

Improved mesoscopic meteorological modelling of the urban climate for building physics applications

Conference Paper**Author(s):**

Stebel, Dominik André ; Derome, Dominique; Kubilay, Aytaç; Carmeliet, Jan

Publication date:

2023-12

Permanent link:

<https://doi.org/10.3929/ethz-b-000647653>

Rights / license:

[Creative Commons Attribution 3.0 Unported](#)

Originally published in:

Journal of Physics: Conference Series 2654, <https://doi.org/10.1088/1742-6596/2654/1/012147>

Improved mesoscopic meteorological modelling of the urban climate for building physics applications

D Strebel^{1,*}, D Derome², A Kubilay¹, J Carmeliet¹

¹ Chair of Building Physics, Swiss Federal Institute of Technology ETH, Zurich, Switzerland

² Department of Civil and Building Engineering, Université de Sherbrooke, Sherbrooke, QC, Canada

* Corresponding author (strebdm@ethz.ch)

Abstract. A meteorological mesoscale model is used to predict the local urban climate at 250 m resolution. The authors propose a hybrid machine learning approach to improve the prediction accuracy and remove simulation bias. Two case studies are presented to show the improvements of the simulation accuracy. Based on the hybrid model results, using cooling degree hours is proposed as an insightful time-dependent index to map local hotspots and assess the difference of cooling loads between rural and urban environments.

1. Introduction

Due to the increasing magnitude and occurrence frequency of heatwaves and their important impact on buildings and urban environments, there is an urgent need for predicting the urban climate and local microclimate more accurately. During heatwaves, outdoor and indoor thermal comfort may deteriorate leading to excessive heat stress on pedestrians and inhabitants. This negative impact may increase with climate change, for example rendering passive cooling by night ventilation ineffective and requiring the use of active cooling. Moreover, we know that cities show an urban heat island (UHI) effect, and that specific neighborhoods become hot spots during heatwaves. There is an urgent need to provide building and urban physics studies with appropriate environmental boundary conditions taking into account the local urban climate.

Mesoscale Meteorological Models (MMM) are commonly used for limited area weather prediction applications. They are driven by lateral boundary and initial conditions from global circulation models. They model the time dependent state of the atmosphere on a regular grid. Subgrid scale physics is represented by parametrized models. These parametrized processes usually include radiation, cloud physics, boundary layer, land use interaction but can also include urban climate effects. There are multiple urban canopy parametrizations with different approaches and model complexities available to use with MMMs. The use of these schemes however requires the availability of detailed data describing the urban environment, like building geometry, land use data and thermal properties of buildings.

Recently, the authors coupled MMMs with their in-house Computational Fluid Dynamics (CFD) urban microclimate model which is solved using OpenFOAM [1]. In this one-way coupling approach,



boundary conditions from MMMs are used to drive the urban microclimate model by interpolating wind and temperature profiles at the boundaries of the CFD domain. The urban microclimate model also considers heat and moisture transport in the urban and building materials, as well as longwave and shortwave radiation exchange and heat and moisture impacts of vegetation. MMM results have also been used by the authors as input to Building Energy Simulation (BES) tools to analyze summer overheating and excess building cooling demand during heatwave days and hot summers [2]. The authors compared WRF (Advanced Weather and Research Model) MMM predictions with weather station measurements and used the forecast data as input for Energy+, finding that building cooling demand can be highly underpredicted during hot summers using standard reference meteorological data.

Here we present a new approach to increase prediction accuracy for local urban climate simulations. This method is based on a hybrid workflow, where MMM simulation output is improved with measurements-based machine learning, in order to simulate city scale climatic conditions and providing accurate spatial boundary conditions for the CFD-based local urban microclimate model.

The paper is organized as follows. First, we present two mesoscale meteorological models, compare their results with measured data from three sources and propose an improvement of the models using machine learning. Then, we analyze the urban heat island based on the MMM results using the cooling degree hours metrics. Finally, we finish with conclusion and outlook for future research.

2. Mesoscale meteorological modelling (MMM) for the urban climate

MMMs, such as WRF (Advanced Weather and Research Model) [3] and COSMO (Consortium for Small-scale Modelling) [4], commonly used for weather prediction [5], can also be configured to study the urban climate at 250 m resolution. The methods are based on a down-nesting of initial and boundary conditions coming from global models like GFS (Global Forecast System), ECMWF (European Centre for Medium-Range Weather Forecasts) or nested models derived from the global ones like MeteoSwiss COSMO-2/1. At the boundaries of the spatial domain, lateral boundary conditions are prescribed which come from higher scale results and a relaxation zone is implemented to blend the forced conditions at the boundaries together with the atmospheric model of a limited area [6]. Figure 1 shows as an example a three down-nesting approach for the city of Zurich, where the largest domain D01 has a grid size of 6.25 km and the smallest domain D03, which is the domain of interest, a grid size of 250 m.

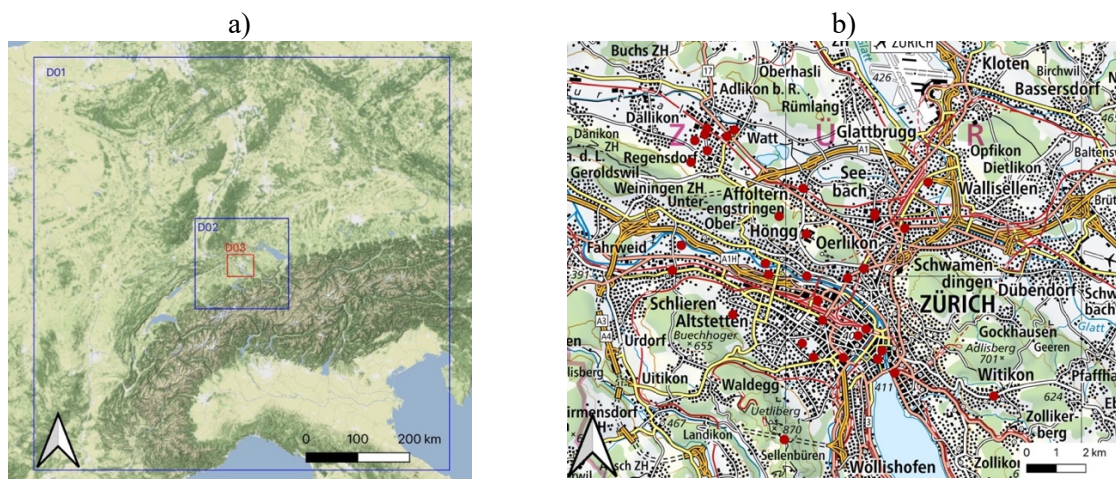


Figure 1a). Nested domain setup for WRF Simulations of Zurich with three domains D01, D02 and D03.

Figure 1b). Ground truth sensor locations in Zurich and its surroundings

Several physical processes occurring at urban scale cannot be taken into the calculations in MMMs at the scale of 250 m and are represented by parametrized models. In COSMO, we use as urban canopy parametrization the double-canyon effect parametrization (DCEP) model for multi-layer urban canopy representation proposed by Schubert et al. [7]. This model takes into account the radiation exchange for two neighboring canyons treating direct and diffuse radiation separately. Building morphology in each $250 \times 250 \text{ m}^2$ area is grouped in four categories with different street orientation, building height and street width. The use of DCEP needs a detailed preparation of input data such as building geometry, and radiative and thermal properties of urban surfaces, which can be processed from data such as LIDAR or building databases. In the present study, COSMO-DCEP simulations are performed with a timestep of 5 seconds and with hourly model output. Mussetti et al. [8] studied the performance of this method for Zurich and showed that the performance of the model improved the prediction of the urban climate when using a higher resolution of 250 m compared to a commonly-used resolution of 1 km.

In WRF, NOAA LSM (Land Surface Model) is used to model the exchange of heat and moisture between soil and the atmosphere. The NOAA scheme is widely used in operational weather forecasting and is considered to be one of the default LSM schemes in WRF [9]. It was developed through an interdisciplinary research project, grouping several universities and US governmental agencies, and continues to be improved by WRF developers. The LSM scheme plays a critical role in predicting surface and near-surface variables like air temperature at 2 m, surface temperature, wind speed and energy fluxes between the atmosphere and the ground. In our WRF setup, we use the NOAA LSM directly, without any additional urban canopy model.

Figure 2a shows the locations of three measurement stations in Zurich used for comparison with the results of WRF and COSMO. These three stations belong to three different environment types. Kaserne station is situated in downtown Zurich and part of the National Air Pollution Monitoring Network (NABEL). SMA station is located on a south-west facing hill around 170 m above the city center. The airport station is situated just outside of the city, in an area not considered fully rural.

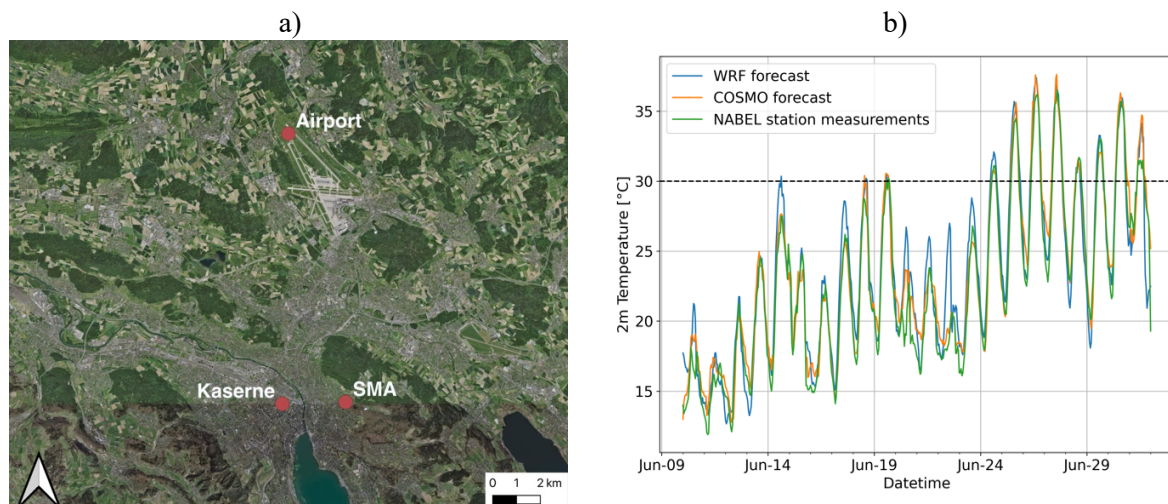


Figure 2 a). Locations of selected measurement stations in Zurich

Figure 2b). Comparison of air temperature at 2 m height between WRF, COSMO and measured data, the latter measured at a meteorological measurement station in central Zurich (NABEL) during a summer period from June 10th 2019 until July 10th 2019

Figure 2b compares the predictions of air temperature from the closest COSMO and WRF cell at 2 m height above ground with the measured data for the heatwave during June 2019 at the urban station Kaserne. The summer period contains a heatwave, with air temperatures above 30° for at least eight consecutive days.

Both models show differences between measurements and predictions. WRF overpredicts daily peaks especially between June 20th and June 23rd but also on June 15th and June 18th. COSMO predicts the daily peaks better but predicts too warm nights, for example from June 20th until June 23rd and on June 26th and 27th. Table 1 shows the RMSE (Root Mean Square Error) values for all stations. None of the RMSE values are bigger than 3°C. COSMO with the computationally-intensive urban canopy performs clearly better. Remarkably, WRF with the default NOAH LSM is not much worse especially in the fully urban environment like the one at Kaserne station.

Table 1. RMSE values of COSMO and WRF predictions with selected measurement stations during heatwave from June 10th to July 10th 2019

Reference station	COSMO	WRF
Kaserne (downtown)	1.26°C	1.87°C
SMA Fluntern (semi urban)	1.82°C	2.76°C
Airport (semi-rural)	1.57°C	2.8°C

It is however clear from the above results that a further improvement of MMM simulation results can be necessary. We present a hybrid method based on Machine Learning (ML) to improve the agreement between simulation and observation data.

Our method is illustrated for a WRF simulation in the domain of interest. The WRF simulation results (referred to as WRF-only) are corrected for the bias between simulation and measured data with a machine learning approach (referred to as WRF+ML). We use two different Swiss cities as case studies: Bern and Zurich. In both cities, a dense network of microclimate stations are built in the last three years to provide diverse ground truth data over the whole city. Figure 2 b) shows the spatial distribution of temperature sensors over Zurich and its surrounding environment. In addition to the sensors in the urban environment, there are also sensors in the rural environment all over the Canton of Zurich [10]. These stations measure the temperature of the local environment, which is also the target variable for the machine learning submodel. The authors experimented with different ML algorithms. It is found that the best RMSE performance is provided using a random forest approach. The WRF simulation was carried out over two summer months from June 1st until July 31st, 2019. This timespan includes several heatwave periods with maximum temperatures over 30°C for more than three consecutive days but also colder days with average or below average temperatures. Data from all stations are used for training the model, but it would also be possible to classify the stations by the associated environment. All dynamic variables of the WRF output are included as predictors. The most significant ones are found to be $T2$ (Temperature at 2 m, feature importance: 0.77), followed by $GRDFLX$ (Ground heat flux, feature importance: 0.049), $TH2$ (Potential Temperature at 2 m, feature importance: 0.014) and MU (perturbation dry air mass in column, feature importance 0.007).

Additional static data like topography, land classification according to cadaster and exposition are also included. None of the additional static data has any positive influence on the prediction accuracy. One possible explanation for this is the indirect inclusion of the static data in the WRF simulation itself, i.e. WRF already uses static datasets, like topography and land use data, for boundary layer, surface and soil parametrization. Hence, including this data into the ML process does not introduce new information.

Table 2 shows the RMSE values between WRF-only (without any bias correction), WRF+ML predictions and the unseen testing dataset from all sensor locations, respectively. A remarkable improvement of the accuracy of the predictions is observed showing a decrease of RMS from 2.82°C to 0.74°C.

Table 2. RMSE of WRF-only and WRF+ML predictions

	WRF	WRF+ML (testing dataset)
RMSE	2.82°C	0.74°C

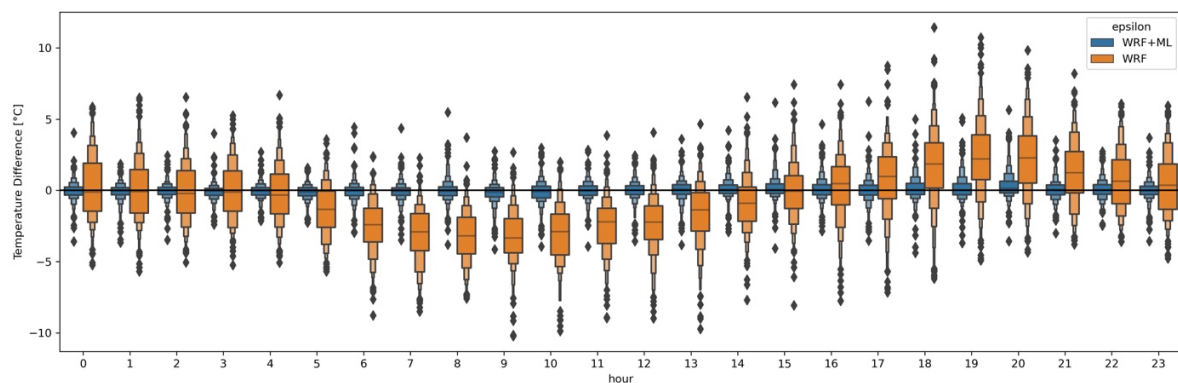
Figure 3 aggregates the temperature differences between measurements and WRF-only and WRF+ML approach to hourly box plots of the distribution over the whole simulation period. One can clearly see that the WRF-only approach underpredicts the temperature in the mornings but overpredicts the evening temperature. This trend is visible for all stations and has been already observed in other simulation cases by the authors. By applying our machine learning model, we can eliminate this bias completely. This leads to lower RMSE values and a significant increase of the prediction accuracy.

We carried out a similar case study for the City of Bern. Gubler et al. built a low-cost sensor network covering the city of Bern to measure air temperatures in different urban neighborhoods [11]. The used data is collected for the summer period 2018. A 250 m x 250 m WRF simulation is carried out for the available longest period, where data is available from all sensors. This ends up being a simulation period from June 25th until July 8th, a comparatively shorter time span.

The RMSE between measurements and WRF-only approach results in a high value of 3.19°C. With our hybrid approach we could improve the predictions to a RMSE value of 0.94°C. However, we note that the simulation period is quite short and there are some uncertainties about the performance of these low-cost sensors, especially about shielding during direct solar irradiation.

A further step is to transfer the fitted model from one city to another. Since we found the best results for the Zurich case, we transfer the trained Zurich model and test it on the Bern data. In this case we found no improvement and RMSE even increased to 3.4°C. There are multiple possible reasons for this observation: the comparatively short simulation period for the city of Bern did not fit the model well; the simulated year, 2018, is different from the 2019 one used to fit the model in Zurich; the sensors in Berne are not the same device types as in Zürich and can react differently in similar conditions. All these issues can influence the model performance.

We conclude that using measured data in the city can significantly improve the RMSE values and accuracy of the predictions. This approach is very promising and computationally much less expensive than the use of parametrized urban canopy models. It also needs no additional static input data. One possible future application would be to use crowdsourcing public measurement networks like Netatmo as input dataset for ML [12].

**Figure 3.** Boxenplot of hourly differences between measurements, WRF-only and WRF+ML for the simulations in Zurich

3. Urban heat island characterization using cooling degree hours (CDH)

Using cooling degree hours (CDH) as metrics, we analyze the cooling energy demand spatially over the Zurich agglomeration and identify the urban heat island (UHI), showing the capacity of this method to locate hot spots. The cooling degree hours for a certain set point temperature is defined as:

$$\text{CDH}_{t_s} = \sum_{k=0}^n \begin{cases} t_k - t_s, & t_k > t_s \\ 0, & t_k \leq t_s \end{cases} \quad (1)$$

where t_s equals the set point temperature, k the simulation hour, n the total hours of the considered period and t_k the predicted temperature at 2 m height. In this analysis, we choose a set point temperature of 22°C, which is considered to be a reasonable temperature for an average insulated building with limited summer cooling measures. A higher setpoint temperature may be used for buildings with passive cooling, decreasing the CDH.

Figure 4a shows the map of CDH for the summer period of 2019 for the Zurich agglomeration. The highest CDH values are observed for the city center, for agglomerations along the lake of Zurich and extensions of the city along the valleys surrounding the city of Zurich. Figure 4b shows a map of soil sealing (asphalt or concrete pavement) or imperviousness in %. We observe a close correlation between sealed soil and CDH, which indicates that dense urban areas with less greenery show higher air temperatures and CDH. Overall, we note that CDH values of approaching 6000 are seen in Zurich compared to the average values lying between 2500 and 4000 for the suburbs and less than 2500 in the rural areas. Figure 5 shows the relative increase of the CDH for the urban environment compared to the rural area. The urban and rural areas are defined based on land use categories. We exclude all urban area and water bodies for the calculation of a rural average. We then calculate the relative difference in CDH between the rural average and the local CDH values. We observe that parts of the city of Zurich and its surrounding agglomeration show increases up to more than 100%, meaning that the cooling demand as estimated by CDH is for a building in the city to be twice the amount than a similar building located in the surrounding rural area.

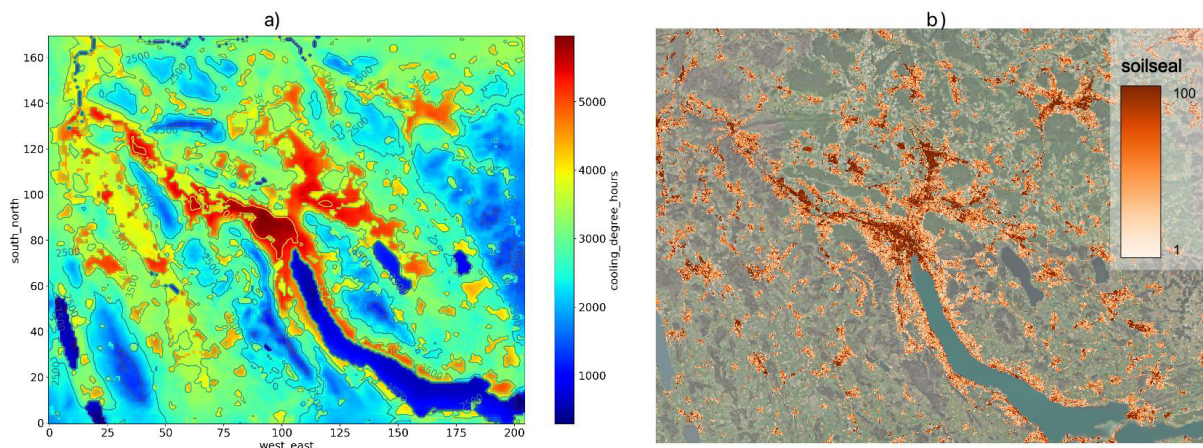


Figure 4 a). CDH map for Zurich with a set point temperature of 22°C from 2019-06-01 until 2019-09-30.

Figure 4 b). soil sealing in [%] of the same domain, based on [13]

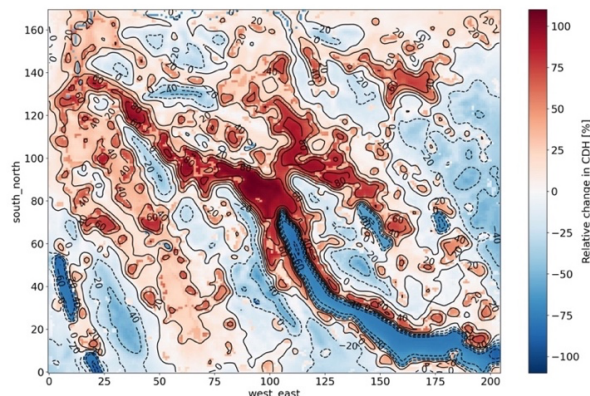


Figure 5. Relative change of CDH between local CDH and average rural environment in %

It is however important to note that the CDH is a proxy for the real cooling energy demand of a real building, which may vary depending on an array of parameters. It should be considered as a general metrics to map cooling energy demand spatially and indicate hot spots to be studied in more details.

4. Conclusions

Understanding the urban heat island and its impact is a key for sustainable and comfortable living in cities around the globe. Due to climate change, this topic is becoming even more important. However, urban climate simulations by mesoscale meteorological models with urban parametrizations can become computationally expensive and spatial climate measurement data is scarce, both factors making it difficult to identify local hotspots in cities.

In this paper, the authors introduce a new hybrid method to improve mesoscale meteorological simulations of the urban environment. This approach is based on common WRF simulations with an added machine learning algorithm trained by a measurement grid in the domain of interest. It is shown that the accuracy of the results can be improved significantly. It is also showed how these improved simulations can be used to calculate cooling degree hours, an insightful time-based index for evaluating hot spots in the urban environment.

These improvements reduce the simulation complexity and can help researchers and practitioners to assess hot spots more quickly in a city. With a broader deployment of sensor networks seen in cities it will be possible to use this method in more cities. It is also possible to harvest citizen science networks like Netatmo and use the data for training and validation of the model.

We also introduce cooling degree hours, a new metric to assess local hotspots on mesoscale in cities like heating degree days or hours in Building Physics. With this number, it is straightforward to detect spatial variability of the urban heat island effect and evaluate different parts of cities for future spatial planning and densification. We show a clear spatial correlation between sealed areas and cooling degree hour values. With a rural average, one can have a relative measurement between city and its environment.

A possible next research step is to transfer the hybrid model from one city to another and investigate the performance in a completely new setting. Furthermore, one could find out the necessary kind of training cities and sensors for broad application to other cities. Finally, these improved results of the local climate can be used for further building physics analyses, such as the study of the durability of building envelopes and heating and cooling energy demand or for public health studies.

5. References

- [1] Kubilay A, Allegrini J, Strebel D, Zhao Y, Derome D and Carmeliet J 2020 Advancement in Urban Climate Modelling at Local Scale: Urban Heat Island Mitigation and Building Cooling Demand

Atmosphere **11** 1313

- [2] Boudali Errebai F, Strebel D, Carmeliet J and Derome D 2022 Impact of urban heat island on cooling energy demand for residential building in Montreal using meteorological simulations and weather station observations *Energy and Buildings* **273** 112410
- [3] Skamarock C, Klemp B, Dudhia J, Gill O, Liu Z, Berner J, Wang W, Powers G, Duda G, Barker D and Huang X 2019 A Description of the Advanced Research WRF Model Version 4
- [4] Doms G and Baldauf M 2011 A description of the nonhydrostatic regional COSMO model. Part I: Dynamics and numerics *Deutscher Wetterdienst, Offenbach*
- [5] Orlandi I 1975 A Rational Subdivision of Scales for Atmospheric Processes *Bulletin of the American Meteorological Society* **56** 527–30
- [6] Davies H C 1976 A lateral boundary formulation for multi-level prediction models *Quarterly Journal of the Royal Meteorological Society* **102** 405–18
- [7] Schubert S, Grossman-Clarke S and Martilli A 2012 A Double-Canyon Radiation Scheme for Multi-Layer Urban Canopy Models *Boundary-Layer Meteorol* **145** 439–68
- [8] Mussetti G, Brunner D, Allegrini J, Wicki A, Schubert S and Carmeliet J 2020 Simulating urban climate at sub-kilometre scale for representing the intra-urban variability of Zurich, Switzerland *International Journal of Climatology* **40** 458–76
- [9] Campbell P C, Bash J O and Spero T L 2019 Updates to the Noah Land Surface Model in WRF-CMAQ to Improve Simulated Meteorology, Air Quality, and Deposition *J. Adv. Model. Earth Syst.* **11** 231–56
- [10] Baum F and Sintermann J 2021 *Stadtklimamessungen: Ausgewählte Resultate Sommer 2020*
- [11] Gubler M, Christen A, Remund J and Brönnimann S 2021 Evaluation and application of a low-cost measurement network to study intra-urban temperature differences during summer 2018 in Bern, Switzerland *Urban Climate* **37** 100817
- [12] Coney J, Pickering B, Dufton D, Lukach M, Brooks B and Neely III R R 2022 How useful are crowdsourced air temperature observations? An assessment of Netatmo stations and quality control schemes over the United Kingdom *Meteorological Applications* **29** e2075
- [13] CORINE Imperviousness — Copernicus Land Monitoring Service