Investigating the Suitability of the WRF Model for Improving Prediction of Urban Climate Boundary Conditions

Kristopher Hammerberg – TU Wien – kristopher.hammerberg@tuwien.ac.at
Oscar Brousse – KU Leuven – oscar.brousse@kuleuven.be
Ardeshir Mahdavi – TU Wien – amahdavi@tuwien.ac.at

Abstract

Urban populations continue to increase in parallel with global temperatures. The result is an increasing number of people affected by increasingly severe urban heat conditions. Understanding these effects and being able to accurately account for the effects of the urban climate on building energy use is important for urban and architectural design decision making. This paper presents part of an on-going research effort to evaluate the Weather Research and Forecasting (WRF) model as a tool for improving prediction of boundary conditions in urban climates. WRF is a regional climate model that is capable of downscaling global weather data to a fine resolution and includes detailed urban canopy models. The use of a numerical model in urban climate studies would allow for computational experiments involving changes to the urban fabric and future climate scenarios. In this study, Vienna, Austria, was used as a test case. The weather was simulated over five 48-hour periods, which were selected using cluster analysis to best represent typical weather conditions in Vienna. The model results were then compared to data collected from a network of 170 weather stations throughout the region of interest. Additionally, the land-use classification and urban parameterization in the model domain were improved using high-resolution GIS data from the city of Vienna. Results show a great deal of variation in the accuracy of the model under different weather conditions. Although individual problems can be identified during specific intervals, there is no obvious trend or bias to the variation across all time periods. The extent of the variation indicates the model results are not suitable for use as boundary conditions for building performance models throughout an entire year.

1. Introduction

In recent years, cities have been increasingly considered to be both contributors and stakeholders in the global climate change discussion, yet scientific precedents for the study of the climate impact of cities stretch back over two centuries (Hebbert and Jankovic, 2013). The commonly understood concept of the urban heat island, for instance, finds its roots in the work of Luke Howard in London (1833) and James Gordon's temperature survey of the Salt River Valley (1921). In our current context of mass urbanization and global warming, urban planners and architectural designers could benefit from improved urban climate representtations. An accurate urban climate model would allow for quantitative decision support in both the response of buildings to altered boundary conditions and the response of the local climate to changes in urban development.

The misrepresentation of external boundary conditions in building simulation can result in fairly large errors. A study in Bahrain showed that using outdated weather files can underestimate annual electricity consumption by 14.5 % and cooling loads by up 8.9 % (Radhi, 2009). Weather data for the typical meteorological year (TMY) are frequently used in building simulation models and are often derived from airport weather stations (Barnaby and Crawley, 2011). These locations do not represent the semantic and geometric complexities of urban environments or the resulting microclimate effects (Pernigotto et al., 2014). Bhandari et al. (2012) investigated the use of synthesized location specific weather files from third party providers and found that monthly building loads can vary up to ± 40 %. In another study, Chan (2011) examined the range of impact of climate change variation by morphing the

TMY based on the projections of a general circulation model (GCM). This revealed the potential for a substantial increase in A/C energy demand ranging from 3.7 to 24 % for residential buildings.

Together, advances in numerical weather models, the increasing availability of highly accurate GIS data, and the proliferation of weather monitoring stations allow for new avenues for exploring the complexities of the urban climate. In this study, we investigated the performance of the WRF model over the city of Vienna, Austria, under various weather conditions with the aim of using the output to improve the boundary conditions of building energy models set in urban areas. The WRF model was chosen as it incorporates a detailed subgrid representation of the urban morphology, called the building environment parameterization and building energy model (BEP+BEM).

2. Methodology

2.1 Study Area

This study was centered on the city of Vienna, Austria. In comparison to other Central European cities with intact historic centers, Vienna is morphologically rather typical. According to the Köppen climate classification, it lies at the edge of the oceanic/subtropical zone (Cfb) that covers most of Western Europe and shares characteristics with both the warm summer continental climate to the East (Dfb) and the humid subtropical conditions (Cfa) of some Northern Italian cities (Kottek et al., 2006). Its municipal boundaries inscribe an area of 414.87 km^2 and approximately 1.8 million residents (Lukacsy and Fendt, 2015). It lies at the eastern edge of the Alps with the western edge of the city rising into the Wienerwald, and the eastern edge stretching over the Danube River and into the Vienna Basin. Our area of interest contains the municipal boundary of Vienna as well as the surrounding suburban and rural areas (Fig. 1).

The availability of both weather data records as well as high-resolution geospatial data informed our decision to select Vienna as our case study. Additionally, with the notable exception of Stuttgart (Fallmann et al., 2014), Berlin (Jänicke et al., 2016), and Madrid (Brousse et al., 2016) relatively few European cities have been examined in WRF modeling studies.

Fig. 1 – Vienna, Austria. Region of interest (red). Municipal boundary (black)

2.2 Study Period

Although numerical weather models can be excellent tools for computational experiments and weather prediction, they are resource intensive. In this study, we hoped to examine the utility of WRF across different weather conditions. Rather than run simulations with season-long or year-long durations, we selected 5 representative periods each with a 48-hour duration.

These representative dates were selected with a k-means cluster analysis method and daily means for the Schwechat Airport provided by Austria's Zentralanstalt für Meteorologie und Geodynamik (ZAMG). This station is well calibrated and sited in an open area that allows for good mixing and less impact of extreme micro-climate effects, making it well suited for classifying regional weather types. Then, three uncorrelated climatic variables were selected to use as key indicators in clustering: temperature, diurnal temperature range, and wind speed. There is no objective or automated method for selecting an optimal number of clusters with the k-means approach. For this study, we needed to achieve a balance between investigating the widest range of weather conditions and reducing the total number of simulations. In order to achieve this

balance, we chose to use 5 weather categories after an iterative testing process (Fig. 2).

The final step was to select the most representative 48-hour period within the cluster for simulation. In order to guarantee that each period best represents the aspects of the category that distinguish it from the other four, we selected days from each category with values furthest from the mean of the other clusters (Table 1).

2.3 Land Cover and Urban Canopy Parameters

2.3.1 WUDAPT Land Cover

The WRF model requires a land cover map to describe the surface condition and calculate its interaction with the atmosphere. For this study we used the World Urban Database and Access Portal Tool (WUDAPT) to provide the land cover map. WUDAPT is a state-of-the-art web portal that produces land cover maps of cities using the Local Climate Zone (LCZ) concept developed by Stewart and Oke (2012). These maps are created for use in urban climate modeling. The method classifies Landsat satellite imagery with a Random Forest Classification algorithm (Bechtel et al., 2105). Training areas, which best represent each LCZ, are selected by the user. Then the algorithm uses the distinct reflective signature in each spectral band to classify each 100x100 m pixel within the region of interest (ROI).

Table 1 – Representative dates and key climate indicators

Dates	Mean Temp. [°C]	Diurnal Range [K]	Mean Wind Speed [m/s]
January 7-9	-1.62	5.13	1.64
February 8-10	-0.32	4.31	5.00
March 20-22	6.93	15.24	0.69
April 21-23	15.00	12.30	2.49
July 5-7	27.26	16.13	1.56

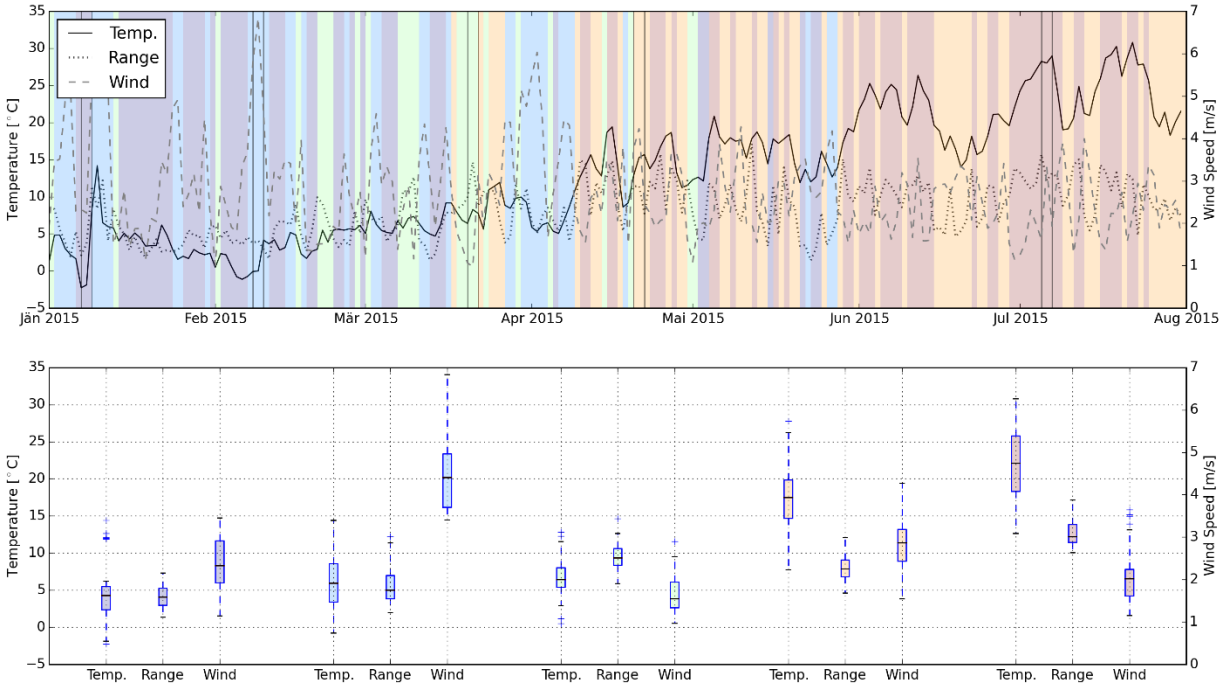

Fig. 2 – Key indicator time series with clustering (top) and boxplot by cluster (bottom)

2.3.2 Urban Canopy Parameters from GIS

In addition to the land cover map, the sub-grid urban model in WRF, the BEP+BEM, requires a morphological description or urban canopy parameters (UCPs) for each urban class in the land cover map. Although there are typical ranges provided by Stewart and Oke (2012), in Vienna we could extract these values directly from a detailed database of geospatial data made available by the City of Vienna ("Open Government in Vienna," 2016). This database includes a high-resolution vector map of all the building and land cover

features within the city. It allows for the precise calculation of most of the required UCPs (e.g. height-to-width ratio, pervious fraction, average building height).

In order to extract these values, an algorithm was created with the Quantum GIS software (Quantum GIS Development Team, 2016). It iterates over the region of interest on a 100x100 m grid and calculates the UCPs for each cell. Building height and land use type are included in the metadata of each polygon. With the use of the metadata and the polygon areas, it was straightforward to calculate the necessary area fractions (i.e. pervious, impervious, and urban fractions). The average building height was calculated by weighting the height values with the building footprint area. Likewise, the histogram of height distribution was calculated using the percentage of total building footprint area in 5 ft bins. The height-to-width ratio was computed with a 4-pass raster method (Fig. 3) adopted from Burian et al. (2003).

This algorithm produced a gridded data set with unique UCP values for each cell. However, this level of detail could not be used directly with WRF, as the GIS data is only available within the municipal boundaries and did not cover the entire ROI. Thus, mean values were used for each LCZC.

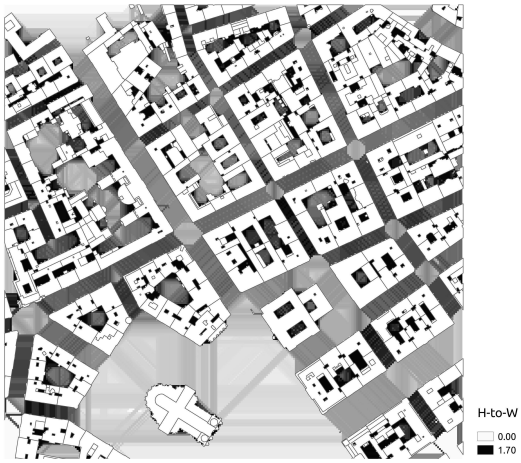

Fig. 3 – Example of the 4 pass height-to-width ratio results

2.4 Model Setup

The BEP+BEM (Martilli, 2002; Salamanca et al., 2010), under the Bougeaut-Lacarrère at 1.5 degree turbulence scheme, was chosen for this study for its multilayer urban canopy representation. This sub-grid urban model is currently the most detailed urban representation within the WRF model. The meteorological boundary conditions used for the simulations were derived from the NCEP Final Analysis (FNL from GFS) data at 1° resolution and taken every 6h for each of the two-day simulations. Three nested domains were used to downscale the results to an inner domain with a 500 m resolution (Fig. 4).

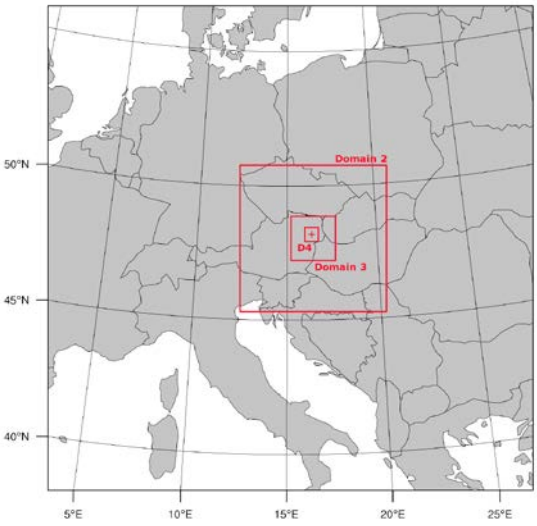

Fig. 4 – WRF nest domain configuration

2.5 Observational Data

We used three different sources for weather data in this study: the Austrian ZAMG, the City of Vienna Department for Environmental Protection (MA22), and Wunderground's Personal Weather Station Network (PWSN). The stations from ZAMG and the MA22 are installed and maintained by official organizations, however they provide sparse coverage (11 stations). On the other hand, the PWSN provides excellent coverage, however these stations are installed and maintained by amateur weather enthusiasts. At the beginning of the study period, January 2015, there were 310 active PSWN stations in our ROI. Although this number has increased to 873 at the time of writing, only those initial stations were included (Fig. 5).

While the PWSN's crowdsourcing approach provides a weather station network with excellent spatial coverage, it introduces several sources of uncertainty. There is no guarantee that recommended standards for installation are followed. In fact, a

number of stations in a similar crowdsourced network in Berlin were setup indoors (Meier et al., 2015).

Furthermore, there is no standard for the type of hardware used. Of the stations that provide information about their hardware type, the majority (78.8 %) are Netatmo stations. While the manufacturer claims acceptable sensor accuracy, these stations are not ventilated or shielded so they have to be set up where they are not in direct sunlight or they risk overheating. Rain and wind gauge accessories are optional and not included with many stations.

In order to address the potential for increased measurement errors as a result of using these unverified stations, we used several data quality control filters to remove erroneous data. These filters reduced the number of stations from 310 to 160. First, all data that greatly exceeded historic global extremes at the Earth's surface were removed. Such extreme values usually represent error codes.

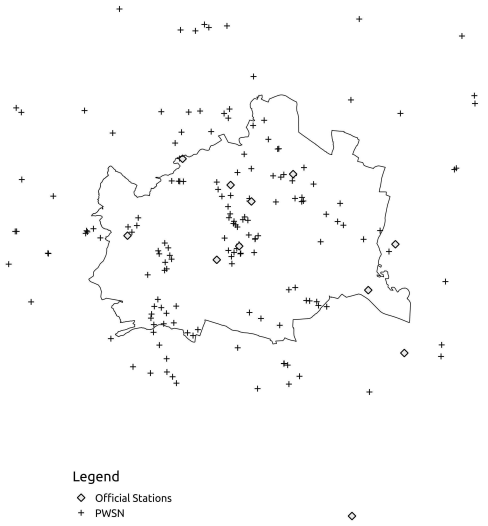

Fig. 5 – Weather station network

Then, outliers were identified with a two-tailed median absolute deviation (double MAD) test to account for skewed distributions and to prevent the outliers from influencing the selection process. The outlying data were then removed and stations that had more than 25 % of their values flagged as outliers were entirely removed from the dataset.

In the final step, the remaining PWSN stations were compared to nearby official stations. The comparison included two additional filters designed to target specific types of error. The first targeted weather stations that were installed indoors by comparing the average daily minimum temperature. The second filter targeted overheating due to lack of radiation shielding, ventilation or improper positioning.

3. Results

3.1 Near-Surface Temperature

The modelled temperatures at 2 meters were compared against a network of official stations and the PWSN. In general, the deviations between the model and observations were large (Table 2).

The model tended to overestimate near surface temperature in the cold study periods and underestimate during the warmer two periods. Also, the model tended to perform better during the day when it had both lower RMSE and less bias. However, at a closer examination of the diurnal variation there is no clear pattern between study periods that might allow for a consistent correction factor or transformation (Fig. 6a): (1) both in January and February, the modeled temperatures rose drastically on the second day; (2) despite that temporary deviation, February had the lowest overall RMSE; (3) in March temperatures were overestimated during the night, but daily temperatures showed good agreement; (4) April is the only period to underestimate the temperatures for the whole duration; (5) overall, July and April cases accurately represented the daily temperature ranges and fluctuation.

Table 2 – 2 m Temperature results [°C]

	RMSE	MB	Mean	SD	Min	Max
Overall	3.04	0.62	10.06	10.48	-8.69	33.71
Day	2.65	-0.10	12.48	11.17	-7.54	33.71
Night	3.40	1.34	7.65	9.15	-8.69	32.23
January	3.62	1.87	0.24	3.91	-8.69	10.34
February	1.94	0.84	0.46	2.25	-5.67	5.19
March	3.76	2.66	9.49	3.48	0.46	16.03
April	2.82	-1.74	13.17	3.92	2.11	19.46
July	2.82	-0.31	26.91	4.39	15.02	33.71

3.2 Global Radiation

Global shortwave radiation was overestimated by the model in nearly every study period. This is likely due to both the obstruction of weather station sensors by surrounding obstacles and cloud formation inaccuracies in the model. The February study period had the most accurate modelled solar radiation and even underestimates the solar radiation on the second day, which indicates that the shortwave radiation was not driving the extreme overestimation of temperature during the same period (Table 3).

Table 3 – Global shortwave radiation results [W/m^2]

	RMSE	MB	Mean	SD	Min	Max
Overall	87.75	26.88	175.53	274.35	0.00	947.25
January	57.27	14.36	38.12	82.46	0.00	311.11
February	43.88	2.12	38.10	73.94	0.00	420.89
March	90.38	35.38	202.76	266.78	0.00	722.62
April	98.07	26.82	258.62	303.51	0.00	867.09
July	123.31	54.47	340.05	358.25	0.00	947.25

3.3 Wind Speed

Wind speed measurements showed consistently large differences between official weather stations and the PWSN. This was likely due to the difficulty of properly positioning wind gauge instruments. Therefore, for wind measurements only the official stations were used in the analysis (Fig. 6b).

Table 4 – Wind speed results [m/s]

	RMSE	MB	Mean	SD	Min	Max
Overall	2.54	0.42	3.76	2.95	0.00	14.17
Day	2.33	0.19	3.70	2.80	0.00	13.14
Night	2.73	0.65	3.82	3.10	0.05	14.17
January	2.25	0.51	3.15	2.04	0.11	12.72
February	3.69	1.16	7.82	2.16	2.83	14.17
March	1.46	0.41	1.57	0.95	0.01	5.17
April	2.67	0.33	4.05	2.50	0.12	11.86
July	2.07	-0.30	2.22	1.89	0.00	8.85

3.4 Comparison to Reference Station

In an effort to give some context to the magnitude of these errors, the deviations of the model results from near surface temperature observations were compared to the deviation of the airport reference station from all other stations. In effect, we wanted to test how well the model performed relative to the naïve use of nearby airport weather station data in predicting urban conditions (Fig. 7).

The median of the absolute model error was only lower in February and the spread of that error was only significantly smaller in July. Therefore, the model failed both in terms of accuracy and precision when compared to the airport reference station.

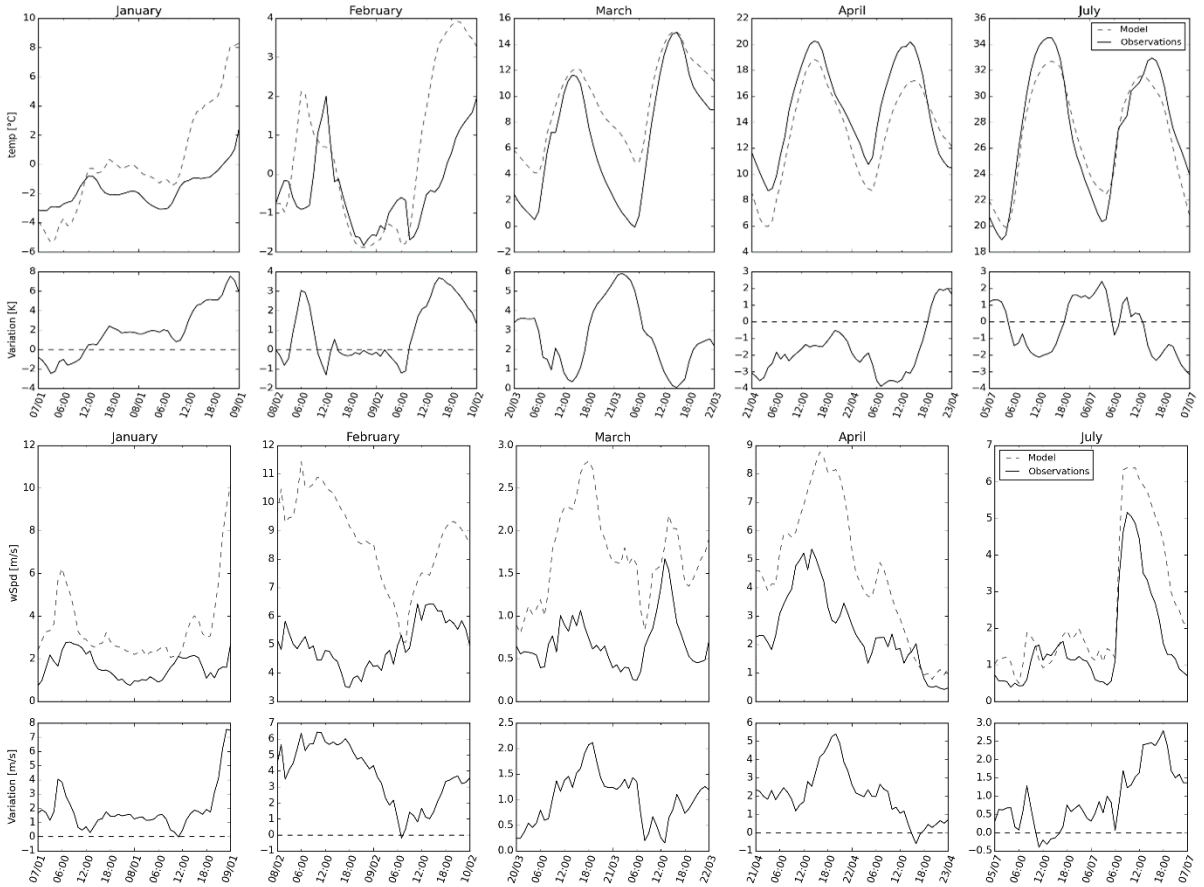

Fig. 6a, 6b – Model vs. Observations for 2 m Temperature (top) and Wind Speed (bottom)

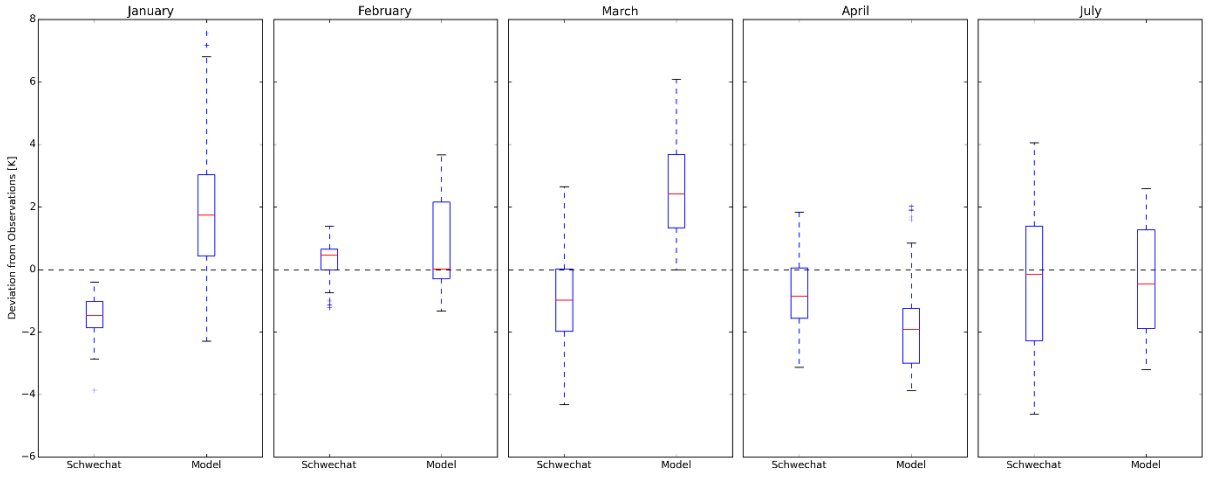

Fig. 7 – Distribution of deviation from 2 m temperature observations for airport reference station and WRF model

4. Discussion

4.1 Suitability for Building Energy Model Boundary Condition

Despite the use of a detailed description of the urban areas, core climatic variables, such as air temperature, were inconsistently reproduced. In fact, as Jänicke et al. (2016) show in their study over the Berlin-Brandenburg region, increasing complex multilayered UCMs might actually be a source of increased error when compared to a simple slab model.

With such a large degree of error and no consistency in the direction of bias that is both seasonally and diurnally stable, the present study cannot identify the WRF model as an appropriate tool for deriving urban boundary conditions for building energy modeling. Furthermore, the model more often provided worse estimates of the urban near surface temperature than the airport reference station. This suggests that morphing approaches that modify reference measurements or TMY records may be more suitable for urban studies.

4.2 Sources of Error

In order to better isolate the sources of modeling error, a correlation analysis was conducted between the overall RMSE per station and station properties that could be contributing to the observed error. Only the station elevation showed any significant correlation with the error with a correlation coefficient of 0.28. It is a weak relationship and may be related more to the land cover typology than elevation as in Vienna, the urban density decreases with elevation to the West and we have seen a higher rate of error associated with lower density LCZCs (e.g. LCZ 6 and 9) in another ongoing research effort.

Due to the use of unofficial personal weather station data, it was also of interest to examine the stations that showed the best and worst agreement with model results. Interestingly, we saw no consistency between study periods or within types of stations (i.e. official vs. amateur). This indicates that any consistent measurement error that might exist is obscured by the magnitude and spatial variation of the modeling error.

References

Barnaby, C.S., D.B. Crawley. 2011. "Weather Data for Building Performance Simulation". In: *Building Performance Simulation for Design and Operation*. Abingdon, UK: Taylor & Francis Group.

Bechtel, B., P.J. Alexander, J. Böhner, J. Ching, O. Conrad, J. Feddema, G. Mills, L. See, I. Stewart. 2015. "Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities". *ISPRS International Journal of Geo-Information* 4(1): 199–219. doi:10.3390/ijgi4010199.

Bhandari, M., S. Shrestha, J. New. 2012. "Evaluation of Weather Datasets for Building Energy Simulation". *Energy and Buildings* 49: 109–18. doi:10.1016/j.enbuild.2012.01.033.

Brousse, O., A. Martilli, M. Foley, G. Mills, B. Bechtel. 2016. "WUDAPT, an Efficient Land Use Producing Data Tool for Mesoscale Models? Integration of Urban LCZ in WRF over Madrid". *Urban Climate* 17: 116–34. doi: 10.1016/j.uclim.2016.04.001.

Burian, S.J., W.S. Han, M.J. Brown. 2003. "Morphological Analyses Using 3D Building Databases: Oklahoma City, Oklahoma". Los Alamos, U.S.A.: LA-UR, Los Alamos National Laboratory.

Chan, A.L.S. 2011. "Developing Future Hourly Weather Files for Studying the Impact of Climate Change on Building Energy Performance in Hong Kong". *Energy and Buildings* 43(10): 2860–68. doi:10.1016/j.enbuild.2011.07.003.

Fallmann, J., S. Emeis, P. Suppan. 2014. "Mitigation of Urban Heat Stress–a Modelling Case Study for the Area of Stuttgart". *DIE ERDE–Journal of the Geographical Society of Berlin* 144 (3-4): 202–16.

Gordon, J.H. 1921. "Temperature Survey of the Salt River Valley, Arizona". *Monthly Weather Review* 49(5): 271–74. doi:10.1175/1520-0493(1921)49<271:TSOTSR>2.0.CO;2.

Hebbert, M., V. Jankovic. 2013. "Cities and Climate Change: The Precedents and Why They Matter". *Urban Studies* 50(7): 1332–47. doi:10.1177/0042098013480970.

Howard, L. 1833. *The Climate of London: Deduced from Meteorological Observations Made in the Metropolis and at Various Places around It*. Vol. 2. London, UK: Harvey and Darton, J. and A. Arch, Longman, Hatchard, S. Highley and R. Hunter.

Jänicke, B., F. Meier, D. Fenner, U. Fehrenbach, A. Holtmann, D. Scherer. 2016. "Urban–rural Differences in near-Surface Air Temperature as Resolved by the Central Europe Refined Analysis (CER): Sensitivity to Planetary Boundary Layer Schemes and Urban Canopy Models". *International Journal of Climatology*, August. doi:10.1002/joc.4835.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, F. Rubel. 2006. "World Map of the Köppen-Geiger Climate Classification Updated". *Meteorologische Zeitschrift* 15(3): 259–63.

Lukacsy, M., C. Fendt. 2015. "Vienna in Figures". Vienna: Vienna City Administration MA23. Accessed on December 15th. https://www.wien.gv.at/statistik/pdf/viennainfigures.pdf.

Martilli, A. 2002. "Numerical Study of Urban Impact on Boundary Layer Structure: Sensitivity to Wind Speed, Urban Morphology, and Rural Soil Moisture". *Journal of Applied Meteorology* 41(12): 1247–66. doi:10.1175/1520-0450(2002)041<1247:NSOUIO>2.0.CO;2.

Meier, F., D. Fenner, T. Grassmann, B. Jänicke, M. Otto, D. Scherer. 2015. "Challenges and Benefits from Crowd-Sourced Atmospheric Data for Urban Climate Research Using Berlin, Germany, as Testbed". In: *Proceedings of ICUC9 - 9th International Conference on Urban Climate jointly with 12th Symposium on the Urban Environment.* Toulouse, France.

"Open Government in Vienna" 2016. *Für Eine Offene Stadt*. Accessed February 9. https://open.wien.gv.at/

Pernigotto, G., A. Prada, D. Cóstola, A. Gasparella, J.L.M. Hensen. 2014. "Multi-Year and Reference Year Weather Data for Building Energy Labelling in North Italy Climates". *Energy and Buildings* 72 :62–72. doi:10.1016/j.enbuild.2013.12.012.

Quantum GIS Development Team. 2016. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Radhi, H. 2009. "A Comparison of the Accuracy of Building Energy Analysis in Bahrain Using Data from Different Weather Periods". *Renewable Energy* 34(3): 869–75. doi:10.1016/j.renene.2008.06.008

Salamanca, F., A. Krpo, A. Martilli, A. Clappier. 2010. "A New Building Energy Model Coupled with an Urban Canopy Parameterization for Urban Climate Simulations—part I. Formulation, Verification, and Sensitivity Analysis of the Model". *Theoretical and Applied Climatology* 99(3-4): 331–44. Doi: 10.1007/s00704-009-0142-9.

Stewart, I.D., T.R. Oke. 2012. "Local Climate Zones for Urban Temperature Studies". *Bulletin of the American Meteorological Society* 93(12): 1879–1900. doi:10.1175/BAMS-D-11-00019.1.