Faculty of Science

Impact of Different Land Uses on Local Climate in the Stockholm Region

Daniel Magnusson

Degree of Bachelor of Science
with a major in Earth Sciences
15 hec

Department of Earth Sciences
University of Gothenburg
2023 B-1222

UNIVERSITY OF GOTHENBURG

UNIVERSITY OF GOTHENBURG
Department of Earth Sciences
Geovetarcentrum/Earth Science Centre

Impact of Different Land Uses on Local Climate in the Stockholm Region

Daniel Magnusson

ISSN 1400-3821

B1222
Bachelor of Science thesis
Göteborg 2023

Mailing address
Geovetarcentrum
S 405 30 Göteborg

Address
Geovetarcentrum
Guldhedsgatan 5A

Telephone
031-786 19 56

Geovetarcentrum
Göteborg University
S-405 30 Göteborg
SWEDEN

Abstract

More than half of the world's population lives in urban areas as a result of the ongoing urbanization. Therefore, it is important to understand the characteristics of urban climate, such as the urban heat island (UHI). To do this, well performing, high resolution climate models are needed. This study evaluates how well the Weather Research and Forecasting (WRF) model can reproduce summer 2 m air temperature (T2m) of 2018 over cities with various sizes and areas with different land use in terms of urban/non-urban over the Stockholm region. Six WRF simulations were carried out with land use classification from Local Climate Zones (LCZ). To evaluate the performance of the model, observation data from eleven Swedish Meteorological and Hydrological Institute (SMHI) weather stations in the area were used. The location of the stations was classified as urban or rural, where the urban stations were divided into city size groups based on the number of inhabitants in the city of the urban station. It was found that the highest correlation was between the simulations with revised urban parameters and the observation for the urban areas with a correlation coefficient between 0.90-0.91. However, an overestimation of the modelled daytime T2m occurred for all city size groups where the overestimation decreased with increased city size. The study concludes that WRF reproduces T2m more accurately for urban areas compared to rural areas and that the model is able to capture the UHI effect for medium to large cities. The modelled overestimation of the T2m might be because of a mismatch in land use, however, further investigation is needed to verify the cause.

Sammanfattning

Som ett resultat av den pågående urbaniseringen bor mer än hälften av världens befolkning i städer. Det är därmed viktigt att förstå egenskaperna hos urbana klimat så som urbana värmeöar. För att lyckas med detta behövs välfungerande och högupplösta klimatmodeller. Den här studien syftar till att utvärdera hur väl Weather Research and Forecasting (WRF) modellen kan återskapa sommartemperaturer 2018 i städer av olika storlekar och över områden med olika markanvändning med fokus på urban/icke-urban markanvändning i Stockholmsregionen. Sex WRF simuleringar genomfördes med markanvändningsklassificering från Local Climate Zones (LCZ). För att utvärdera modellens prestationsförmåga användes observationsdata från elva väderstationer från SMHI. Placeringen av väderstationerna klassificerades som urban eller icke-urban där de urbana väderstationerna indelades i stadsgrupper baserat på antalet invånare i staden där stationen var placerad. Den högsta korrelationen hittades mellan simuleringarna med reviderade stadsparametrar och observationen för de urbana platserna med en korrelation koefficient på 0,90-0,91. En överestimering påvisades dock för den dagliga temperaturen för alla stadsgrupper där överestimeringen minskade med ökad stadsstorlek. Studien drar som slutsats att modellen återskapar temperatur bättre över urbana områden jämfört med icke-urbana områden och att den kan fånga effekten av den urbana värmeön för mellanstora och stora städer. Modellens överestimering av temperaturen kan bero på en felaktig klassificering av markanvändning. Vidare undersökning behövs dock för att verifiera orsaken till överestimeringen.

Table of Contents

1. Introduction 1
1.1 Background 1
1.2 Aim and Research Questions 2
2. Data and Methods 3
2.1 Site Description 3
2.2 Weather Research and Forecasting Model 3
2.3 Observation Data 6
2.4 Analysis 6
2.4.1 Evaluation of the WRF Model 6
2.4.2 WRF and City Size 7
3. Result 8
3.1 Performance of the WRF Model 8
3.2 WRF and City Size 10
4. Discussion 13
4.1 Evaluation of the WRF Model 13
4.2 WRF and City Size 14
4.3 Improvement of the WRF Model 15
5. Conclusion 17
6. Acknowledgments 18
7. Appendix A 19
8. References 21

1. Introduction

1.1 Background

In 2018, 55 % of the world's population lived in urban areas and this number is projected to increase in the future with the ongoing urbanization (United Nations. Department of Economic and Social Affairs, 2019). When more people are living in cities, construction of new houses and buildings is needed to support the inhabitants with their housing needs. The impact of urbanization on the spatial structure of cities can be discerned in mainly two ways, expansion and densification (Ustaoglu & Jacobs-Crisioni, 2022). Expansion occurs through a land use change from rural to urban when surrounding, previously non-built-up, areas are used to construct new buildings. Densification, on the other hand, occurs when more buildings are built within already existing urban areas (Broitman & Koomen, 2015). Urbanization, such as in Sweden, is characterized by a shift in land use from vegetation to built-up areas in order support inhabitants with their housing needs (Amorim et al., 2020).

When a larger proportion of the inhabitants is living in cities, it is of great importance to understand city climate and its impact on human life and well-being. One phenomenon associated with urban climate is the urban heat island (UHI), where the temperature is higher in cities compared to the surrounding rural areas. There are several factors affecting the UHI, such as changes in land use, urban geometry, and background climate (Zhao et al., 2014). Vegetated and green areas in cities have a cooling effect on the urban climate and decrease the UHI effect, which makes it more favourable for people living in the city (Steeneveld et al., 2011). When vegetated areas are transformed into built-up areas, there is a loss of evaporative cooling and the UHI will be aggravated (Zhao et al., 2014). Artificial materials, such as concrete and asphalt, can store a larger amount of energy from solar radiation compared to natural surfaces like vegetation and soil. The absorbed energy during daytime is released at night which contributes to the nocturnal UHI (Zhao et al., 2014).

The urban geometry can impact city climate in several ways. In highly built-up areas with reflecting facades, the solar radiation can be reflected and bounce between buildings and the solar radiation can reach the ground even though the buildings provide shade at pedestrian-level (Wang et al., 2021). The solar energy will then be stored in the buildings through multiple absorption which contribute to an increase in urban temperature. Urban geometry does also have an impact on wind characteristics in cities. When wind are flowing from the open rural areas towards a city, the urban geometry increase the roughness which alter the wind direction to an up-ward direction (Arkon & Özkol, 2014). This decreases the ground-level wind speed in the urban area. When the urban wind speed is decreased, the warm air masses at pedestrian-level is not replaced with new air, and the UHI exacerbates (Kim & Baik, 2004).

With the ongoing climate change, the temperature is projected to increase and extreme events, such as heat waves, are expected to occur more frequently and with higher intensity (Hoegh-Guldberg et al., 2018). This together with intense UHI effect due to larger and denser cities can lead to health issues for the inhabitants in the cities. Heat related health issues in urban areas, such as heat stress, can lead to heat related mortality, especially for children and elderly people (Lundgren Kownacki et al., 2019).

It is important to understand urban climate and its relation to urban geometry to be able to make proper city plans for the growing cities which provides thermal comfort for people so they can live there with relatively less heat related health issues. One key to get more understanding of urban climate and to predict UHI effects in cities is to have high performing and accurate climate models (Chen et al., 2022). These climate models can be used to test different urban planning scenarios with different projection of the future climate to predict future urban climates.

A previous study by Amorim et al. (2020) investigated the impact of land use changes from vegetation to construction on summer temperature in Stockholm. It was found that for two urban planning scenarios, the temperature in the built-up areas would increase by, on average, 0.29 °C and 0.46 °C in 2030 and 2050, respectively. Steeneveld et al. (2011) studied UHI in cities of variable sizes in the Netherlands and found that the 95th percentile of the UHI was well correlated with population density. Furthermore, Zhou et al. (2017) investigated UHI in relation to city size for cities in Europe and found that UHI intensity is positively correlated with city size. Ribeiro et al. (2021) studied how well the Weather Research and Forecasting (WRF) model could reproduce different meteorological features over the urban region of Barcelona, Spain, and found that the simulation BEP (building energy parameterization) had the highest correlation with the meteorological observations, such as temperature.

1.2 Aim and Research Questions

The overall aim of this study is to quantify the impact of urbanization on the local climate. To do this, a model that can reproduce local climate well is needed. Therefore, this study will investigate how well the Weather Research and Forecasting (WRF) model version 4.4 (Skamarock et al., 2021) can reproduce summer temperature of 2018 in cities with various sizes and in areas of different land use, in terms of urban/rural land use, in the Stockholm region. This will be done by answering the following questions:

1. How well can the regional climate model WRF reproduce summer temperature of cities with various sizes?
2. How big differences in temperature could be for different land use in terms of urban/rural?
3. What can be done to improve the WRF model over the Stockholm region?

2. Data and Methods

2.1 Site Description

The target region for this study is the Stockholm region, the area in the vicinity of the lake Mälaren, in eastern Sweden (Figure 1). There are several larger cities in the region, such as Stockholm, Uppsala, and Eskilstuna. The Stockholm region is a relatively densely populated region in Sweden where more than 30 % of the inhabitants of Sweden live (SCB, 2023). The climate is classified as Dfb in the Köppen-Geiger climate classification which is characterised by a cold climate, without dry season, and with warm summers (Peel et al., 2007). There are two large waterbodies in and around the region, the Baltic Sea to the east and the lake Mälaren in the central parts.

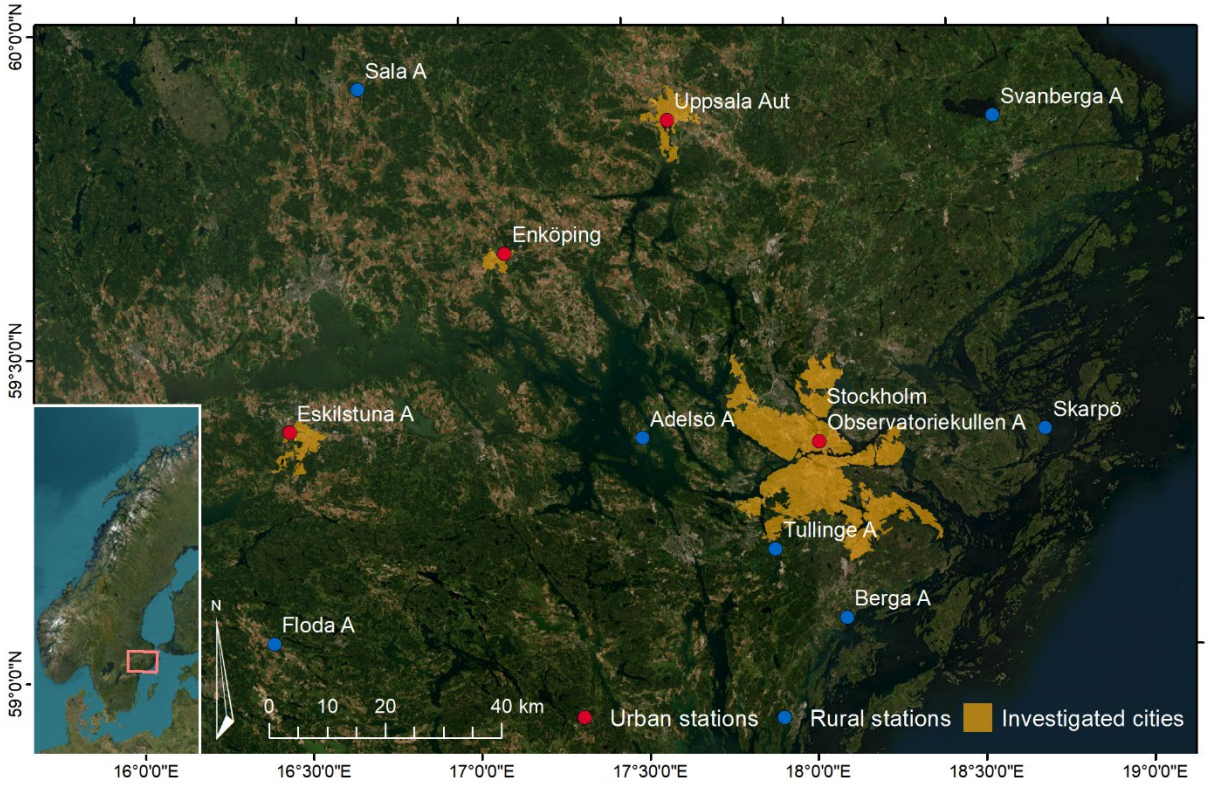

Figure 1. Map over the target region. The coloured dots indicate the location of the weather stations used in this study, red for urban stations and blue for rural stations. The orange areas indicate the cities (from left to right) Eskilstuna, Enköping, Uppsala, and Stockholm.

2.2 Weather Research and Forecasting Model

For this study, the open-source community model Weather Research and Forecasting (WRF) version 4.4 will be used (Skamarock et al., 2021). It is an atmospheric modelling system created for scientific and numerical weather prediction purposes. The WRF model uses fundamental physical laws for high resolution modelling that reproduces climate variables with high spatial resolution (Skamarock et al., 2021). This model has been widely used to investigate urban climates and to reproduce and predict UHI effects in cities (Chen et al., 2022; Jahangir & Moghim, 2019; Ribeiro et al., 2021; Vinayak et al., 2022).

The WRF model is configured with three nested domains with the outer domain having a resolution of 9 km (Figure 2). The region which this study is focusing on is relatively small,

with an area of 31 500 km^2. Therefore, it is of great importance to have modelled temperature data with high spatial resolution to be able to investigate the climate at city level. Hence, a dynamical downscaling of the WRF model was done to get a higher resolution of the modelled output. In this study, the inner modelling domain with a spatial resolution of 1 × 1 km was used, consisting of 150 × 210 grid cells (Figure 2). The model output contained data with an hourly temporal resolution for the months June, July, and August 2018.

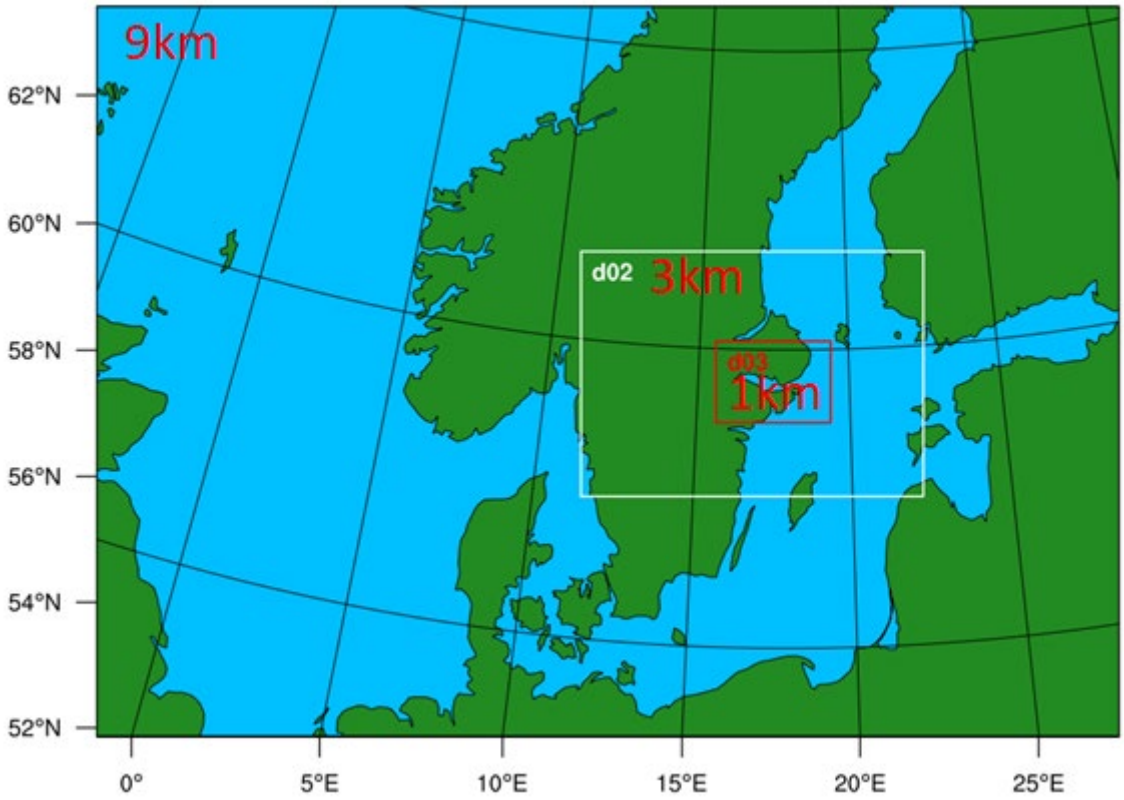

Figure 2. WRF model domain with the horizontal resolution for each domain indicated. The red rectangle indicates the spatial extent of the model domain used in this study.

In the WRF, the RRTMG Shortwave and Longwave Schemes (Iacono et al., 2008), Thompson Microphysics Scheme (Thompson et al., 2008), Noah–MP Land Surface Model (Niu et al., 2011), Eta Similarity Surface Layer Scheme (Janjić, 1996), and Mellor–Yamada–Janjic Scheme (MYJ) Planetary Boundary Layer (PBL) (Janjić, 1994) schemes were used in the simulations. The initial and boundary conditions, including sea surface temperature, were provided by ERA5 (Hersbach et al., 2020), in which the boundary condition is updated every 3 hours.

The urban parameters used in the WRF model was extracted from the 100 m resolution global maps of Local Climate Zones (LCZs), generated by Demuzere et al. (2022b) based on the classification system of LCZ developed by Stewart and Oke (2012). To convert LCZs over the study region to urban parameters for WRF, the python-based W2W package was used developed by Demuzere et al. (2022a). The default urban parameters used in W2W are listed in Table 1, where the built types are based on the study by Stewart and Oke (2012).

Table 1. Default LCZ-based urban canopy parameters.

Built Types	Urban Parameters					
	FRC	MH_MIN	MH	MH_MAX	BLDFR	H2W
1	0.95	25.00	50.00	75.00	0.50	2.50
2	0.90	10.00	17.50	25.00	0.55	1.25
3	0.85	3.00	6.50	10.00	0.55	1.25
4	0.65	25.00	50.00	75.00	0.30	1.00
5	0.70	10.00	17.50	25.00	0.30	0.50
6	0.60	3.00	6.50	10.00	0.30	0.50
7	0.85	4.00	5.00	6.00	0.75	1.50
8	0.85	3.00	6.50	10.00	0.40	0.20
9	0.30	3.00	6.50	10.00	0.15	0.15
10	0.55	5.00	10.00	15.00	0.25	0.35

Urban parameters: FRC: urban fraction field; MH_MIN: minimum building height; MH: mean building height; MH_MAX: maximum building height; BLDFR: building surface fraction; H2W: height to width ratio.

However, the default urban parameters (Table 1) are not representative for the cities in the Stockholm region as the building height is higher in the default settings compared to the buildings in the Stockholm region. Therefore, a revision of the statistics in the urban parameters were done to get a building height that mimics the actual building height in the region (Table 2). This revision of the urban parameters was done using the building height from the Global Human Settlement (GHS) (Schiavina et al., 2022) with the methodology established by Pesaresi et al. (2021).

Table 2. Revised LCZ-based urban canopy parameters.

Built Types	Urban Parameters					
	FRC	MH_MIN	MH	MH_MAX	BLDFR	H2W
1	0.95	4.00	18.00	40.00	0.50	2.00
2	0.90	7.00	16.00	30.50	0.55	1.25
3	0.85	2.50	7.00	17.00	0.55	1.25
4	0.65	2.50	11.30	27.20	0.30	1.00
5	0.70	2.50	7.70	23.00	0.30	0.50
6	0.60	2.50	4.00	17.50	0.30	0.50
7	0.85	2.50	3.50	6.00	0.75	1.50
8	0.85	2.50	6.60	20.00	0.40	0.20
9	0.30	2.50	2.80	4.00	0.15	0.15
10	0.55	3.00	14.00	28.00	0.25	0.35

Urban parameters: FRC: urban fraction field; MH_MIN: minimum building height; MH: mean building height; MH_MAX: maximum building height; BLDFR: building surface fraction; H2W: height to width ratio.

In the WRF model, six simulations with different settings and parameters were integrated: (1) Building Energy Model (BEM), (2) BEM with a lake simulation being coupled to the WRF simulation (BEMLake), (3) Building Effect Parameterization (BEP) (Martilli et al., 2002), (4) all urban land use have been removed (NoUrban), (5) BEM with revised building height (BEMLowBH), and (6) BEP with revised building height (BEPLowBH). For the simulations BEM, BEMLake and BEP, the default urban parameters (Table 1) were used to convert the WRF urban parameters from LCZs. For the NoUrban simulation, the urban land use was replaced with the dominant adjacent non-urban land use type and neither BEM nor BEP were included in this simulation. The NoUrban simulation was included in the investigation to test the possible impact of urbanization on the local climate. For BEMlowBH and BEPlowBH, the revised urban parameters (Table 2) were used to convert the WRF urban parameters from LCZs.

2.3 Observation Data

To be able to validate the model outputs, observation data from the investigated region is needed. The observational data used in this study was retrieved from SMHI's weather stations (https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer/#param=airtemperatureInstant,stations=core) and consisted of 2 m mean surface air temperature (T2m) with an hourly resolution. Not all available SMHI stations in the area with data for 2018 was used. The stations of Södertälje and Västerås were not included as they only had two measured values per day, which is insufficient for evaluating the diurnal cycle of T2m. The weather stations located at an airport were also excluded as these areas are not representative for a typical urban or rural area. Data from a total of eleven weather stations was used, seven stations in rural areas and four stations in urban areas. To classify if the weather station were located in an urban or rural area, the stations were displayed in ArcMap along with a shape-layer containing urban areas downloaded from Statistics Sweden (https://www.scb.se/vara-tjanster/oppna-data/oppna-geodata/tatorter/). Where the location of the stations overlapped the urban area, it was classified as an urban station, and where the station was outside the urban area, it was classified as a rural station. The location of the SMHI's weather stations used in this study can be seen Figure 1 and the classification of the stations can be seen in Table 3.

Table 3. SMHI weather stations used in this study, with percentage of missing values.

Station name	Station id	Longitude	Latitude	Urban (U) / Rural (R) site	Missing values (%)
Floda A	96040	16.3944	59.0558	R	6.4
Eskilstuna A	96190	16.4526	59.3821	U	0.3
Tullinge A	97100	16.6001	59.5972	R	0.1
Adelsö A	97280	17.9093	59.1785	R	0.1
Enkoping	97370	17.6291	59.2141	U	1.6
Uppsala Aut	97510	17.1200	59.6500	U	1.9
Berga A	98040	17.6320	59.8471	R	9.2
Skarpö	98160	17.5930	59.9000	R	0.0
Svanberga A	98490	18.1150	59.0680	R	2.4
Stockholm-Observatoriekullen A	98230	18.6313	59.8317	U	0.0
Sala A	96560	19.5022	59.4423	R	0.0

For all stations, except from Skarpö, Stockholm-Observatoriekullen A, and Sala A, missing values were found in the observation data for the summer months of 2018. The percentage of missing values for the stations ranged from 0-9.2 % (Table 3). To account for the missing values, two methods were used depending on whether one or more consecutive missing values were found. If one consecutive missing value was found, an average of the preceding and following values was used to interpolate the missing value. If several consecutive missing values were found, an average of each hour was computed for the entire dataset, where the missing values were replaced with the respective hourly value.

2.4 Analysis

2.4.1 Evaluation of the WRF Model

To evaluate how well the WRF model reproduces summer temperature over the investigated region, a correlation analysis was conducted. First, the corresponding grid cell of the location

of each SMHI weather station was identified in the modelled grid data. The modelled data for these grid cells were then extracted and correlated with the data for the corresponding weather station, this was done for all six simulations. Additionally, the stations were grouped based on their location into urban and rural, where the correlation were computed for the two groups to evaluate how well the WRF model can reproduce urban T2m compared to rural T2m. The statistics were plotted into three Taylor diagrams for a better visualization of the correlation coefficient, standard deviation, and the centred root-mean-square error (CRMSE), one for all locations, one for the urban locations, and one for the rural locations. The evaluation of the model focused on all simulations except for the NoUrban. The NoUrban simulation was used to test the possible impact of urbanization on the local climate as it does not have representative land use compared to the actual land use in the region. To further investigate how well the model is able to reproduce T2m, hourly mean values were computed for all simulations and the observation respectively to get the mean T2m diurnal cycle for each location.

2.4.2 WRF and City Size

The cities investigated in the current study are of various sizes, both when it comes to spatial extent and number of inhabitants. A compilation of city size, population, and population density for the investigated cities can be seen in Table 4 with statistics from Statistics Sweden (https://www.scb.se/vara-tjanster/oppna-data/oppna-geodata/tatorter/). To investigate how well the WRF model can reproduce T2m over cities with various sizes, the location of each weather station was classified into one of four groups based on population in the city: (1) rural, (2) small-sized city, (3) medium-sized city, and (4) large city. The city classification used was from the Swedish Municipality Division (Sveriges Kommuner och Regioner, 2022) which classified a city as a small-sized city if the population was 15 000–39 999, medium-sized city if the population was 40 000–199 999, and large city if the population ≥200 000. The locations where the population <15 000 were classified as rural. The classification of the cities can be seen in Table 4. For each city size group, the average T2m for the different simulations and observation were calculated to quantify the magnitude of deviation between the simulations and observation for the respective city size group.

Table 4. City area, population, population density and city size classification for the investigated cities with an urban weather station.

City	City area (km2)	Population	Population density	City size group
Enköping	12.0	25 071	2091.6	Small
Eskilstuna	29.6	69 948	2360.4	Medium
Stockholm	433.4	1 617 407	3732.3	Large
Uppsala	43.6	166 698	3822.9	Medium

To get an overview of the modelled data over the region, the mean T2m were calculated for each grid cell for all simulations. Additionally, the difference between the mean T2m for NoUrban and BEPlowBH were computed to see if BEPlowBH were able to capture the UHI effect over the investigated cities. The simulation BEPlowBH was chosen because it has shown that simulations using BEP has high correlation with meteorological observations in previous studies, such as Ribeiro et al. (2021).

3. Result

3.1 Performance of the WRF Model

The Taylor Diagrams (Figure 3) are showing the performance level of the simulations, Figure 3a for all locations, Figure 3b for the urban locations, and Figure 3c for the rural locations. When including all locations, the simulations BEMlowBH, BEPlowBH, BEM, and BEMLake had a similar level of performance, where BEMlowBH had the best performance with a correlation coefficient of 0.71 and a CRMSE of 0.39 (Table A1, in Appendix A). The BEP simulation showed a slightly lower performance than the simulations previously mentioned. The CRMSE of the simulations ranged from 0.39-0.46 (Table A1). For more details of the performance statistics, see Table A1 in Appendix A.

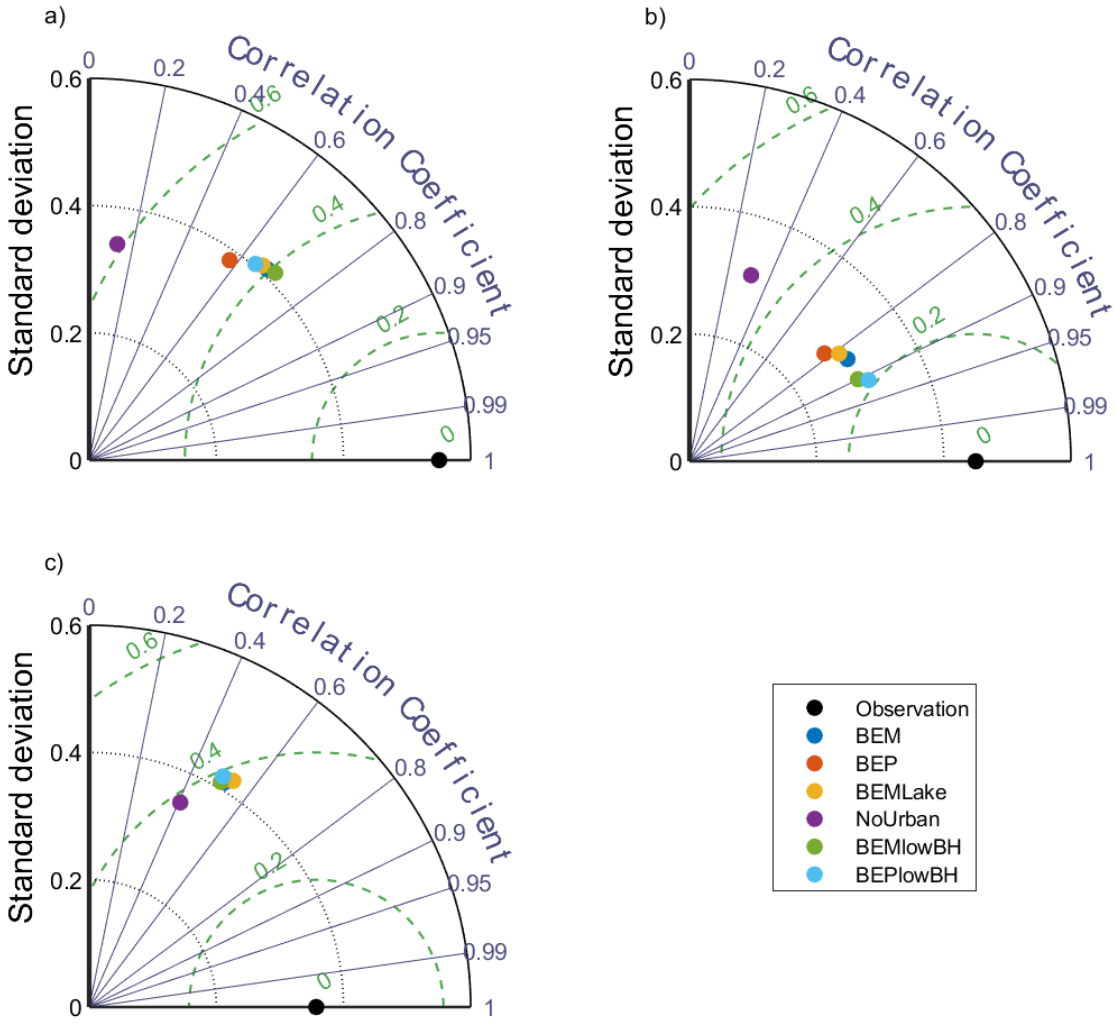

Figure 3. Taylor diagrams showing the performance of the six simulations for the WRF model compared to the observation. The green dashed lines indicate centred root mean square error. (a) all locations, (b) urban locations, and (c) rural locations.

BEMlowBH and BEPlowBH showed the highest performance level for the urban locations (Figure 3b), with a correlation coefficient of 0.90 and 0.91 along with a CRMSE of 0.23 and 0.21, respectively (Table A2, in Appendix A). The simulations BEM, BEP, and BEMLake showed a slightly lower performance with a correlation coefficient ranging from 0.78-0.84 with a CRMSE of 0.26-0.29. A more detailed list of the performance statistics for the urban locations

is presented Table A2 in Appendix A. For the rural locations (Figure 3c), the simulations BEM, BEP, BEMLake, BEMlowBH, and BEPlowBH had a similar performance level with a correlation coefficient of 0.50-0.54 and a CRMSE of 0.38-0.39. More details of the performance statistics of the rural locations can be seen in Table A3 in Appendix A.

The diurnal cycle for each location shows that the simulations can reproduce T2m with various precision during the hours of the day (Figure 4, 5). All simulations have a similar diurnal cycle for the respective urban location (Figure 4). For the urban locations, an overestimation of the T2m occurred during daytime (9:00-18:00) compared to the observed T2m (Figure 4). During the morning hours (4:00-8:00), the model reproduced a T2m close to the observed for all simulations. At night-time (19:00-3:00), all simulations showed a too high T2m for Eskilstuna and Enköping (Figure 4a, 4b). For the two largest cities, Stockholm and Uppsala, the simulation BEP was able to reproduce a nocturnal T2m close to the observation, but for the remaining simulations, a slight nocturnal overestimation of T2m occurred (Figure 4c, 4d).

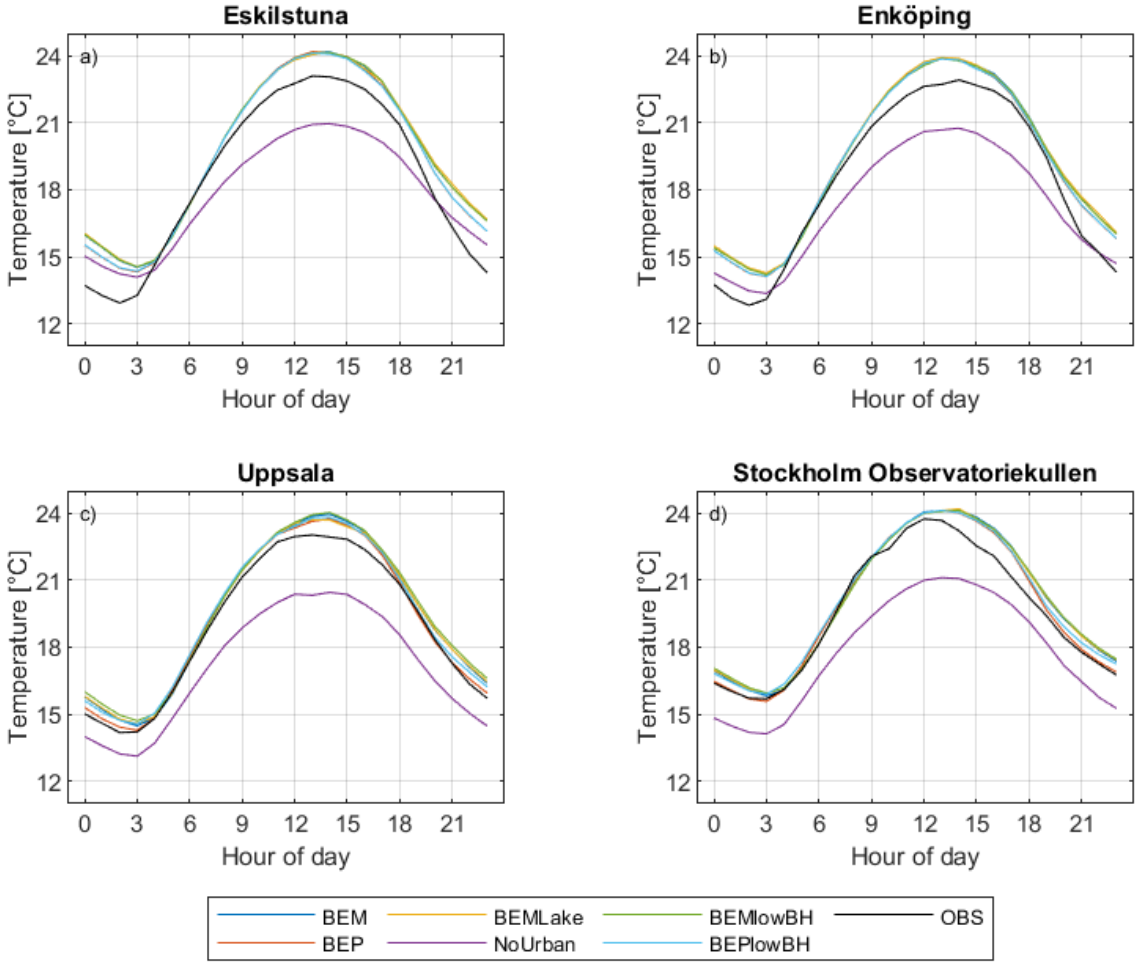

Figure 4. Diurnal cycle for the simulations and the observation for the four urban locations.

For the rural locations, the simulations showed an overestimation of the T2m during the day (9:00-18:00) (Figure 5). For the night (19:00-3:00), the simulations captured a higher T2m than the observed for all rural locations except from Skarpö, where all simulations except from NoUrban captured a T2m similar to the observation (Figure 5e). During the morning hours (4:00-8:00) for all rural locations, the simulations captured a T2m similar to the observation.

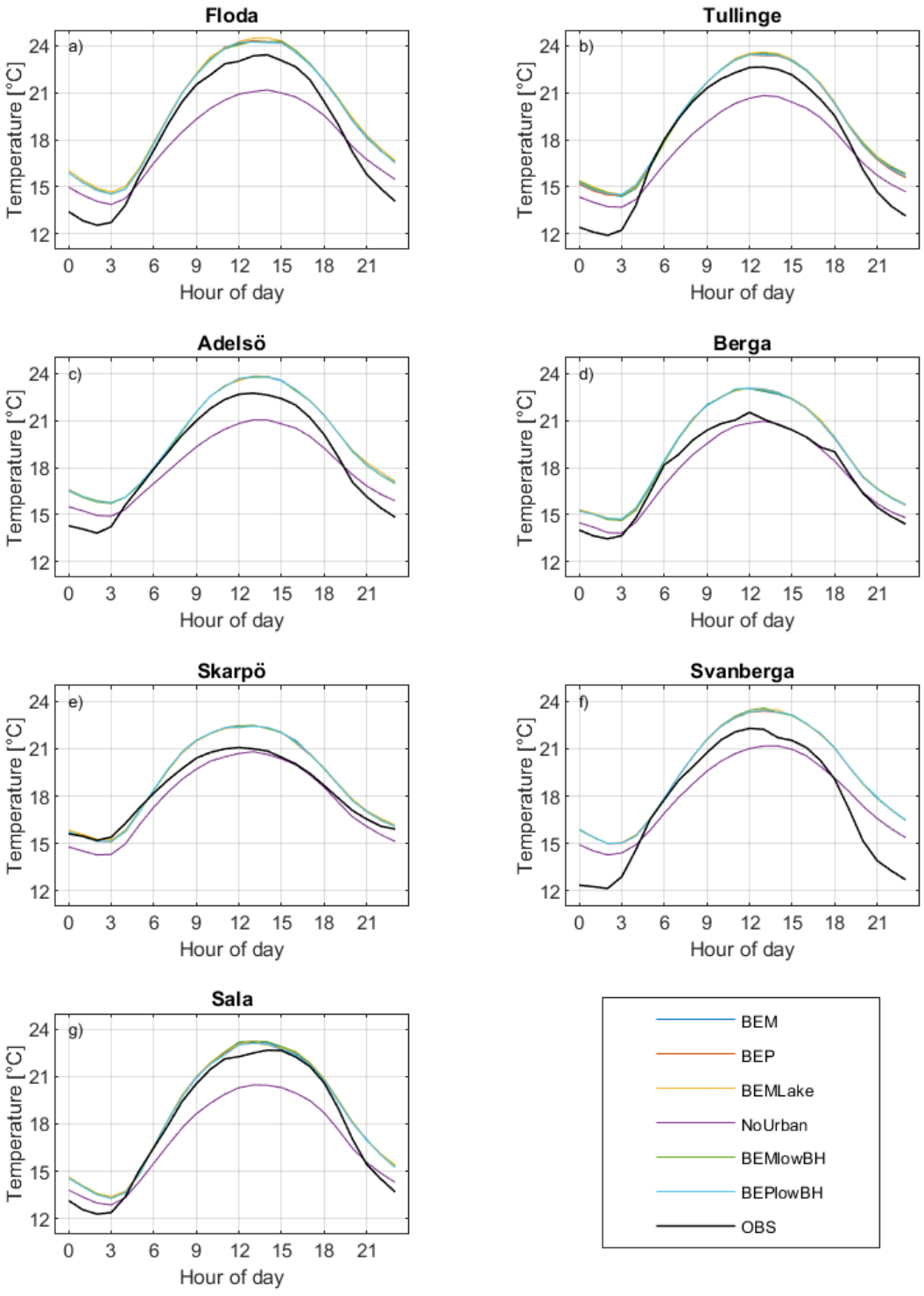

Figure 5. Diurnal cycle for the simulations and the observation for the seven rural locations.

3.2 WRF and City Size

When all investigated locations are divided into groups based on city size, all simulations show an overestimation of the mean T2m for all city size groups compared to the observation (Figure 6). The overestimation of the mean T2m is declining as the city size increases.

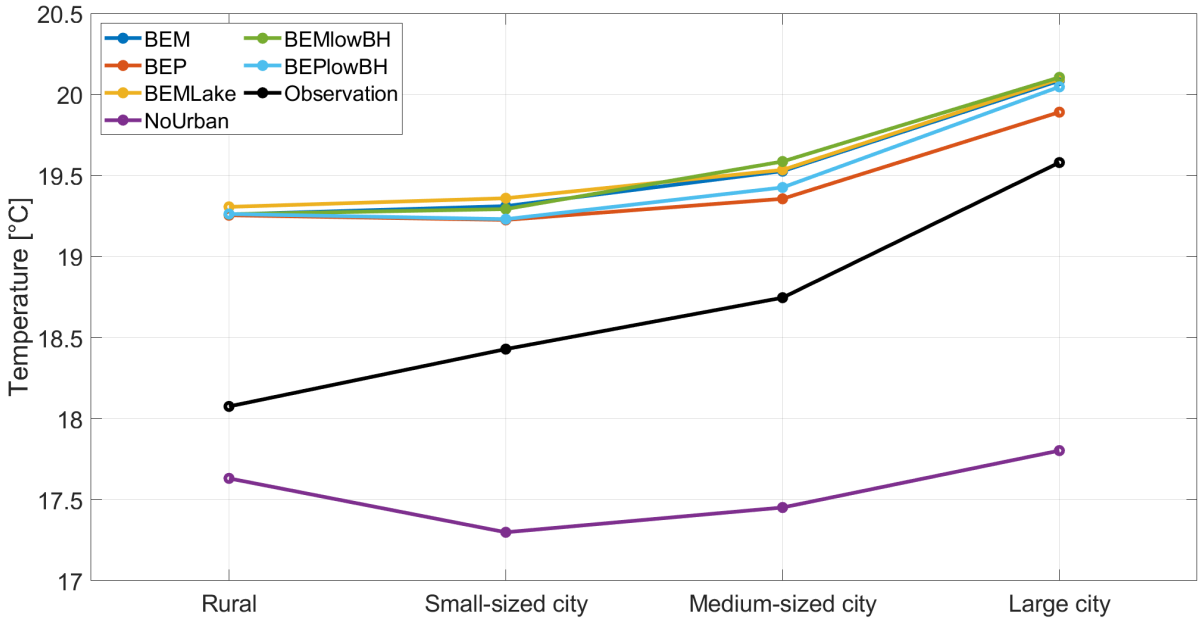

Figure 6. WRF simulated mean T2m for the simulations along with observed T2m based on city size.

For the rural locations, all WRF simulations, except for the NoUrban, showed a similar overestimation of the mean T2m ranging from 1.18-1.23 °C above the observation, where the NoUrban showed a mean T2m of 0.44 °C below the observation (Table 5). For the small-sized cities, the NoUrban showed a mean T2m of 1.13°C below the observation and the remaining simulations showed a mean T2m between 0.80-0.93 °C above the observation. For the medium-sized cities, the NoUrban showed a T2m of 1.29 °C below the observation and the remaining simulations showed a T2m between 0.61-0.84 °C above the observation. Lastly, for the large cities, the simulation BEP showed a mean T2m of 0.31 °C above the observation where the rest of the simulations, except for the NoUrban which showed a T2m of 1.78 °C below the observed, showed a mean T2m between 0.47-0.53 °C above the observation (Table 5).

Table 5. Modelled mean T2m deviation from observation based on city size.

Simulation	**Temperature deviation (°C) from observation based on city size group**			
	Rural	**Small-sized city**	**Medium-sized city**	**Large city**
BEM	1.18	0.88	0.78	0.50
BEP	1.18	0.80	0.61	0.31
BEMLake	1.23	0.93	0.79	0.51
NoUrban	-0.44	-1.13	-1.29	-1.78
BEMlowBH	1.19	0.86	0.84	0.53
BEPlowBH	1.19	0.80	0.68	0.47

To test the impact of urbanization, the simulations NoUrban was incorporated in the investigation. NoUrban showed an underestimation of the T2m in the diurnal cycles, mainly during daytime (6:00-18:00), compared to the observation (Figure 4, 5). An underestimation of the mean T2m for all city size groups compared to the observation is also seen for the NoUrban (Figure 6), where the underestimation is increasing with increasing city size. The underestimation from the observation ranged from -0.44 °C for the rural locations to -1.78 °C for the large city group (Table 5).

An overview of the mean T2m over the entire target area for NoUrban and BEPlowBH can be seen in Figure 7a and Figure 7b, respectively. The mean T2m ranges between 15-18 °C for NoUrban and 15-21 °C for BEPlowBH, with the exception for some areas in the western part of the region (blue areas) where the mean T2m ranged between 10-13 °C (Figure 7a, 7b). The difference between the mean T2m of the NoUrban and BEPlowBH (ΔT2m) can be seen in Figure 7c. The ΔT2m is ranging between 0-2.5 °C with the exception for the western part where some areas showed a ΔT2m >3 °C and others a ΔT2m<-1 °C. The rectangles (Figure 7c) indicate the three largest cities included in this investigation, from left to right, Eskilstuna, Uppsala, and Stockholm, where the ΔT2m over the cities is slightly higher compared to the surrounding areas. For an overview of the simulated mean T2m for all simulations, see Figure A1 in Appendix A.

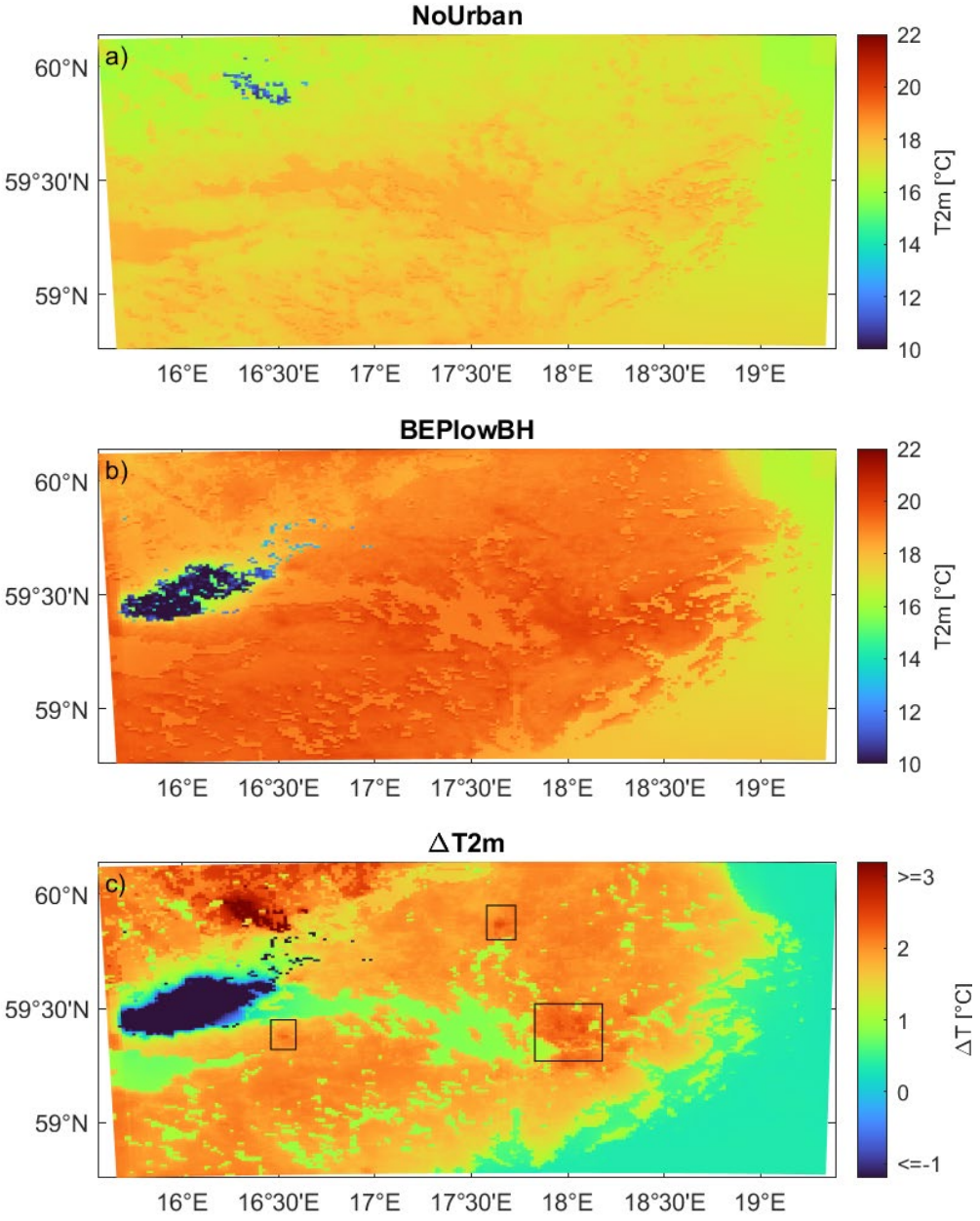

Figure 7. (a) Mean T2m for the NoUrban simulation. (b) Mean T2m for the BEPlowBH simulation. (c) Difference in mean T2m between the NoUrban and BEPlowBH simulations. The rectangles are indicating the cities (from left to right) Eskilstuna, Uppsala, and Stockholm.

4. Discussion

4.1 Evaluation of the WRF Model

When looking at the model's performance level presented in the Taylor diagrams (Figure 3), it is clear that the model does reproduce T2m for locations with urban land uses better compared to locations with non-urban land uses. The simulations with the best performance statistics were the ones with revised urban parameters (BEMlowBH and BEPlowBH) for the urban locations (Table A2, in Appendix A). Here, the importance of representative urban parameters for the investigated area is evident. When revising the urban parameters to match the actual building height in the area, the performance of the model increases. When setting the result in comparison to the study of Amorim et al. (2020), which found correlation coefficients between 0.93-0.99 for urban and rural locations, the performance level of this work show a lower performance level regarding correlation coefficient. For the urban locations, a slightly lower performance for BEMlowBH and BEPlowBH was received (r = 0.90-0.91), and for the rural locations, a significantly lower performance was received for all simulations (r = 0.41-0.54). Therefore, it can be concluded that WRF has difficulties to reproduce T2m of areas with non-urban land uses. However, WRF is able to reproduce T2m of areas with urban land uses well when comparing the result with other studies, such as Amorim et al. (2020).

All WRF simulations succeeded to reproduce a diurnal cycle close to the one of the observed (Figure 4, 5). For the largest cities, the model was able to reproduce a diurnal cycle close to the observation with high accuracy (Figure 4). However, an overestimation during the day and night occurred for the majority of the investigated locations (Figure 4, 5). This shows that the model has difficulties to reproduce T2m during both day and night. However, during the morning hours, all simulations showed a similar T2m to the one of the observed. In a study by Ribeiro et al. (2021), the WRF model with LCZ as land use input was used to reproduce temperature over the metropolitan area of Barcelona, Spain. A similar overestimation of the temperature during the day was observed for one of their investigated locations of what the diurnal cycles of the present study show. For this specific location, the authors found that the overestimation of the temperature was due to a mismatch of the land use classification. In this case, the specific grid cell got a land use class with a too low vegetation fraction which resulted in increased modelled temperature. Additionally, Steeneveld et al. (2011) and Amorim et al. (2020) found that a change in land use from vegetation to urban will locally increase the urban temperature. Thus, a mismatch of land use, with too low vegetation fraction, might be the case for the overestimation of the T2m during the day in the present study (Figure 4, 5). Further investigation is needed to see how well the LCZ land classification used in this study match the actual land cover to find out if this could be the case for the overestimation. Land use data with higher resolution could possibly reduce the overestimated T2m during the day as well.

For the location of Berga, the diurnal cycle of the observed T2m looks a bit pointy (Figure 5d) compared to the other observed diurnal cycles (Figure 5). This might be because there were 9.2 % missing values in the observed T2m which were interpolated. The interpolation of the missing values might in some cases have given a too high value and in other cases a too low value, which can have an impact on the T2m in the diurnal cycle. It can be good to have this in mind when looking at the diurnal cycles for the locations with missing values.

4.2 WRF and City Size

Overall, the WRF model were able to reproduce the mean T2m well for the various city size groups as the deviation from the observed ranged between –1.78-1.23 °C for the different simulations (Table 5). The simulation BEP were able to reproduce the mean T2m for the small, medium, and large-size groups with highest precision with a mean T2m of 0.80 °C, 0.61 °C, and 0.31 °C above the observed, respectively. Here it can be seen that the model reproduces mean T2m more accurately for cities of larger size. For the rural locations, the simulation NoUrban seems to give the best representation with a mean T2m of 0.44 °C below the observation (Table 5).

The result of the present study indicates that the WRF model can reproduce summer T2m with greater precision for large cities than for rural areas and small- to medium-sized cities (Figure 6). However, for all city size groups, the simulations did an overestimation of the T2m, except from the NoUrban which did an underestimation of the T2m. One thing to keep in mind is that this study only included four cities in the analysis, one small-sized city, two medium-sized cities and one large city. Further analysis based on a larger number of cities for each city size group would be needed to better evaluate how well the WRF model can reproduce T2m over cities of various sizes. However, the result of this study still gives an indication of how well the model can reproduce summer T2m over cities of various sizes despite the low number of cities included.

In this study, the simulations BEMlowBH and BEPlowBH, which had a more realistic representation of the actual building height, did not show any major differences in mean T2m compared to the BEM, BEP, and BEMLake simulations (Figure 6), despite the higher correlation (Figure 3). Therefore, the revised urban parameters (Table 2) did not make any significant difference in mean T2m compared to the default urban parameters (Table 1). This can raise the question about how building height impact the temperature in cities. Previous studies have shown to be conflicting when it comes to how building height impact the surface temperature in cities. Some shows that increased building height will increase the surface temperature (Rezaei Rad et al., 2019) and others show that increased building height will decrease the surface temperature (Danniswari et al., 2022; Zheng et al., 2019). The result of this study does not accept or reject any of the two ideas, but it shows that a better representation of the urban geometry gives higher correlation. Further research is needed to get a better understanding of how building height impact the surface temperature in the cities in the Stockholm region. It is important in a human comfort aspect to know how changes in building height impact the surface temperature in cities to be able to implement urban planning which reduces heat-related mortality in future heat waves.

For the observed mean T2m, there is an increase in T2m for every city size group and roughly a 1.5 °C temperature difference between the rural and the large city size group (Figure 6). This increase in T2m is interpreted to be the result of the UHI effect in the cities. This is in line with the result of Zhou et al. (2017), which found that the UHI intensity increases with increasing city size. In Figure 7c, it is visible that the model can capture the UHI effect as the T2m for the marked cities (the rectangles) shows a higher T2m compared to the surrounding area. However, when looking at how well the model can capture the UHI effect, the modelled mean T2m difference between the rural and the large city size group ranged from about 0.5 °C to 1.0°C for the different simulations (Figure 6), which is lower than for the observation. Additionally, the

model did not show a difference in T2m between the urban and small city size group and did therefore not capture the UHI effect for the small city size group at all (Figure 6). It can therefore be concluded that the model can capture the UHI effect for the medium and large city size group but not fully to the degree of the observed UHI effect.

The NoUrban simulation was used in this study to test the impact of urbanization on the local climate. As it does not include any urban land use, it is not representative for the current land use in the region. This can be seen in the performance level for the NoUrban where the correlation coefficient ranged between 0.13-0.41 (Table A1, A2, A3). However, the NoUrban simulation can act as an indication of how the local climate looked like before the urbanization. In the diurnal cycles during daytime (Figure 4, 5) and the simulated mean T2m based on city size (Figure 6), the NoUrban showed a lower T2m compared to the observation. This shows that less urban land use leads to lower temperature, which is in line with the studies of Ribeiro et al. (2021) and Amorim et al. (2020). When comparing the mean T2m for NoUrban and the observation of the present study (Figure 6), it indicates that the urbanization and increased urban land use have led to an increase in T2m in the Stockholm region. Based on the result, the possible impact of urbanization in the Stockholm region is an increase in mean T2m ranging from 0.44 °C for rural areas to 1.78 °C for large cities (Table 5).

4.3 Improvement of the WRF Model

The model was not able to reproduce the T2m for the rural sites to the same degree as for the urban sites (Figure 6). The simulations BEM, BEP, BEMLake, BEMlowBH and BEPlowBH showed a mean T2m between 1.18-1.23 °C above the observation for the rural sites (Table 5). Therefore, one area where the WRF model can be improved, is to make it reproduce T2m for areas with non-urban land use more accurately. In the present study, urban parameters used in the WRF had a resolution of 100 m. With a 100 m resolution, smaller features in villages and cities might not be represented entirely correct which can have an impact on the model output. One way to improve the model can be to use urban parameters with a higher resolution to represent features in villages and cities better. With a better representation of how the world really looks like, the model should be able to give an output closer to the observation.

For the BEMlowBH and BEPlowBH simulations, the same revised urban parameters (Table 2) were used for all investigated cities even though the cities have different sizes and therefore different urban geometry. When a smaller city receives urban parameters based on a larger city, the urban parameters are not fully representative which can result in an overestimation of the temperature. The simulated mean T2m showed an overestimation for all city size groups where the overestimation decreased with increased city size (Figure 6), which can be a result of the urban parameters used. The simulated mean T2m for the large city group is closer to the observation compared to the small-sized city group, which can indicate that the urban parameters are more in line with the urban geometry of the largest city, the city of Stockholm in this case. To reproduce a mean T2m closer to the observation for the small, and medium-sized city groups, representative urban parameters for all cities can be included in the model for a better representation of the urban geometry for the respective cities. This will possibly decrease the overestimation and improve the performance of the WRF model.

In the western part of the modelled region, 59°50’N, 16°30’E and 59°30’N, 16°E for NoUrban and BEPlowBH respectively, an area which displayed a significantly lower T2m than for the rest of the region was found (Figure 7a, 7b). For NoUrban, the T2m was about 5 °C lower

compared to the surrounding areas, and for the BEPlowBH, the T2m was about 9 °C lower compared to the surrounding area. This underestimation of the T2m was found over the same area in BEM, BEP, BEMLake, and BEMlowBH as in BEPlowBH (Figure A1 in Appendix A). Here, it is evident that the model failed to reproduce the T2m over this area as it is unrealistic that the mean T2m is so much lower in these areas compared to the surrounding areas. The areas which showed lower T2m are mainly over land with the exception for a smaller part of the area which extends over the lake Mälaren. Further investigation is needed to understand why there is an underestimation of T2m in these areas, but it is possible that this can have an impact on the rest of the modelled area and that it can affect the overall performance of the model.

Improvement of the WRF is needed to get a model that accurately can reproduce T2m for both urban and rural areas. The BEMlowBH and BEPlowBH simulations had a high correlation (0.90-0.91) for the urban locations but was significantly lower (0.50) for the rural locations (Table A2, A3 in Appendix A). The deviation from the observation were roughly 0.7-0.9 °C higher for the rural locations compared to the large size city group (Table 5). Hence, these simulations have difficulties to reproduce T2m of rural areas compared to urban locations where further investigation is needed to improve the model.

5. Conclusion

This study was investigating how well the WRF model was able to reproduce summer T2m over the Stockholm region 2018 for cities with various sizes. Furthermore, the model's performance over different land uses in terms of urban/non-urban land use was investigated. Six different simulations were carried out (BEM, BEP, BEMLake, NoUrban, BEMlowBH, and BEPlowBH). It was found that the simulations with lowered building height according to the actual building height in the region (BEMlowBH and BEPlowBH) had the highest performance level for the urban locations with a correlation coefficient of 0.90 and 0.91 along with a CRMSE of 0.23 and 0.21, respectively. However, BEMlowBH and BEPlowBH did not reproduce T2m of the rural areas very well, with a correlation coefficient of 0.50 and 0.50 along with a CRMSE of 0.38 and 0.39, respectively. The model was able to reproduce a diurnal cycle close to the observation, however, an overestimation of the T2m both during day and night occurred for the majority of the locations. This overestimation could be because of a mismatch in land use; however, further investigation is needed to determine the cause of the overestimation. The model did reproduce mean T2m better for cities of larger size. Additionally, WRF was able to capture the UHI effect for medium and large cities. This study concludes that WRF can reproduce T2m over urban areas with higher precision than over rural areas. Moreover, correct land use data is of great importance to get a high performance level of the model and that further investigation is needed to improve the model regarding recreating T2m of rural areas.

6. Acknowledgments

The author would like to thank the supervisor Tinghai Ou, at the Department of Earth Science at the University of Gothenburg, for providing the model output, giving valuable guidance throughout the process, and for giving feedback on the first draft of the manuscript.

7. Appendix A

The performance statistics standard deviation, CRMSE, and correlation coefficient from the WRF model displayed in the Taylor diagrams (Figure 3) can be seen in Table A1, A2, and A3 for all locations, urban locations, and rural locations, respectively.

Table A1. Performance statistics of the six different simulations for all locations.

Simulation	Standard deviation	CRMSE	Corelation Coefficient
BEM	0.41	0.40	0.68
BEP	0.38	0.46	0.58
BEMLake	0.41	0.41	0.67
NoUrban	0.34	0.61	0.13
BEMlowBH	0.41	0.39	0.71
BEPlowBH	0.40	0.42	0.65

Table A2. Performance statistics of the six different simulations for the urban locations.

Simulation	Standard deviation	CRMSE	Corelation Coefficient
BEM	0.30	0.26	0.84
BEP	0.27	0.29	0.78
BEMLake	0.29	0.27	0.81
NoUrban	0.31	0.46	0.31
BEMlowBH	0.29	0.23	0.90
BEPlowBH	0.31	0.21	0.91

Table A3. Performance statistics of the six different simulations for the rural locations.

Simulation	Standard deviation	CRMSE	Corelation Coefficient
BEM	0.41	0.38	0.52
BEP	0.42	0.38	0.52
BEMLake	0.42	0.38	0.54
NoUrban	0.35	0.39	0.41
BEMlowBH	0.41	0.38	0.50
BEPlowBH	0.42	0.39	0.50

For the six simulations, the mean T2m for each grid cell were computed for the entire summer of 2018 which is displyed in Figure A1.

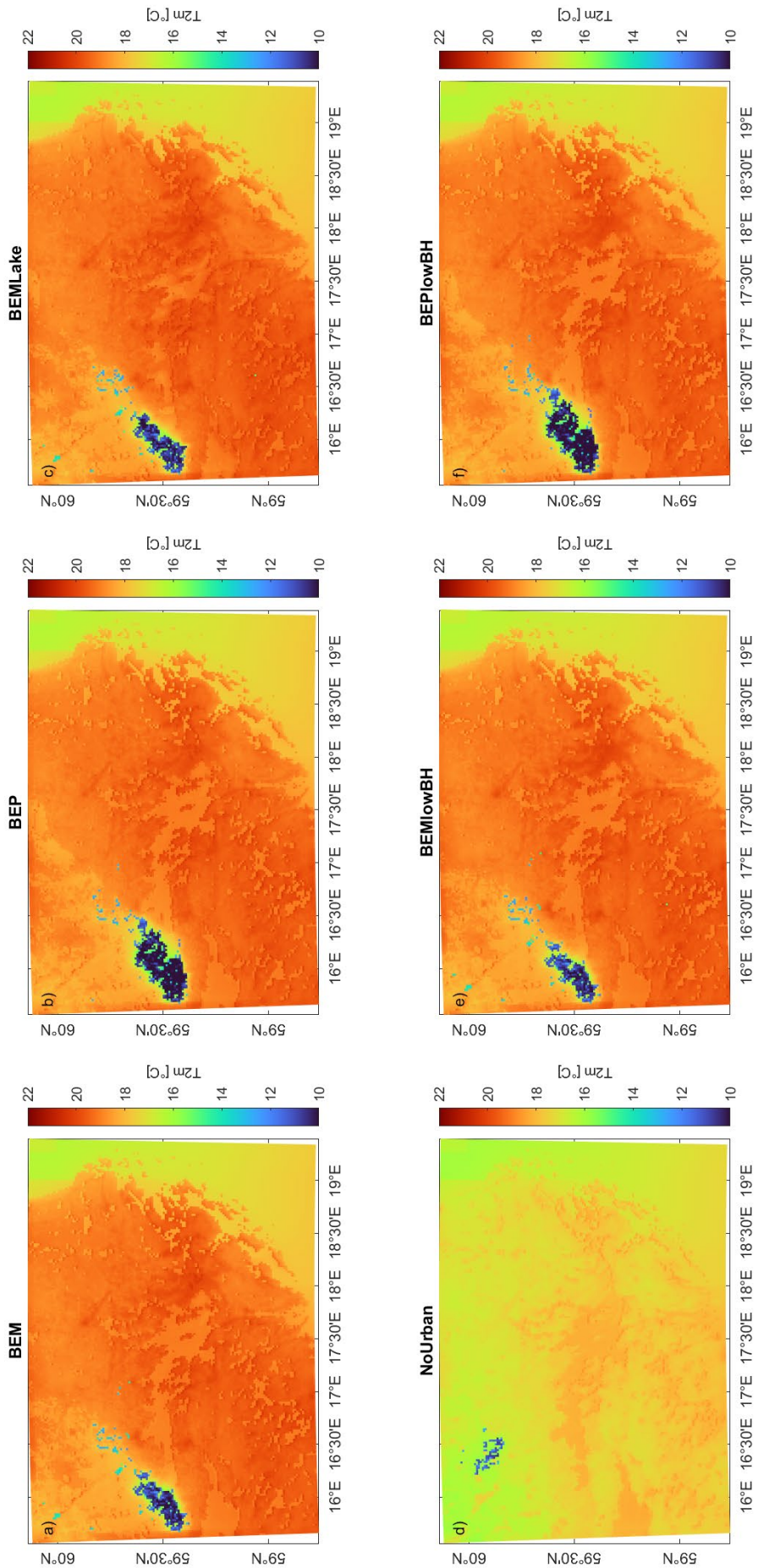

Figure A1. WRF simulated mean T2m for the summer 2018 for all simulations.

8. References

Amorim, J. H., Segersson, D., Körnich, H., Asker, C., Olsson, E., & Gidhagen, L. (2020). High resolution simulation of Stockholm's air temperature and its interactions with urban development. *Urban Climate*, *32*, 100632.

Arkon, C. A., & Özkol, Ü. (2014). Effect of urban geometry on pedestrian-level wind velocity. *Architectural Science Review*, *57*(1), 4-19.

Broitman, D., & Koomen, E. (2015). Residential density change: Densification and urban expansion. *Computers, Environment and Urban Systems*, *54*, 32-46.

Chen, F., Yu, B., Wu, M., & Yang, X. (2022). Improved Urban Finescale Forecasting During a Heat Wave by Using High-Resolution Urban Canopy Parameters. *Frontiers in Climate*, *3*, 197.

Danniswari, D., Honjo, T., Kato, A., & Furuya, K. (2022). Utilizing Open-Source Satellite Data for the Relationship between Building Height and Land Surface Temperature. *Journal of Environmental Information Science*, *2021*(2), 1-10.

Demuzere, M., Argüeso, D., Zonato, A., & Kittner, J. (2022a). W2W: A Python package that injects WUDAPT's Local Climate Zone information in WRF. *Journal of Open Source Software*, *7*(76), 4432.

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., & Bechtel, B. (2022b). A global map of local climate zones to support earth system modelling and urban-scale environmental science. *Earth System Science Data*, *14*(8), 3835-3873.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., & Schepers, D. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049.

Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., & Guiot, J. (2018). Impacts of 1.5 C global warming on natural and human systems. *Global warming of 1.5° C.*

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, *113*(D13).

Jahangir, M. S., & Moghim, S. (2019). Assessment of the urban heat island in the city of Tehran using reliability methods. *Atmospheric Research*, *225*, 144-156.

Janjić, Z. I. (1994). The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly weather review*, *122*(5), 927-945.

Janjić, Z. I. (1996). The surface layer in the NCEP Eta Model. Eleventh conference on numerical weather prediction, Norfolk, VA, 19–23 August 1996. Amer Meteor Soc, Boston, MA, pp 354–355.

Kim, Y.-H., & Baik, J.-J. (2004). Daily maximum urban heat island intensity in large cities of Korea. *Theoretical and Applied Climatology*, *79*, 151-164.

Lundgren Kownacki, K., Gao, C., Kuklane, K., & Wierzbicka, A. (2019). Heat stress in indoor environments of scandinavian urban areas: A literature review. *International journal of environmental research and public health*, *16*(4), 560.

Martilli, A., Clappier, A., & Rotach, M. W. (2002). An urban surface exchange parameterisation for mesoscale models. *Boundary-layer meteorology*, *104*, 261-304.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., & Rosero, E. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research: Atmospheres*, *116*(D12).

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and earth system sciences*, *11*(5), 1633-1644.

Pesaresi, M., Corbane, C., Ren, C., & Edward, N. (2021). Generalized Vertical Components of built-up areas from global Digital Elevation Models by multi-scale linear regression modelling. *Plos one*, *16*(2), e0244478.

Rezaei Rad, H., Rafieian, M., & Sozer, H. (2019). Evaluating the effects of increasing of building height on land surface temperature. *Journal of Urban Management and Energy Sustainability*, *1*(1), 52-57.

Ribeiro, I., Martilli, A., Falls, M., Zonato, A., & Villalba, G. (2021). Highly resolved WRF-BEP/BEM simulations over Barcelona urban area with LCZ. *Atmospheric Research*, *248*, 105220.

SCB. (2023). *Folkmängd i riket, län och kommuner 31 december 2022 och befolkningsförändringar 2022*. Retrieved 2023-04-21 from https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/pong/tabell-och-diagram/folkmangd-och-befolkningsforandringar---helarsstatistik/folkmangd-i-riket-lan-och-kommuner-31-december-2022-och-befolkningsforandringar-2022/

Schiavina, M., Melchiorri, M., Pesaresi, M., Politis, P., Freire, S., Maffenini, L., Florio, P., Ehrlich, D., Goch, K., & Tommasi, P. (2022). GHSL Data Package 2022.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., & Huang, X.-Y. (2021). A Description of the Advanced Research WRF Model Version 4.3. *National Center for Atmospheric Research: Boulder, CO, USA*, *No. NCAR/TN-556+STR*. https://doi.org/doi:10.5065/1dfh-6p97

Steeneveld, G.-J., Koopmans, S., Heusinkveld, B., Van Hove, L., & Holtslag, A. (2011). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research: Atmospheres*, *116*(D20).

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, *93*(12), 1879-1900.

Sveriges Kommuner och Regioner. (2022). *Kommungruppsindelning 2023*. https://skr.se/download/18.ef4ba7d1849a2f55db2898a/1669978414789/Kommungruppsindelning-2023.pdf

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly weather review*, *136*(12), 5095-5115.

United Nations. Department of Economic and Social Affairs. (2019). *World Urbanization Prospects: The 2018 Revision* (ST/ESA/SER.A/420). https://population.un.org/wup/publications/Files/WUP2018-Report.pdf

Ustaoglu, E., & Jacobs-Crisioni, C. (2022). What Drives Residential Land Expansion and Densification? An Analysis of Growing and Shrinking Regions. *Land*, *11*(10), 1679. https://www.mdpi.com/2073-445X/11/10/1679

Vinayak, B., Lee, H. S., Gedam, S., & Latha, R. (2022). Impacts of future urbanization on urban microclimate and thermal comfort over the Mumbai metropolitan region, India. *Sustainable Cities and Society*, *79*, 103703.

Wang, J., Liu, S., Meng, X., Gao, W., & Yuan, J. (2021). Application of retro-reflective materials in urban buildings: a comprehensive review. *Energy and Buildings*, *247*, 111137.

Zhao, L., Lee, X., Smith, R. B., & Oleson, K. (2014). Strong contributions of local background climate to urban heat islands. *Nature*, *511*(7508), 216-219.

Zheng, Z., Zhou, W., Yan, J., Qian, Y., Wang, J., & Li, W. (2019). The higher, the cooler? Effects of building height on land surface temperatures in residential areas of Beijing. *Physics and Chemistry of the Earth, Parts A/B/C*, *110*, 149-156.

Zhou, B., Rybski, D., & Kropp, J. P. (2017). The role of city size and urban form in the surface urban heat island. *Scientific reports*, *7*(1), 1-9.