010. Thermal comfort: a review paper. Renew. Sust. Energ. Rev. 14 (9), 2626–2640. https://doi.org/10.1016/j.rser.2010.07.040. ISSN 1364-0321. https://www.sciencedirect.com/science/article/pii/S1364032110002200.

Foken, T., 2017. Micrometeorology, 2 edition. Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-25440-6.

Gehrke, K.F., Sühring, M., Maronga, B., 2020. Modeling of land-surface interactions in the palm model system 6.0: land surface model description, first evaluation, and sensitivity to model parameters. URL. https://data.uni-hannover.de/dataset/ee0a4763-d6ad-423f-a573-323efa8ff841.

Gronemeier, T., Surm, K., Harms, F., Leitl, B., Maronga, B., Raasch, S., 2021. Evaluation of the dynamic core of the palm model system 6.0 in a neutrally stratified urban environment: comparison between les and wind-tunnel experiments. Geosci. Model Dev. 14 (6), 3317–3333. https://doi.org/10.5194/gmd-14-3317-2021. https://gmd.copernicus.org/articles/14/3317/2021/.

K. Hayhoe, S. Sheridan, L. Kalkstein, and S. Greene. Climate change, heat waves, and mortality projections for Chicago. *J. Great Lakes Res.*, 36:65–73, 2010. ISSN 0380-1330. Doi: https://doi.org/10.1016/j.jglr.2009.12.009. URL http://www.sciencedirect.com/science/article/pii/S0380133009002275. Potential Climate Impacts on Chicago and Midwest.

Hertwig, D., Patnaik, G., Leitl, B., 2017. Les validation of urban flow, part ii: eddy statistics and flow structures. Environ. Fluid Mech. 17 (3), 551–578. https://doi.org/10.1007/s10652-016-9504-x. ISSN 1573-1510.

Howell, J.F., Mahrt, L., 1997. Multiresolution flux decomposition. Bound.-Layer Meteorol. 83 (1), 117–137. https://doi.org/10.1023/A:1000210427798. ISSN 1573-1472.

Kaspar, F., Friedrich, K., 2019. Rückblick auf die temperatur in deutschland im jahr 2019 und die langfristige entwicklung. techreport. Deutscher Wetterdienst.

Kellnerová, R., Fuka, V., Uruba, V., Juráčková, K., Nosek, Štěpán, Chaloupecká, H., Jaňour, Z., 2018. On street-canyon flow dynamics: advanced validation of les by time-resolved piv. Atmosphere 9, 161. https://api.semanticscholar.org/CorpusID:54807698.

A. M. Kempf. Les validation from experiments. *Flow, turbulence and combustion*, 80(3): 351–373, 2008. ISSN 1573-1987. Doi: https://doi.org/10.1007/s10494-007-9128-9.

Kloog, I., 2019. Air pollution, ambient temperature, green space and preterm birth. Curr. Opin. Pediatr. 31 (2) (ISSN 1040-8703.).

A. Kubilay, J. Allegrini, D. Strebel, Y. Zhao, D. Derome, and J. Carmeliet. Advancement in urban climate modelling at local scale: urban heat island mitigation and building cooling demand. Atmosphere, 11(12), 2020. ISSN 2073-4433. doi: https://doi.org/10.3390/atmos11121313. URL https://www.mdpi.com/2073-4433/11/12/1313.

N. Liu and L. Morawska. Modeling the urban heat island mitigation effect of cool coatings in realistic urban morphology. J. Clean. Prod., 264(121560), 8 2020. ISSN 0959-6526. doi: https://doi.org/10.1016/j.jclepro.2020.121560.

Mahrt, L., Thomas, C.K., Prueger, J.H., 2009. Space–time structure of mesoscale motions in the stable boundary layer. Q. J. R. Meteorol. Soc. 135 (638), 67–75. https://doi.org/10.1002/qj.348. https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.348.

Maronga, B., Gryschka, M., Heinze, R., Hoffmann, F., Kanani-Suehring, F., Keck, M., Ketelsen, K., Letzel, M.O., Suehring, M., Raasch, S., 2015. The parallelized large-eddy simulation model (palm) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. Geosci. Model Dev. 8 (8), 2515–2551. https://doi.org/10.5194/gmd-8-2515-2015. https://www.geosci-model-dev.net/8/2515/2015/.

B. Maronga, G. Gross, S. Raasch, S. Banzhaf, R. Forkel, W. Heldens, F. Kanani-Suehring, A. Matzarakis, M. Mauder, D. Pavlik, J. Pfafferott, S. Schubert, G. Seckmeyer, H. Sieker, and K. Winderlich. Development of a new urban climate model based on the model palm - project overview, planned work, and first achievements. Meteorol. Z., 28(2): 105–119, 06 2019. doi: https://doi.org/10.1127/metz/2019/0909.

Martilli, A., Santiago, J.L., 2009. How to Use Computational Fluid Dynamics Models for Urban Canopy Parameterizations. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 31–37. https://doi.org/10.1007/978-3-642-00298-4_4. ISBN 978-3-642-00298-4.

Mehrotra, S., Bardhan, R., Ramamritham, K., 2019. Outdoor thermal performance of heterogeneous urban environment: an indicator-based approach for climate-sensitive planning. Sci. Total Environ. 669, 872–886. https://doi.org/10.1016/j.scitotenv.2019.03.152. URL. ISSN 0048-9697. http://www.sciencedirect.com/science/article/pii/S0048969719311349.

Nowak, D.J., Greenfield, E.J., 2020. The increase of impervious cover and decrease of tree cover within urban areas globally (2012–2017). Urban For. Urban Green. 49, 126638. https://doi.org/10.1016/j.ufug.2020.126638. URL. ISSN 1618-8667. https://www.sciencedirect.com/science/article/pii/S161886671930295X.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press.

OpenStreetMap contributors, 2017. Planet dump retrieved from. https://planet.osm.org.https://www.openstreetmap.org.

Parsons, K., 2002. Human Thermal Environments. CRC Press. https://doi.org/10.1201/b16750.

Raabe, A., Foken, T., 2003. Die hoehe der internen grenzschicht im windfeld nach einersprunghaften aenderung der aerodynamischen rauhigkeit der unterlage - beispiel kueste. Beitraege zur klima- und meeresforschung (F.-MChmielewski & T. Foken,Eds.), pp. 227–237.

Resler, J., Eben, K., Geletič, J., Krč, P., Rosecký, M., Sühring, M., Belda, M., Fuka, V., Halenka, T., Huszár, P., Karlický, J., Benešová, N., Ďoubalová, J., Honzáková, K., Keder, J., Nápravníková, Š., Vlček, O., 2021. Validation of the palm model system 6.0 in a real urban environment: a case study in dejvice, Prague, the Czech Republic. Geosci. Model Dev. 14, 4797–4842. https://api.semanticscholar.org/CorpusID:237589610.

Revi, A., Satterthwaite, D., Aragón-Durand, F., Corfee-Morlot, J., Kiunsi, R., Pelling, M., Roberts, D., Solecki, W., 2014. Urban Areas in Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 535–612. IPCC Chaper 8.

Salim, M.H., Schubert, S., Resler, J., Krč, P., Maronga, B., Kanani-Sühring, F., Sühring, M., Schneider, C., 2020. Importance of radiative transfer processes in urban climate models: a study based on the palm model system 6.0. Geosci. Model Dev. Discuss. 1–55. https://api.semanticscholar.org/CorpusID:225453614.

Santamouris, M., 2023. Chapter 5 - urban overheating—Energy, environmental, and heat-health implications. In: Paolini, R., Santamouris, M. (Eds.), Urban Climate Change and Heat Islands. Elsevier, pp. 165–225. https://doi.org/10.1016/B978-0-12-818977-1.00007-7. ISBN 978-0-12-818977-1. https://www.sciencedirect.com/science/article/pii/B9780128189771000077.

Seo, E., Dirmeyer, P.A., 2022. Understanding the diurnal cycle of land–atmosphere interactions from flux site observations. Hydrol. Earth Syst. Sci. 26 (20), 5411–5429. https://doi.org/10.5194/hess-26-5411-2022. https://hess.copernicus.org/articles/26/5411/2022/.

Staiger, H., Bucher, K., Jendritzky, G., 1997. Gefühlte temperatur. die physiologisch gerechte bewertung vonwärmebelastung und kältestress beim aufenthalt im freien in der maßzahl grad celsius. Ann. Meteorol. 33, 100–107.

Steeneveld, G.J., Wokke, M.J.J., Groot Zwaaftink, C.D., Pijlman, S., Heusinkveld, B.G., Jacobs, A.F.G., Holtslag, A.A.M., 2010. Observations of the radiation divergence in the surface layer and its implication for its parameterization in numerical weather prediction models. J. Geophys. Res. Atmos. 115 (D6). https://doi.org/10.1029/2009JD013074 https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD013074.

C. K. Thomas. Variability of sub-canopy flow, temperature, and horizontal advection in moderately complex terrain. *Bound.-Layer Meteorol.*, 139(1):61–81, 2011. ISSN 1573-1472. Doi: https://doi.org/10.1007/s10546-010-9578-9. doi: https://doi.org/10.1007/s10546-010-9578-9.

Toparlar, Y., Blocken, B., Maiheu, B., van Heijst, G., 2017. A review on the cfd analysis of urban microclimate. Renew. Sust. Energ. Rev. 80, 1613–1640. https://doi.org/10.1016/j.rser.2017.05.248. URL. ISSN 1364-0321. https://www.sciencedirect.com/science/article/pii/S1364032117308924.

Vermessungsverwaltung, B., Dec. 2024. Alkis - tatsaechliche nutzung (tn). https://geoservices.bayern.de/od/wms/alkis/v1/tn.

Vogel, J., Afshari, A., Chockalingam, G., Stadler, S., 2022. Evaluation of a novel wrf/palm-4u coupling scheme incorporating a roughness-corrected surface layer representation. Urban Clim. 46, 101311. https://doi.org/10.1016/j.uclim.2022.101311. URL. ISSN 2212-0955. https://www.sciencedirect.com/science/article/pii/S2212095522002292.

Wetterdienst, D., 2024. Lufttemperatur: vieljährige mittelwerte 1991–2020. https://www.dwd.de/DE/leistungen/klimadatendeutschland/mittelwerte/temp_9120_SV_html.html?view=nasPublication&nn=16102.

Winklmayr, C., Muthers, S., Niemann, H., Mücke, H.-G., M. an der Heiden., 2022. Heat-related mortality in Germany from 1992 to 2021. Dtsch. Arztebl. Int. 119 (26), 451–457. https://doi.org/10.3238/arztebl.m2022.0202. https://www.aerzteblatt.de/int/article.asp?id=225956.