AUTHOR(S)
Saranko, O., Suomi, J., Partanen, A.-I., Fortelius, C., Gonzales-Inca, C., & Käyhkö, J.

TITLE
Comparison of Physically Based and Empirical Modeling of Nighttime Spatial Temperature Variability during a Heatwave in and around a City

YEAR
2024

DOI
10.1175/jamc-d-23-0149.1

CITATION
Saranko, O., Suomi, J., Partanen, A.-I., Fortelius, C., Gonzales-Inca, C., & Käyhkö, J. (2024). *Comparison of Physically Based and Empirical Modeling of Nighttime Spatial Temperature Variability during a Heatwave in and around a City*. Journal of Applied Meteorology and Climatology, 63(10), 1053–1074. https://doi.org/10.1175/jamc-d-23-0149.1

VERSION
Accepted Manuscript

Comparison of physically based and empirical modeling of night-time spatial temperature variability during a heatwave in and around a city

Olli Saranko[a], Juuso Suomi[b], Antti-Ilari Partanen[c], Carl Fortelius[a], Carlos Gonzales-Inca[b], Jukka Käyhkö[b]

[a] *Meteorological Research, Finnish Meteorological Institute, Helsinki, Finland*

[b] *Section of Geography, Department of Geography and Geology, University of Turku, Turku, Finland. Postal address: FI-20014 TURUN YLIOPISTO, FINLAND.*

[c] *Climate System Research, Finnish Meteorological Institute, Helsinki, Finland*

Corresponding author: Olli Saranko, olli.saranko@fmi.fi

ABSTRACT

THE numerical weather prediction model HARMONIE-AROME and a multiple linear regression model (referred to in this article as the TURCLIM model after the local climate observation network) were used to model surface air temperature for 25 to 31 July 2018 in the City of Turku, Finland, to study their performance in urban areas and surrounding rural areas. The 2 am (local standard time) temperatures modeled by the HARMONIE-AROME and TURCLIM models were compared to each other and against the observed temperatures to find the model best suited for modeling the urban heat island effect and other spatial temperature variability during heatwaves. Observed temperatures were collected from 74 sites, representing both rural and urban environments. Both models were able to reproduce the spatial night-time temperature variation. However, HARMONIE-AROME modeled temperatures were systematically warmer than the observed temperatures in stable conditions. Spatial differences between the models were mostly related to the physiographic characteristics: for the urban areas, HARMONIE-AROME modeled on average 1.4 °C higher temperatures than the TURCLIM model, while for other land cover types the average difference was 0.51 °C at maximum. The TURCLIM model performed well when the explanatory variables were able to incorporate enough information on the surrounding physiography. Respectively, systematic cold or warm bias occurred in the areas in which the

File generated with AMS Word template 2.0

thermophysically relevant physiography was lacking or was only partly captured by the model.

SIGNIFICANCE STATEMENT

As more and more people are living in an urban environment, the demand for accurate urban climate modeling is growing. This study aims to understand the differences between the numerical weather prediction and multiple linear regression modeling, and their limitations in modeling surface air temperature in sub-kilometer scale. The case-study shows that models are capable of predicting the spatial variation of 2 am night-time temperature during a heatwave in a high-latitude coastal city. Both models are therefore valuable assets for city planners who need accurate information about the impacts of the physiography on the urban climate. The results indicate that to improve the performance of the models, more accurate physiographic description and higher spatial resolution of the models are needed.

1. Introduction

Development of modeling methods for city scale spatial air temperature differences, together with increased availability of relevant GIS-based input data (Dai et al., 2018; Guerri et al., 2023), has broadened the spectrum of modeling approaches applied during the last decades. Linear regression has been used in urban temperature modeling since the 1990's (e.g., Eliasson, 1992; Kim and Baik, 2002). The modeling settings have varied, and along with increased amount of high-quality open access data, one rather common approach has been the use of multiple linear regression model incorporating multiple GIS data based explanatory variables, and temperature as a dependent variable (Bottyán and Unger, 2003; Ketterer and Matzarakis, 2015; Porangaba et al., 2021). Such regression models are based on the assumptions on thermal impacts of land cover, urban morphology, topography, water bodies and other relevant environmental factors on air temperature. In the explanatory variable formation, these factors can already as such be used in a numerical form (e.g., elevation) or they can be converted to a numerical form by e.g., calculating their spatial coverage (%) inside a buffer zone around the temperature observation site (Foissard et al., 2019; Oukawa et al., 2022).

Another approach to model urban scale spatial temperature differences is to use physically based modeling, as in numerical weather prediction models and in climate models (see Hamdi et al. 2020 for review on physically based urban modeling, and Ronda et al. 2017 for high-resolution urban modeling). Numerical weather prediction (NWP) (Bauer et al., 2015) is the primary method for operational weather forecasting, using physics based numerical models to simulate the physical phenomena occurring in the atmosphere and near the surface. The resolution as well as the size of the modeled domain vary, depending on the need and the available computational resources.

One urban scale spatial temperature difference phenomenon is the urban heat island (UHI) which means the warmness of urban areas compared to their rural surroundings. The UHI results mainly from differences in 1) solar heat storage, 2) anthropogenic heat release, and 3) evaporation between urban and rural areas (Oke, 1987). When referring to UHI in this study, we mean the canopy layer UHI which means the urban-rural temperature difference in the lower part of the atmospheric canopy layer (WMO, 2023). In high latitudes in winter, spring and autumn, the consequences of UHI, such as lower heating demand and longer growing season, are often considered beneficial phenomena (Santamouris, 2014; Schatz and Kucharic, 2016; Yang et al., 2020). In high latitudes in summer, and in lower latitudes throughout the year, the high temperatures can cause health problems and even increase mortality (Hajat and Kosatky, 2010; Kollanus et al., 2021). As a consequence of rise in temperature due to climate change, heat-related health problems will get more common, and in urban areas, UHI will for its part probably worsen the situation, as already in current climate it is likely increasing the mortality (Kivimäki et al., 2023). Better knowledge on the city-scale heat-related health risks and their spatial distribution call for more accurate UHI modeling (Ruuhela et al., 2021). The purpose of this study is to compare and analyze the performance of two different kind of spatial temperature prediction models, namely a multiple linear regression model (later the ‘TURCLIM model’) and a NWP model HARMONIE-AROME (Bengtsson et al., 2017) in modeling the UHI and other spatial temperature variations during a July heatwave. In literature, the UHI is reported to be strongest 3-5 hours after sunset (Oke and Maxwell, 1975; Oke, 1987; see also Bokwa et al., 2015), and also the spatial temperature variability in general is often largest at night or early morning (Suomi and Käyhkö, 2012; Suomi, 2018). Consequently, we compared the surface air temperatures at 2 am local standard time (UTC +2), which fits inside the time frame of diurnally strongest UHI, during a one-week long study period in July 2018 in the city of

Turku and its surroundings on the south-western coast of Finland. In all contexts of this study, when we discuss temperature, we refer to the surface air temperature. In the case of the TURCLIM model, this means 3 m height, and 2 m height for the HARMONIE-AROME model.

The main objectives of this study are to:

1. Investigate, how the night-time surface air temperatures and their spatial variability modeled with the HARMONIE-AROME model and the TURCLIM model correspond to the observations
2. Assess how do the night-time surface air temperatures modeled by the HARMONIE-AROME model and the TURCLIM model differ spatially and temporally from each other
3. Assess which factors explain these differences, and what are the implications in future model development

2. Study area

Our study area in south-western Finland on the coast of the Baltic Sea consists of a middle-size (ca. 200,000 inhabitants) city of Turku (city center: 60°27’N, 22°16’E) and parts of its neighboring municipalities (Fig. 1). The second largest city in the study area is Salo (52,000 inhabitants), located 50 km to the ESE of Turku. The extent of the rectangular study area delineated for the modeling purposes is 113 km (E-W) times 76 km (N-S). The climate of the Turku region is a combination of coastal and inland types because of the relatively large islands next to the city center and an extensive archipelago consisting of thousands of islands stretching tens of kilometers to the sea. The location and movements of large weather systems determine temporally whether continental or marine characteristics dominate (Alalammi, 1987). Towards the inland from the city center, the climate becomes gradually more continental.

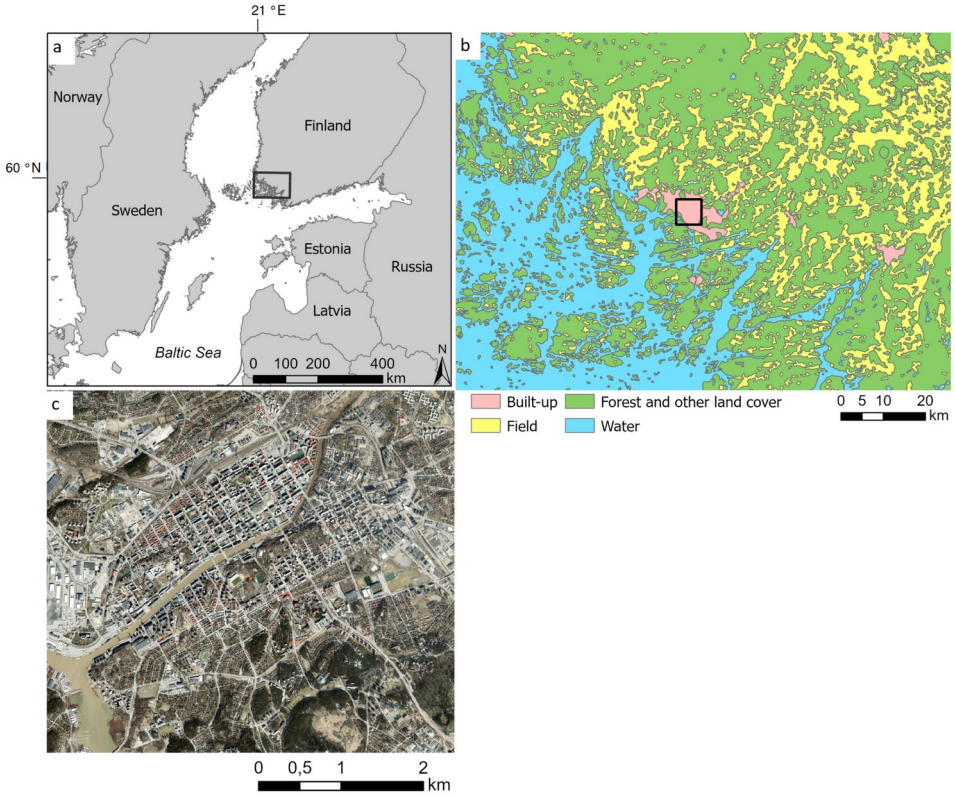

Fig. 1. The study area in the southwestern Finland. In the subfigure b, city of Turku is the built-up area in the center of the subfigure, and the urban area in the center right is the Salo city center. The aerial photo (subfigure c) presents the 6 x 6 km area in the Turku city center. The respective area is marked with a black square in the subfigure b, which in turn is marked with a black rectangle in the subfigure a.

In Köppen's climate classification, Turku belongs to the hemiboreal and humid Dfb class, which extends to the Baltic countries, the southern parts of Scandinavian Peninsula, the eastern part of Europe and the mid-latitudes of western Asia, and the North American Great Lakes region. By using 1991-2020 as a reference period, the annual average temperature at Turku Airport, approximately 7 km to the north of the city center, is 5.8 °C (Jokinen et al., 2021). February is the coldest month based on the average temperature (-4.5 °C), and the average daily minimum and maximum temperatures (-7.1 °C and -1.2 °C, respectively). The warmest month is July, based on the average temperature (17.5 °C), and the average daily minimum and maximum temperatures (12.5 °C and 22.6 °C, respectively). The highest and lowest observed temperatures in 1991-2020 were +33.0 °C (Jul 2018), and -28.2 °C (Jan 2013), respectively. The mean annual rainfall in Turku is 684 mm, with April (32 mm) as the driest month and July (74 mm) the wettest. The length of the permanent snow cover period increases from the archipelago towards the inland, being approximately 2-3 months around

the Turku city center, starting typically at the end of December (FMI, 2022). The average wind speed in the area is 3.4 m s^{-1}, being strongest in December (3.7 m s^{-1}), and weakest (3.1 m s^{-1}) in July, August and September. The dominant wind direction (16 %) is southeasterly, while northerly and northeasterly winds are the rarest (8 %). During the study period, 25 to 31 July 2018, the city of Turku experienced a strong heatwave. The average temperature of the study period was 7.9 °C warmer than the average temperature of 17.5 °C in July in the period 1991-2020. Altogether, the heatwave lasted almost uninterrupted from 12 July to 3 August, but to eliminate the potential impact of the day of the week on UHI, we selected one week period for the analyses (see e.g. Earl et al., 2016).

The extent of the Turku city center is approximately 1.5 km (SE-NW) times 4 km (SW-NE). The streets are oriented from SW to NE and from SE to NW. Otherwise the grid plan area mostly consists of 6-8 storey high blocks of flats, and scattered parks. Commercial activities are concentrated in the surroundings of the market place located in the middle of the grid plan area. The 50-100 m wide Aura River flows from NE to SW across the grid plan. The city center area extends over relatively flat terrain 5-10 m above sea level, with scattered bedrock hills 30-50 m in height. Beyond the grid plan area, the land cover consists of a mosaic of built-up areas, forests and fields (City of Turku, 2022).

3. Data and methods

a. Temperature data

Temperature observations have been recorded as part of the Turku Urban Climate Research Project (TURCLIM) of the Geography Division at the University of Turku. The TURCLIM observation network currently consists of altogether 83 Onset Hobo Pro v2 U23-001 temperature and relative humidity data loggers, placed inside Onset RS1 radiation shields on poles at 3 m elevation above the ground. The elevation differs from the standard 2 m elevation in order to minimize the risk of vandalism. The observation interval is 30 minutes. According to the manufacturer, the accuracy of the instrument is ±0.2 °C at 0–70 °C, and ±0.25 °C at -40–0 °C, while the resolution is 0.04 °C. In this study, we use temperature observations of 74 sites.

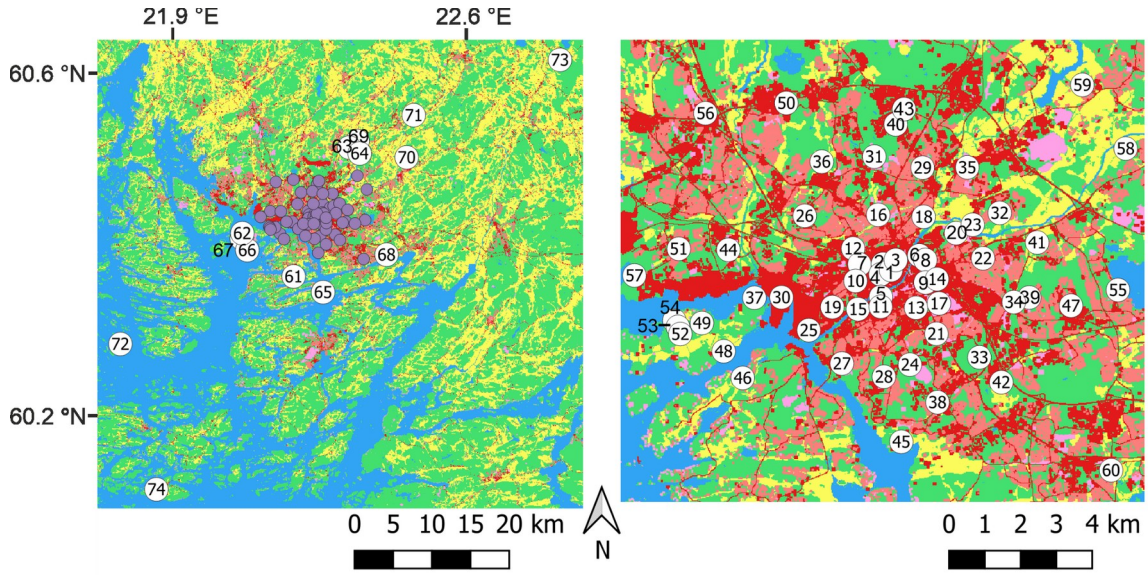

Fig. 2. Observation sites in the study area. The order of the numbering is based on the beeline distance from the Turku city center (site no. 1, Market place). The sites that are presented on the right-hand side map are marked with purple dots on the left-hand site map. For more information on the sites, see Table A1.

b. TURCLIM model

Spatially continuous temperature maps were produced with the TURCLIM model (Fig. 3). The observed temperatures of the observation site network acted as response variables, while the GIS-based variables representing the impact of topography, water bodies and land use acted as explanatory variables. The sizes of the circle-shaped footprint areas of the explanatory variables were determined case-specifically based on the Pearson's correlation coefficients between the response and explanatory variables. The correlation analyses were performed with the original resolutions of the dataset, i.e., in the case of topography and water bodies with 10 m resolution and in case of land use with 20 m resolution. Either the footprint area with the highest correlation coefficient or the correlation coefficient representing clear saturation of the coefficient in relation to the change in footprint area was selected to be used in the respective regression model. Also, the thermophysical mechanisms relevant in the climatic impact of the variable in question were considered in determining the footprint area. E.g., the impact of a variable that represents land use is mostly transmitted via heat flux that originates from anthropogenic activity and solar heat storage and releases from urban construction material. Before deciding the footprint area, the area suggested by the correlation analysis was first compared to the literature (see e.g. Oke, 2006), and if considered to be in line with earlier studies, accepted as a footprint area of the respective variable in this study.

The variable reflecting the impact of water bodies was formulated based on the SLICES land use classification (spatial resolution 10 x 10 m) (National Land Survey of Finland, 2020). Sea areas, lakes and rivers were combined to represent surface waters in the area. Numerically, the variable represented the area of water bodies in relation to the area of the footprint area as a whole. The variable reflecting the impact of land use was formulated from CORINE Land Cover (CLC) 2018 dataset (spatial resolution 20 x 20 m) (SYKE, 2023). CLC classes representing continuous urban fabric, commercial units and road and rail networks and associated lands were reclassified to a single class reflecting the impact of urban land use on temperatures. The reclassification was based on metadata of the CLC classification and Pearson's correlation coefficients between temperatures and spatial coverage of single CLC classes in the vicinity of temperature measurement sites. After preliminary analyses, coverage of the reclassified urban land use class of the total footprint area represented the numeric value of the variable. Relative elevation acted as a variable reflecting the local climatic impact of topography in the area. The variable was formulated from the Digital Elevation Model (DEM, spatial resolution 10 x 10 m) (National Land Survey of Finland, 2019), and indicated whether the site was in a lower or higher position relative to its surroundings. Numerically, the variable described how many meters higher position (positive values) or lower position (negative values) the site is relative to the average elevation of the footprint area. In the TURCLIM model, these numeric explanatory variables reflect thermophysical impacts of the environmental factors: for example high proportion of urban land use (and high value of respective explanatory variable) in the vicinity means higher temperature of those areas compared to the areas with less urban land use. The variable formulation was performed with the ArcGIS Pro 2.9 software, and the correlation analyses were made with the IBM SPSS Statistics 27 software.

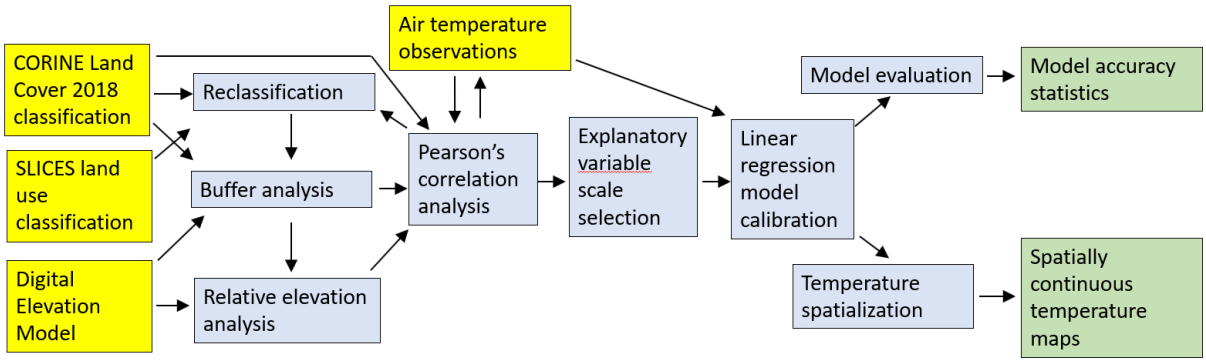

Fig. 3. A flow diagram representing the steps of the TURCLIM model including data (yellow), methods (blue) and results (green).

TURCLIM model including explanatory variables for water bodies, relative elevation and urban land cover, was calibrated with the 'Enter' setting for each case (each day in 25 to 31 July 2018 at 2 am local standard time + average of the previously mentioned times = 8 cases) with the temperature data of all the available 71 observation sites as a dependent variable. With the 'Enter' setting, all explanatory variables are entered in a model in a single step in spite of their statistical significance. These model results were rasterized to 100 m resolution spatially continuous temperature maps. To evaluate the accuracy of the models with independent evaluation data, the model calibration was also performed with the data of 60-61 observation sites, whereas the remaining 10-11 sites were used in model accuracy evaluation. When evaluating the model accuracy, the model calibration was repeated 7 times per case so that each observation site was involved in model calibration and model accuracy evaluation. Additionally, temperature data of 3 remote observation sites that are located approximately 30-40 km from the city center (see Fig. 2) and that were not used in any model calibrations, were also included in the model accuracy evaluation, resulting in altogether 74 sites used for model evaluation. For the 3 remote observation sites, the model calibration was performed with the data of the previously mentioned 71 observation sites. The TURCLIM model was run by the IBM SPSS Statistics 27 software and the model accuracy estimation was performed with the MS Excel 2016 software.

In the TURCLIM model evaluation, the explanatory variables for temperature observation sites were extracted based on the 100 m resolution grid cell structure that was used in the spatialization of the modeling results. For comparison of the spatial differences between the TURCLIM model and HARMONIE-AROME model, the impact of different spatial resolutions between the models was first neutralized by generalizing the 100 m resolution temperatures predicted by the TURCLIM model to a 750 m resolution of the HARMONIE model. As 750 is not divisible with 100,the 100 m resolution TURCLIM model temperatures were first resampled to 50 m. After that, the average temperature of the TURCLIM model temperature was calculated for each 750 m resolution HARMONIE-AROME model grid cell. The spatialization and generalization of the modeling results was done with ArcGIS Pro 2.9 software.

c. HARMONIE-AROME

HARMONIE-AROME (Bengtsson et al., 2017) is a short-range limited-area numerical weather prediction system used for high-resolution forecasts at the Finnish Meteorological Institute, and at several other European meteorological services. It is developed in the frame of the international ACCORD consortium formed by 26 European services (ACCORD, 2022). HARMONIE-AROME includes a non-hydrostatic, convection-permitting limited-area model of the atmosphere. Sub-grid scale atmospheric physical processes treated in the model include cloud microphysics and precipitation, radiative transfer, turbulence and shallow convection.

The interaction between surface and atmosphere through the exchange of energy, water and momentum is calculated by the surface and soil module SURFEX (Masson et al., 2013). SURFEX includes sub models for four different surface types: Sea and ocean, inland water, vegetated surfaces and urban surfaces. For example, biophysical processes of soil and vegetation related variables, and their interactions with the atmosphere are treated by the Interactions between Soil–Biosphere–Atmosphere model (ISBA) (Noilhan and Planton, 1989, Masson et al., 2013).

In the HARMONIE-AROME model, the built-up areas are handled by the Town Energy Balance model (TEB) (Masson, 2000). TEB is a single layer urban canopy model which calculates energy budgets for three surfaces: roads, roofs, and walls. The model assumes a homogeneous street layout inside a grid cell. TEB takes into account several physical processes related to urban environment: 1) shortwave and longwave trapping effect of the canyon geometry; 2) anthropogenic sensible heat flux from heated or cooled buildings or from traffic and industry; 3) water and snow interception by roofs and roads; 4) heat conduction and heat storage in buildings and roads; and 5) interactions between the built surfaces and the canyon air (temperature, specific humidity, wind, turbulence). TEB has been shown to be capable of predicting the UHI (Hamdi 2010, Hamdi et al. 2012).

When both town and nature are present in a model grid cell, the diagnostic surface parameters, such as temperature at 2 meter height, are calculated separately in ISBA and TEB, and then averaged together based on the grid tile fractions. In ISBA, diagnostic quantities are interpolated from atmospheric forcing, and surface variables. In TEB, thermodynamic fluxes for the canyon air are assumed to be in equilibrium, and anthropogenic flux is taken into account (Masson, 2000).

In SURFEX 8.1 (CNRM 2017) and onwards, ECOCLIMAP Second Generation physiographic data base (CNRM 2018) is used for topographic information. ECOCLIMAP-SG land cover map has a resolution of 300 m, and has 33 cover types. The 10 urban cover types represent the 10 urban local climate zones (LCZ) (Stewart and Oke, 2012).

In this study, HARMONIE-AROME was used as a forecasting system to generate short-range hind casts of the weather conditions during a heatwave in July 2018. The experiment was run with a grid size of 750 m and with 65 vertical levels. The applied version of HARMONIE-AROME (Cy43h2.1) uses the surface interaction sub model SURFEX 8.1 and thus ECOCLIMAP-SG physiographic data base was used after minor corrections to the sea and lake tiles based on local data. Analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) were used as boundary conditions at the lateral edges of the domain, and to supply the sea surface temperature which is not a prognostic variable of HARMONIE-AROME. Atmospheric variables were initialized by combining forecasts from the previous cycle to large-scale features from the hosting analyses. In order to keep the model as close as possible to the actual prevailing conditions during the experiment, surface and soil variables (road surface temperature in an urban environment) were updated by data assimilation using observations of near surface temperature and relative humidity from reporting weather stations (located outside the cities). The experiment produced 6-hour forecasts four times for each day. See Fig. 4 for a flow diagram describing the steps of HARMONIE-AROME.

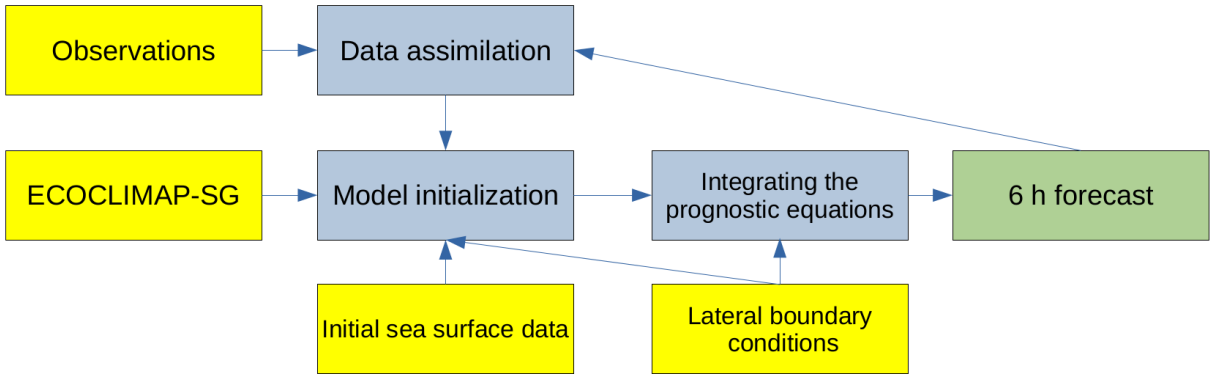

Fig. 4. A flow diagram representing the steps of the HARMONIE-AROME model including data (yellow), methods (blue) and results (green).

d. Weather conditions during the study period

The average temperature of the study period 25 to 31 July 2018 was 25.4 °C in the Turku Artukainen weather station 5 km to the west of the city center. Daily maximum temperatures

were close to 30 °C, and daily minimum temperatures around 20 °C. During the first part of the study period, a high-pressure ridge extended over Finland from the east resulting in clear sky and weak wind conditions in the study area. This resulted in relatively large spatial temperature differences in the study area (Fig. 5). On the 29 and 30 July, the weather was cloudier and windier than on the other days of the study period, and during those days the spatial temperature differences were suppressed.

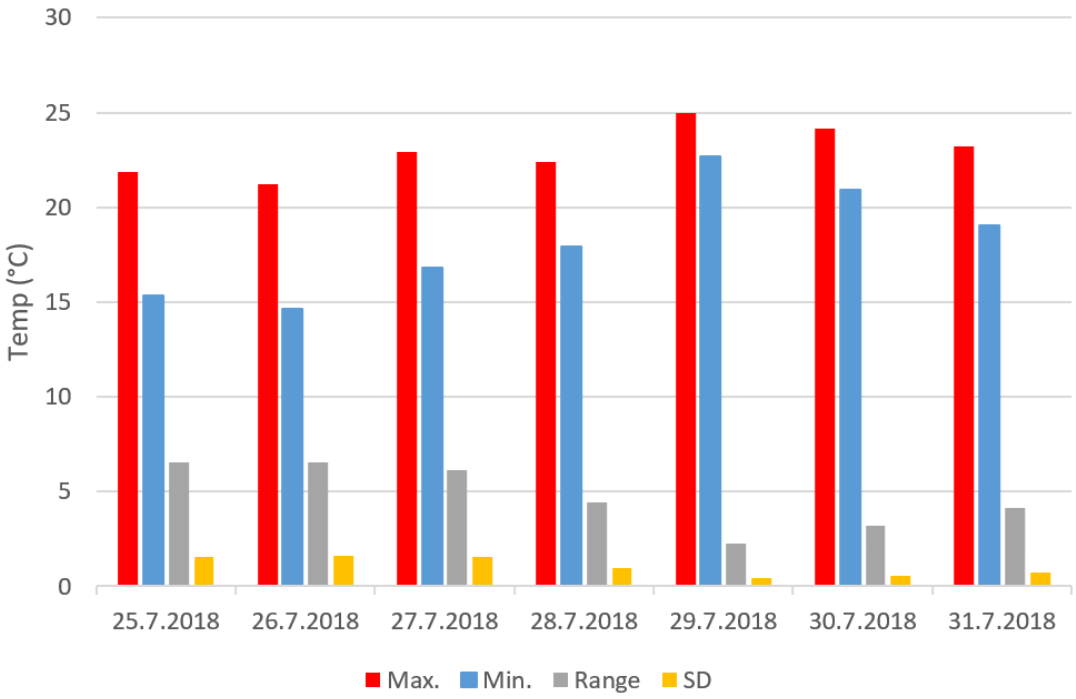

Fig. 5. Observed maximum and minimum temperature, spatial temperature range and standard deviation of temperatures (°C) at 2 am Finnish local standard time (UTC+2) for each day of the week as observed at the 71 sites used in the TURCLIM model calibration.

4. Results

a. Difference between observed and modeled temperatures

For the TURCLIM and the HARMONIE-AROME model, the mean absolute error (MAE), root mean square error (RMSE), and the maximum difference between the observed and modeled temperature were generally larger in the beginning of the study period, when also the spatial temperature differences were the largest (Table 1). Adjusted R^2 values, calculated for TURCLIM models only, followed a similar pattern. Regarding the average differences between the observed and modeled temperatures, HARMONIE-AROME and

TURCLIM model behaved differently; for HARMONIE-AROME model, the average differences were largest in the beginning and in the middle of the study period, whereas for the TURCLIM model, the respective differences were largest at the end of the study period.

Table 1. Mean absolute error (MAE), root mean square error (RMSE), average difference (Av. diff.), maximum difference (Max. diff), site with maximum difference, Index of Agreement and Pearson's correlation coefficients between the observed and modeled temperatures of the HARMONIE-AROME (HAR.) model and TURCLIM (TUR.) model (100 m resolution) + adjusted R^2 value of the TURCLIM model. Regarding the TURCLIM model, R^2 values represent the model calibrations performed with the data of all 71 observation sites. MAE, RMSE, Av. diff., Max. diff., Index of Agreement and Pearson's R values are determined based on the separate evaluation data. Each MAE value represents the average MAE of the 8 time-specific model calibrations, and the Max. diff. the largest difference of the 8 time-specific model calibrations. For information on the locations of the sites (Column Site with max. diff), see Fig. 2.

Error statistics	Model	25.7. 2 am	26.7. 2 am	27.7. 2 am	28.7. 2 am	29.7. 2 am	30.7. 2 am	31.7. 2 am	Average*	25.-31.7. **
MAE (obs-mod), °C	HAR.	1.83	2.34	1.66	2.39	0.48	0.4	1.21	1.47	
	TUR.	0.74	0.8	0.71	0.66	0.27	0.33	0.4	0.56	0.45
RMSE (obs-mod) °C	HAR.	2.12	2.61	1.92	2.55	0.53	0.5	1.3	1.65	
	TUR.	0.93	1	0.89	0.79	0.37	0.49	0.55	0.72	0.59
Av. diff. (obs-mod), °C	HAR.	-1.71	-2.26	-1.57	-2.38	0.39	-0.29	-1.22	-1.29	
	TUR.	0	-0.01	-0.02	-0.02	-0.02	-0.04	-0.02	-0.02	-0.01
Max. diff. (obs-mod), °C	HAR.	-4.43	-5.18	-4.84	-4.33	-1.13	-1.71	-2.91	-3.5	
	TUR.	2.35	2.92	-2.3	-2.14	-1.28	-1.98	-2.2	-0.66	-1.59
Site with max. diff.	HAR.	50	50	42	42	54	64	64		
	TUR.	52	52	2, 73	64	54	64	64		73
Index of agreement	HAR.	0.58	0.49	0.64	0.20	0.69	0.78	0.48	0.55	
	TUR.	0.88	0.88	0.90	0.72	0.70	0.63	0.79	0.78	0.88
Pearson's R	HAR.	0.66	0.67	0.74	0.50	0.65	0.73	0.79	0.68	
	TUR.	0.80	0.80	0.83	0.62	0.61	0.52	0.69	0.69	0.81
Adj. R^2	TUR.	0.67	0.66	0.72	0.41	0.4	0.36	0.5	0.53	0.69

** = Calculated based on each days' 2 am model results.*

*** = A model that is calibrated for the average temperature of 25 to 31 July at 2 am instead of separate model calibration for each day. Due to the differences in modeling methods, this can be performed only with the TURCLIM model.*

Considering the direction of the difference between the observed and modeled temperature, for HARMONIE-AROME and the TURCLIM model the average difference between the observed and modeled temperature is without one exception negative, i.e., the modeled temperature is higher than the observed temperature. In case of maximum difference, the respective difference is always negative in case of HARMONIE-AROME model, whereas for the TURCLIM model the respective difference is positive during the first two nights and during the rest of the nights negative. On average the average and maximum differences are negative for both models.

Regarding comparison between the HARMONIE-AROME and the TURCLIM model, MAE and average difference between the observed and modeled temperature was smaller for the TURCLIM model during each seven nights and on average. In the case of maximum difference between the observed and modeled temperature, the difference was larger for the HARMONIE-AROME model during five nights, and for the TURCLIM model during two nights. On average, the maximum difference was larger for the HARMONIE-AROME model. These model-specific characteristics are detectable also in the scatterplots representing the observation site specific differences between the observed and modeled temperatures (Fig. 6).

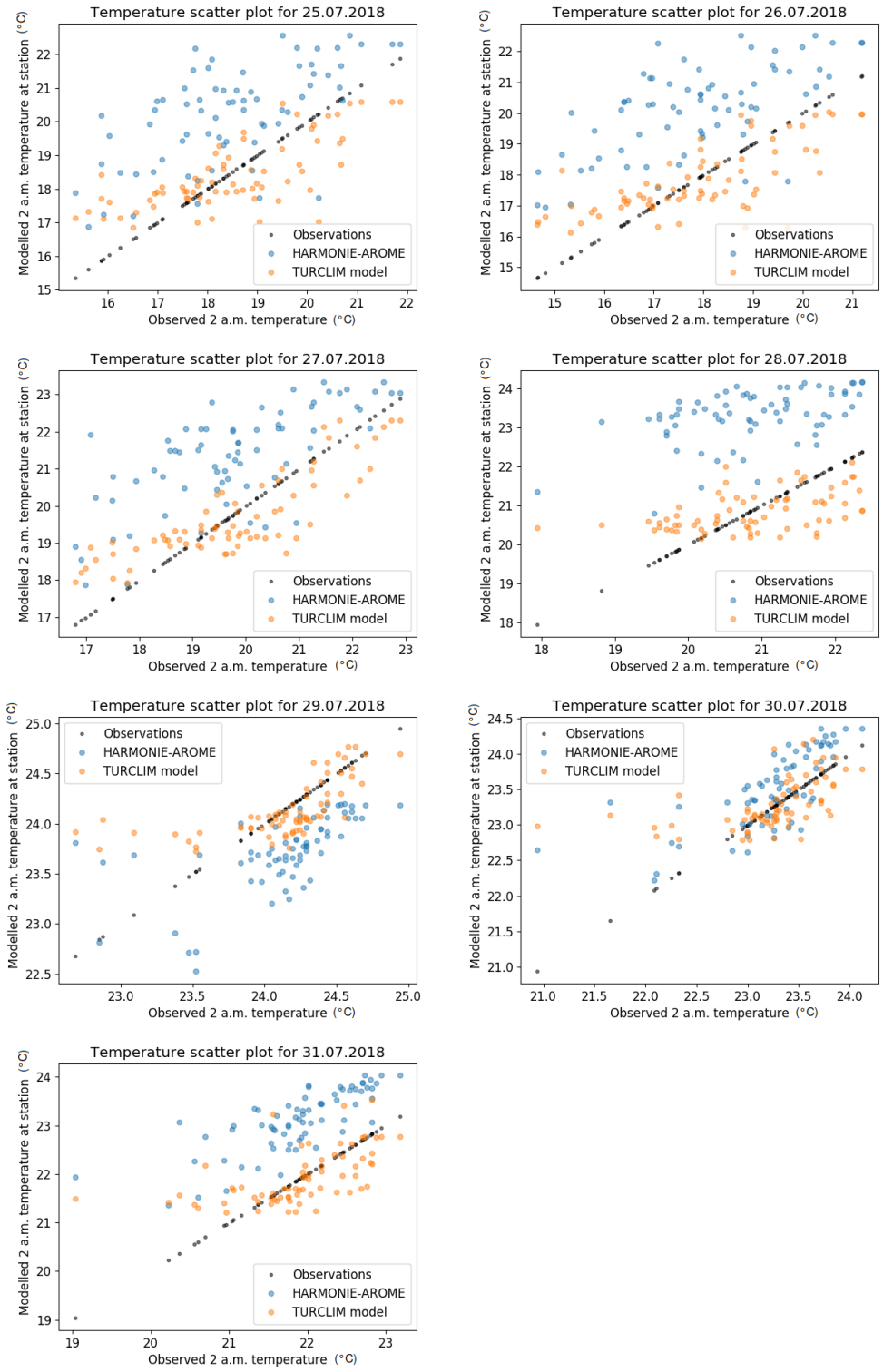

Fig. 6. The scatterplots on the modeled and observed temperatures. The black X = Y dots represent the observed temperatures, the blue dots the temperatures predicted by the HARMONIE-AROME model and the orange dots the temperatures predicted by the TURCLIM model.

b. Spatial pattern of the differences between the observed and modeled temperatures

Temperatures predicted by both models generally correspond rather well with the observed temperatures of the study area. E.g. the UHI of the Turku and Salo city centers are clearly visible in both modeling results (Figs 7 and 8), and the negative and positive residuals (difference between the observed and the modeled temperature) are not clearly concentrated on any certain areas (Figs 9 and 10), but the differences are mostly explained by the characteristics of the immediate neighborhood of the observation sites. There are also time-specific differences; some points that have negative residuals at certain night can have positive ones at some other night.

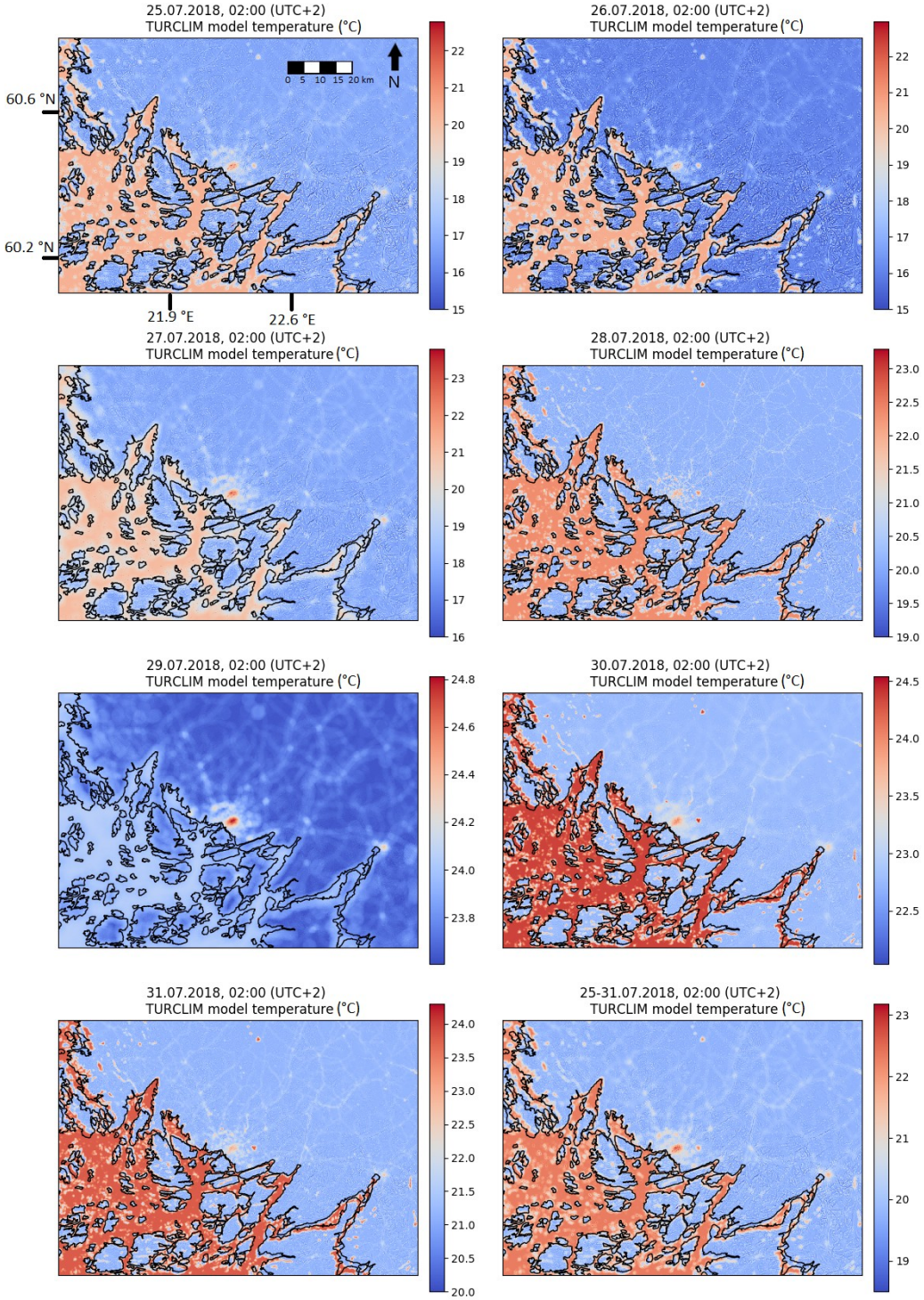

Fig. 7. 100 x 100 m resolution temperatures modeled by the TURCLIM model. The UHIs of the Turku and Salo city centers (for the locations, see Fig. 1) are detectable, as well as the warming impact of the sea. The minimum values for the 25th, 26th, 27th and 31st of July are 8.8, 7.6, 12.7 and 19.3 °C, respectively, but the scales of the maps were trimmed for clearer visualization. Similarly, the modeled average temperatures have a minimum temperature of 16.6 °C. These low minimum values are due to the local pronounced cooling impact of topography in over a 100 m deep open-pit limestone mine approximately 18 km to the south of the Turku city center, and thus do not represent the natural topography-based temperature variation of the study area.

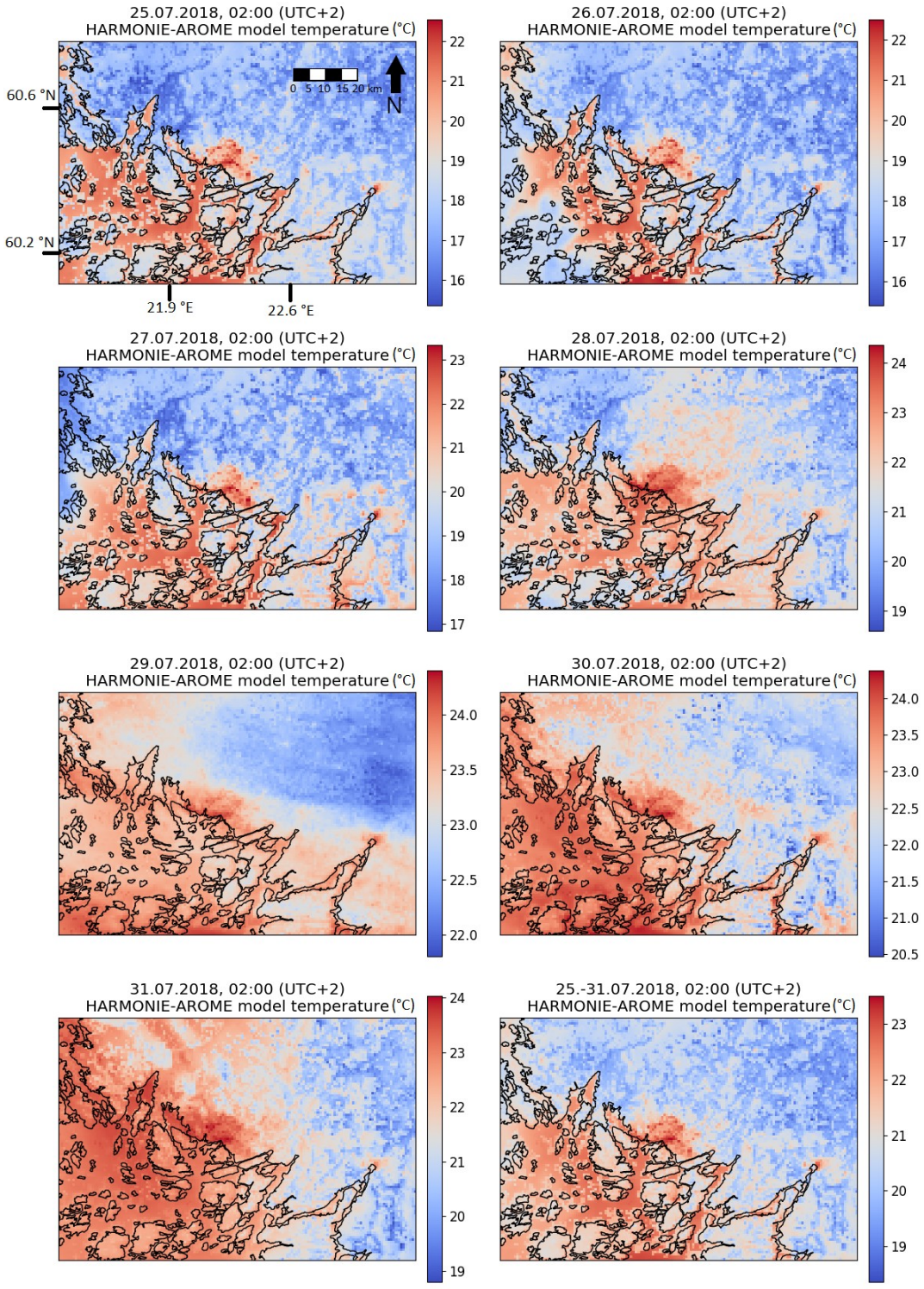

Fig. 8. The spatial distribution of 2 am temperatures modeled by HARMONIE-AROME for each day from 25 to 31 July 2018 and the average for the seven days.

Despite the generally logical results of both models, there are, however, relatively large differences between the observed and modeled temperatures at certain sites. For HARMONIE-AROME and TURCLIM models, the largest site-specific differences between

the observed and modeled temperature, 5.18 °C and 2.92 °C, respectively, occurred during the second night of the study period, and the second largest difference, 4.84 °C and 2.35 °C, during the third night of the study period for HARMONIE-AROME model and during the first night of the study period for TURCLIM model. For the HARMONIE-AROME model, the largest difference occurred at the observation site 50 (Mylly shopping mall, Fig. 2; Table A1) approximately 5.5 km to the NW of the Turku city center, and the second largest difference at the observation site 42 (Huhkola), located approximately 4.5 km to the SE of the Turku city center. For the TURCLIM model, the two largest differences occurred in the same site, namely in the observation site 52 (Hiiriluoto, inland) in the middle of Ruissalo island approximately 6 km to WSW of the city center. For the TURCLIM model the two largest differences were positive, i.e., the observed temperatures were warmer than the modeled ones, whereas for the HARMONIE-AROME model the differences were negative, i.e., the observed temperatures were colder than the modeled ones.

For the HARMONIE-AROME model, during the last five days of the study period, the difference between the observed and modeled temperature was largest during 27 and 28 July in the site 42 (Huhkola), during 29 July in site 54 (Hiiriluoto, shore), and during 30 and 31 July in site 64 (Niuskala). Site 54 is located on the seashore at the northern part of Ruissalo island approximately 6 kilometers to the WSW of the city center and approximately 400 meters to the NNW of the site 52 (Hiiriluoto, inland) referred earlier in this chapter. The site 64 is located in an uninhabited forested area approximately 10 km to the NNE of the Turku city center. On average, the difference between the observed and modeled temperature was largest (-2.70 °C) in site 42 (Huhkola).

Considering the rest of the week, for the TURCLIM model, most often, i.e., during three nights, the largest difference between the observed and modeled temperature occurred in site 64 (Niuskala). During one night the difference was largest both in site 2 (Puolalanmäki) at the city center and in site 73 (Karinainen) approximately 37 km to the NE of the city center, and during one night in observation site 54 (Hiiriluoto, shore) in the northern shore of the Ruissalo island, During all of these five nights, the observed temperatures were colder than the modeled ones. In addition to the single days, the TURCLIM model was calibrated also for the average temperature of single days (25 to 31 July) 2 am temperatures. For that model, the difference between the observed and modeled temperature was largest in site 73 (Karinainen), where the observed temperature was -1.59 °C colder than the modeled temperature. To summarize, during the last three nights of the study period, the maximum

difference occurred in the same site (twice at site 64 and once at site 54) for the HARMONIE-AROME and TURCLIM models.

Regarding the differences between the observed and modeled temperatures during the whole week, the majority of the sites (in case of HARMONIE-AROME 70 out of 74, and in case of the TURCLIM model 57 out of 74) had both positive and negative residuals, depending on the day. For the HARMONIE-AROME there were, however, 4 sites that had negative residuals each night, whereas for the TURCLIM model there were 7 sites that had negative residuals each night and 10 sites that had positive residuals each night. Regarding the TURCLIM model, for example, on the sites 25 (Turku castle) and 30 (Vapaavarasto) next to the Turku harbor the observed temperatures were systematically warmer than those predicted by the TURCLIM model (see Figs 2 and 10). This could be explained by the fact that the harbor area, located next to the both sites, consists of relatively large buildings and asphalted surfaces that store heat rather well in summer. That area, which is defined as ‘‘Port areas’ in the CLC is, however not included in the TURCLIM model as an individual explanatory variable or as a part of any other variable as it exists only in a limited area next to the sea and is thus irrelevant for the vast majority of the study area. Consequently, the warming impact of the Turku harbor is not reflected by the modeled temperature, which is seen as approximately 1 °C too cold modeled temperature for the sites 25 and 30.

On the contrary, examples on the site in which the TURCLIM model based temperatures are systematically too warm are sites 29 (Valkiasvuori) and 64 (Niuskala). Of those sites, site 29 is the only site for which both models predict systematically too warm temperatures (see Figs 2, 9 and 10). The site 29 is located in a grassy area in the yard of a detached house. The site is bordered from the south and southwest by a forested hill that results in unfavorable radiative conditions and increased proneness to cold air drainage (see Suomi et al., 2024). Systematically too warm modeled temperature for the site indicates that the local cooling factors are not properly incorporated in either model. Analogous challenge also relates to the site 64 (Niuskala), even if the HARMONIE-AROME -based modeled temperature for the site is at one night slightly (0.03 °C) colder than the respective observed temperature. That specific night (29 July) is an exception in the HARMONIE-AROME residuals also in general; whereas during the other nights, the residuals are dominantly negative, on 29 July the residuals are close to zero and on average slightly positive (see Fig. 9).

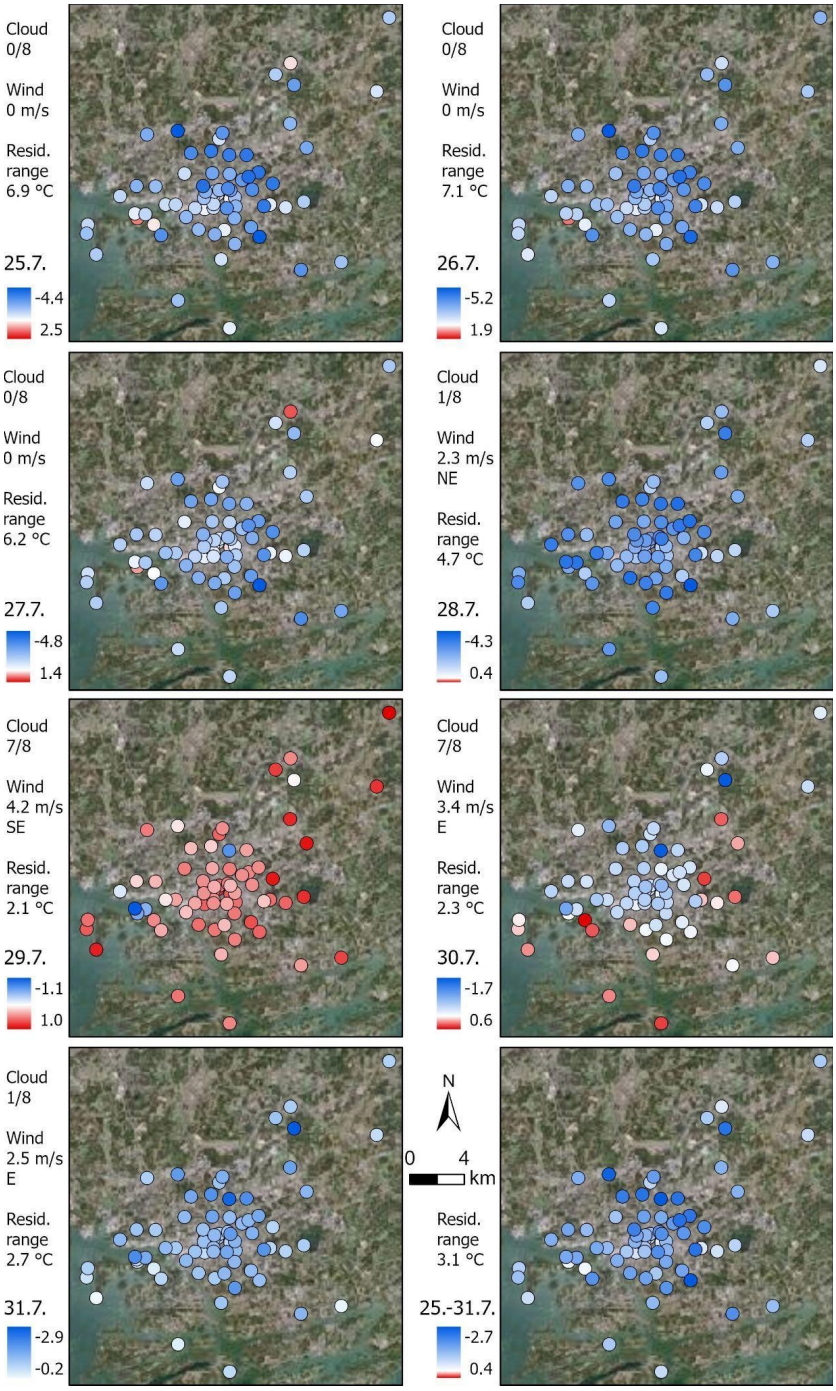

Fig. 9. Spatial distribution of the difference between the observed temperatures and HARMONIE-AROME based modeled temperatures at 2 am local standard time. Blue tones denote that the observed temperature is colder than the modeled, and red tones vice versa. Pale tone denotes only a minor difference between the observed and modeled temperatures. For better readability, the view (22.08 °E -22.50 °E (W-E) x 60.35 °N - 60.59 °N (S-N)) focuses on the 71 observation sites nearest to the Turku city center. The cloudiness and wind speed values represent the 2 am conditions at the Turku Artukainen weather station located approximately 5 km to the west of the Turku city center. Background map: Orthophoto, National Land Survey of Finland.

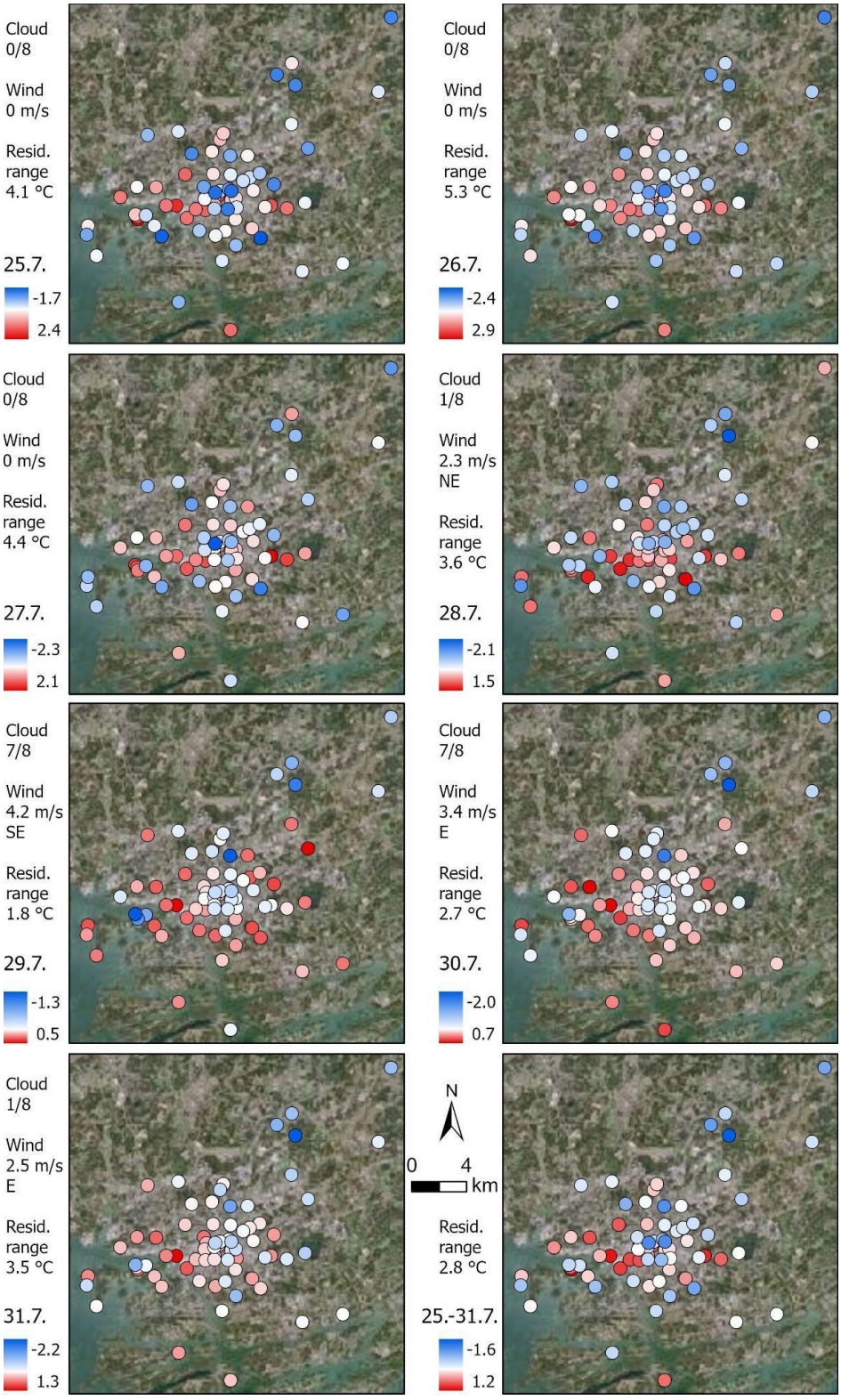

Fig. 10. Spatial distribution of the difference between the observed temperatures and TURCLIM based modeled temperatures 25 to 31 July 2018 at 2 am local standard time. Blue tones denote that the observed temperature is colder than the modeled one, and reddish vice versa. Pale tone denotes only a minor difference between the observed and modeled temperatures. For better readability, the view (22.08 °E -22.50 °E (W-E) x 60.35 °N - 60.59 °N (S-N)) focuses on the 71 observation sites nearest to the Turku city center. The cloudiness

and wind speed values represent the 2 am conditions at the Turku Artukainen weather station located approximately 5 km to the west of the Turku city center. Background map: Orthophoto, National Land Survey of Finland.

c. Differences between the temperatures modeled by HARMONIE-AROME model and TURCLIM model

For comparative purposes, the temperatures predicted by the TURCLIM model temperatures dealt with in this section have been generalized to 750 m resolution. The temperatures predicted by the HARMONIE-AROME model are on average slightly warmer than the respective TURCLIM model based temperatures. The scatter plots (Fig. 11) and histograms (Fig. 12) also demonstrate generally larger variability in temperatures predicted by HARMONIE-AROME than by the TURCLIM model.

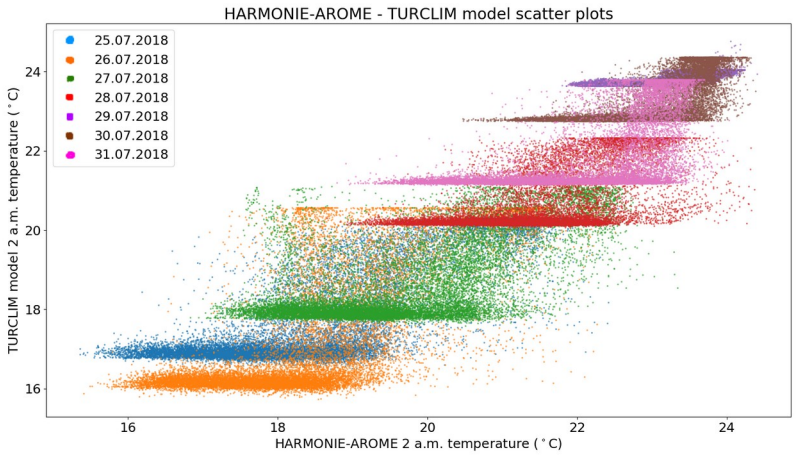

Fig. 11. A scatter plot between 2 am predicted temperatures from HARMONIE-AROME and the TURCLIM model. Each point on the plot represents one 750 m resolution grid cell inside the modeled area. Each color represents a different day in the 25 to 31 July 2018 period.

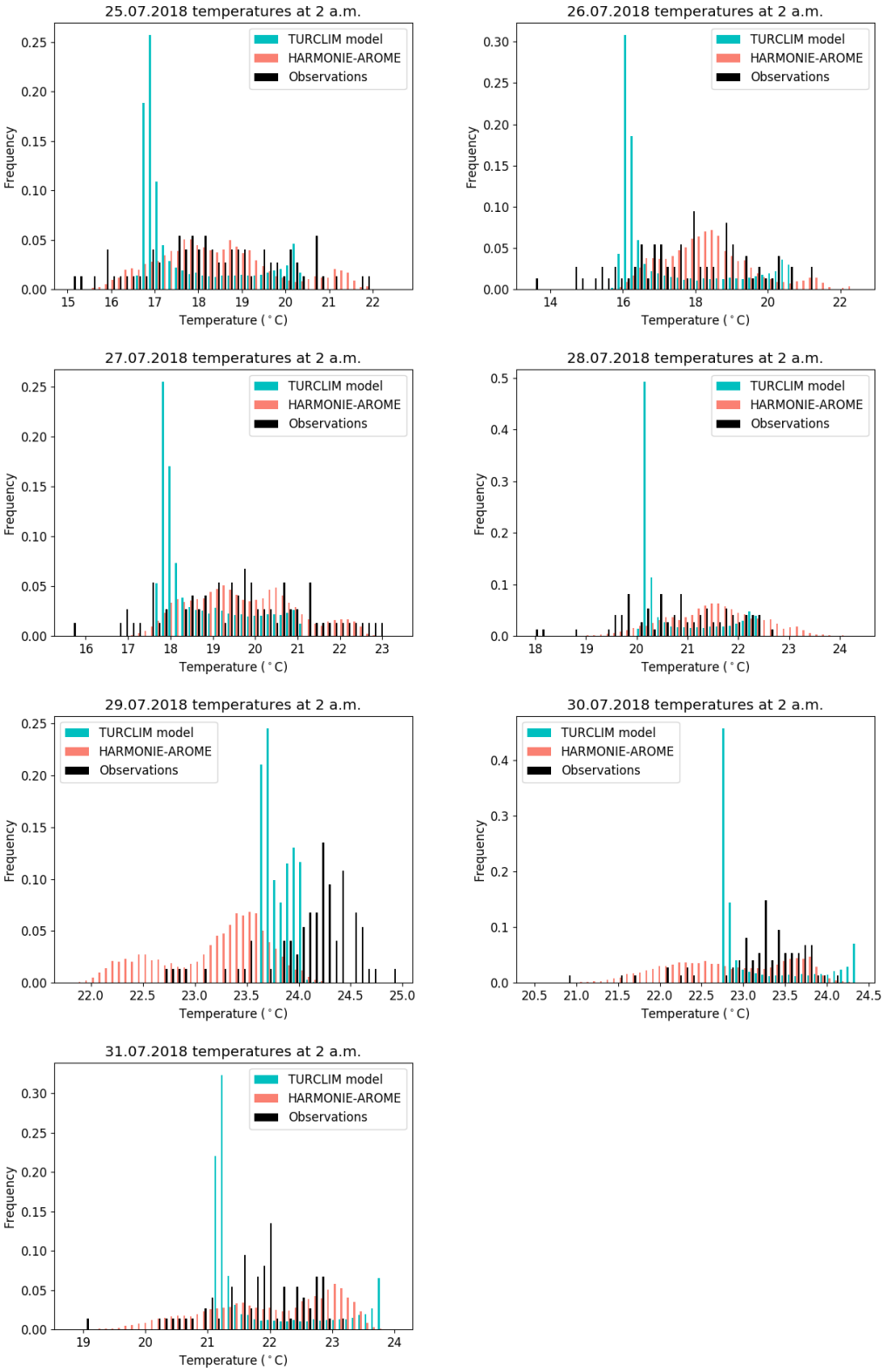

Fig. 12. Temperature distributions of the TURCLIM model (blue), HARMONIE-AROME model (red), and observations (black) at 2 a.m. for each day of the period 25 to 31 July 2018. The modeled datasets consist of 750 m resolution grids resulting in altogether 151 x 101 individual temperature values. The observed temperatures are measured at 74 sites.

d. Spatial pattern of the differences between the temperatures modeled by HARMONIE-AROME model and TURCLIM model

The difference between the temperatures modeled by HARMONIE-AROME and TURCLIM (see Fig. 13) are largest in the Turku city center and in the beginning of the week also in the densely built areas in neighboring municipalities Raisio and Kaarina. In these areas, the HARMONIE-AROME temperatures are warmer than the TURCLIM temperatures. In uninhabited inland areas there is no uniform pattern in the differences in the beginning of the week, but during the windiest nights 29 and 30 July the HARMONIE-AROME temperatures are principally slightly cooler than the TURCLIM based temperatures.

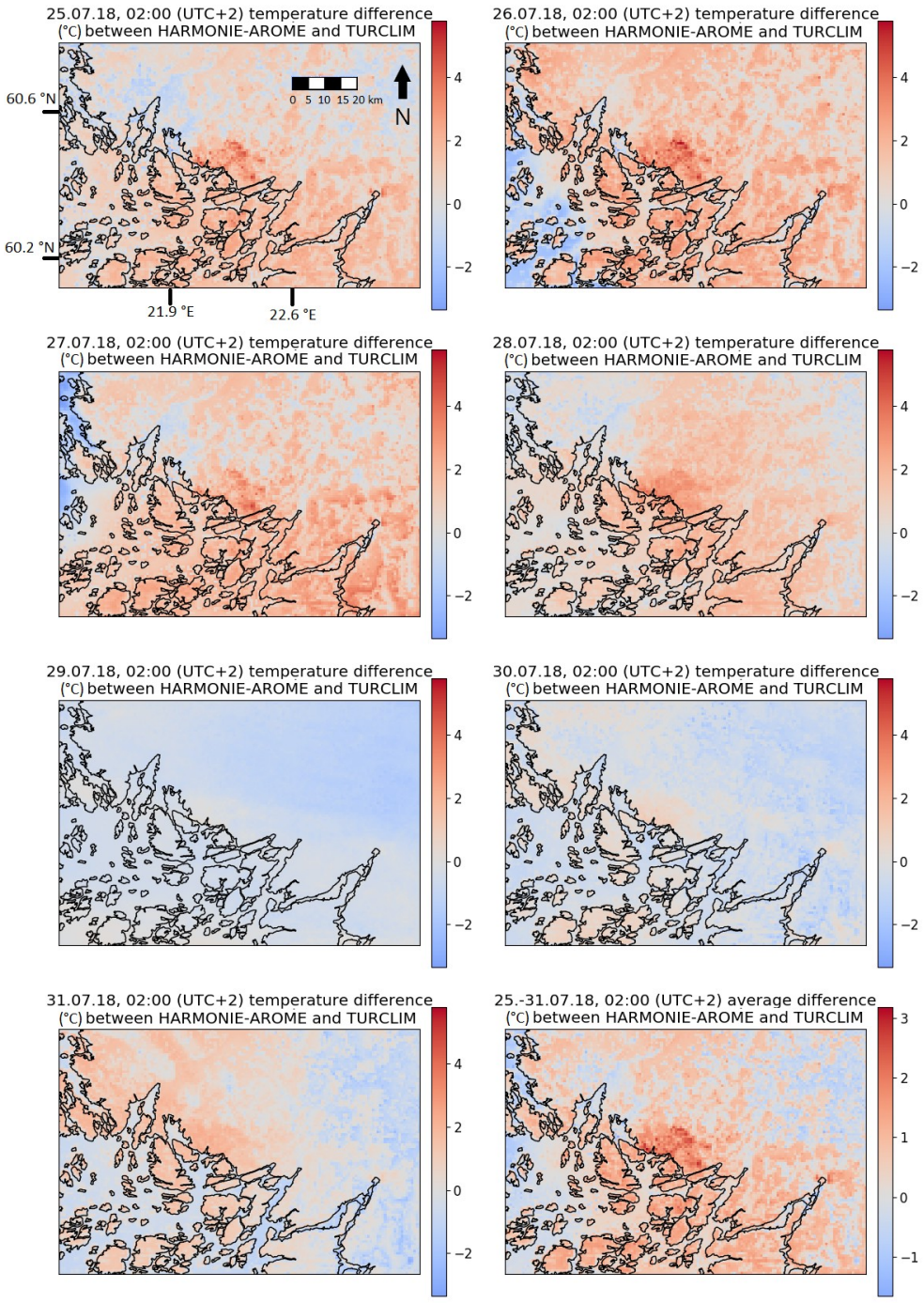

Fig. 13. The differences between 2 am temperatures modeled by HARMONIE-AROME and the TURCLIM model for each day in the period 25 to 31 July 2018. The last figure is the average difference between the modeled 2 am temperatures. Positive values mean that HARMONIE-AROME predicted warmer temperatures. Note that the individual days have identical scales, but the average map has a differing scale for better visualization.

The temperatures modeled by HARMONIE-AROME and TURCLIM were almost the same for the sea tile of the HARMONIE-AROME model indicating only slightly (0.2 °C) warmer HARMONIE-AROME temperatures (Fig. 14). For the lake and nature tiles were

0.49 °C and 0.51 °C, respectively. The clearest difference existed in town tiles, in which the HARMONIE-AROME temperatures were on average 1.4 °C warmer than the respective TURCLIM temperatures. The more detailed comparison of LCZs of the town tile indicates that the difference between the modeled temperatures is smallest in sparsely built areas (LCZ9) and largest in areas dominated by heavy industry (LCZ10). Note that the coverage of some LCZs is quite low (values listed in the figure caption of Fig. 14) and therefore the average values of those covers are less reliable than others, but the trend of the LCZs seem to be quite convergent.

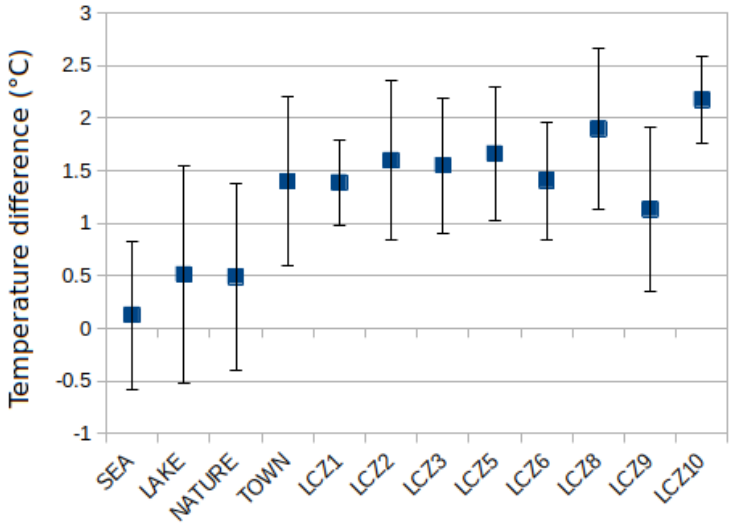

Fig. 14. The average differences of 2 am temperatures between the two models (HARMONIE-AROME minus TURCLIM) in the period 25 to 31 July 2018 for the SEA, LAKE, NATURE and TOWN tiles and the urban cover types. Standard deviations are shown as error bars. The coverage (in grid cells) of different tiles and LCZs in the examined area surrounding Turku are: SEA: 3339.8, LAKE: 45.8, NATURE: 11414.3, TOWN: 248.9, LCZ1: 3.5, LCZ2: 3.4, LCZ3: 1.5, LCZ5: 2.7, LCZ6: 0.3, LCZ8: 82.9, LCZ9: 532.5, LCZ10: 8.0.

5. Discussion

The HARMONIE-AROME and TURCLIM models performed principally well in predicting night-time spatial temperature variability in the study area during a heatwave. The TURCLIM model's R^2 values are, especially during the calm and clear conditions in the beginning of the week, in line with the linear regression models applied elsewhere with a similar kind of model settings (see e.g. Unger, 2006; Szymanowski and Kryza, 2009; Wicki et al., 2018). Despite the generally good performance of the models, at some locations the differences between the observed and modeled temperatures were rather large, especially

during certain weather conditions. In previous studies, similar weather-related variability in the multiple linear regression model performance has also been reported e.g. by Alonso and Renard (2019) and in the HARMONIE-AROME model performance by Sandu et al. (2013)

The largest difference between the observed temperature and modeled temperature by the HARMONIE-AROME model occurred in the observation site 50 (Mylly shopping mall). As shown by Suomi et al. (2024), the same site also had the largest average difference between the observed and modeled hourly temperatures throughout the diurnal cycle during a one-week long observation period. The apparent large warm bias can mostly be explained by the fact that the temperature observation site is in the middle of a grassy patch in an otherwise densely built and asphalted area. Due to the resolution of ECOCLIMAP-SG, HARMONIE-AROME did not have the information about grassy patch (approximately 7% of the total grid cell area), as the respective grid cell is entirely covered by the LCZ8 (large low-rise) land cover.

The two largest differences between the observed and modeled temperature by the TURCLIM model (25 July: 2.35 °C; 26 July: 2.92 °C) occurred at the site 52 (Hiiriluoto, inland). This can probably be explained by the inadequate incorporation of the warming effect of the sea to the north and to the south of the site with the 500 m radius buffer used in the TURCLIM models of 25 July and 26 July. This is supported by the fact that e.g., in the observation site 54 (Hiiriluoto, shore) the respective differences were only 0.54 °C and 0.01 °C. On that seashore site the 500 m buffer was adequate to capture the sea areas immediately to the north-west of the observation site. On the south, the sea is further than in the case of site 52 (Hiiriluoto inland), and respectively, the absence of the sea areas in the southern sector of the 500 m buffer is less relevant for site 54 (Hiiriluoto, shore). On the whole, this is a good example in demonstrating that even if the model has been run with the optimal footprint areas of the explanatory variables regarding the whole observation network, it is probable that for some sites some other scale would have been more appropriate.

For TURCLIM model, the third largest difference between the observed and modeled temperature occurred in two rather different sites, i.e., site 2 (Puolalanmäki) in the city center, and rural site 73 (Karinainen) located approximately 37 kilometers to the NE of the city center. For both sites the modeled temperature was 2.3 °C warmer than the observed one, whereas for site 52 (Hiiriluoto, inland) that had the two largest differences, the modeled temperature had been colder than the observed one. For site 2 (Puolalanmäki), too warm modeled temperature can probably be partly explained by poor representability of the

observation site surroundings in relation to the land cover of the 700 m radius buffer that was used as a source area of the urban land cover variable of the model. Immediately to the W and NW of the observation site is park, whereas over 86 % of the land cover inside the 700 m buffer consists of urban land cover. Consequently, the cooling impact of the park is only partly incorporated in the model. The observation is analogous with that of the HARMONIE-AROME model at the site 50 (Mylly shopping mall). In the case of site 73 (Karinainen), the difference of similar magnitude can possibly be due to regional scale temperature variation that cannot be captured by the TURCLIM model.

For both of the models, the differences between the observed and modeled temperatures were largest in the beginning of the week when clear skies and weak winds favored the formation of a strongly stratified nocturnal boundary layer, and the spatial temperature differences in the area were at their largest. This intra-week variation in the difference between the observed and modeled temperature was more pronounced for the HARMONIE-AROME model than for the TURCLIM model, and is related to the tendency of HARMONIE-AROME to overestimate the turbulent mixing and consequently surface air temperature during conditions favoring strong stratification. The overestimation is not present in the model during the cloudy and windy nights of 29 and 30 July. This weather-related dependency of model accuracy is in line with earlier observations that stable atmospheric stratification and related calm or weak wind conditions are challenging for the HARMONIE-AROME model and for numerical weather prediction models in general (Sandu et al., 2013; Kalverla et al., 2019; Sekula et al., 2019). Even though the differences between the observed and modeled temperatures were largest during the calm weather, the explanatory power of the TURCLIM model was also largest during those conditions. The observation is analogous to that of Straub et al. (2019), who studied the performance of linear regression model and random forests method in modeling the UHI of Augsburg, Germany, and to that of Coseo and Larsen (2014), who used linear regression model in studying the impacts of land use / land cover and other relevant factors on the UHI of Chicago, the United States. Better explanatory power during calm or weak wind conditions is logical in the sense that during windier weather the spatial temperature differences are diminished by horizontal mixing, and consequently, the impact of explanatory variables such as land cover, topography and water bodies are less clearly reflected by temperatures (see e.g. Wicki et al., 2018).

One difference between the HARMONIE-AROME and TURCLIM model is that the TURCLIM model predicts rather similar temperatures for unbuilt or sparsely built inland

areas, whereas for the HARMONIE-AROME model the respective variation is larger. The TURCLIM model has variables for land cover, topography and water bodies. Therefore, as in the unbuilt and sparsely built inland areas the topographical variation is small, the land cover is rather homogeneous dominated by fields and forests, and there are no large water bodies, it is only logical that the spatial differences in predicted temperature are negligible and clearly smaller than in the coastal areas and topographically varying Turku city center. For the similar kind of reasons, the spatial variability of TURCLIM model based temperatures is small or even absent also in the sea areas, where topographic differences and urban land cover may be totally lacking also inside the buffer zone distances that have been used in model calibration. In addition to the small spatial variability, the temperatures predicted by the TURCLIM model may also have some warm bias in the open sea areas, as the model has been calibrated based on the observations recorded in land areas where the proportion of water bodies around the observation sites is zero or remarkably lower than in the open sea areas. Furthermore, the relatively shallow coastal waters near the observation sites reach higher temperatures during the heatwaves than in the outer archipelago. Consequently, the warming impact of water bodies may be overestimated in the open sea areas where the water temperature is lower than in the coastal zone.

In addition to the higher resolution of the TURLIM model, one potential factor for the models' slightly better performance in general could be the calibration of the TURCLIM model with the temperature observations of the study area. Even though the model accuracy analyses are performed with a separate validation dataset, the model calibration with the actual on-site temperatures succeeds to incorporate the effects of relevant environmental factors rather well especially in the study area in question.

Based on the results of this study, even though both models generally perform rather well in an urban environment, the models also call for further development. A night-time warm bias is a general feature of HARMONIE-AROME noted by forecasters and is clearly visible here. Similar modeling results with the MUKLIMO_3 urban climate model were found by Hürzeler et al. 2022, who reported a bias of 0.2 to 0.7 K at the lowest air temperatures at 7 a.m., while a larger bias of 1.5 to 2.8 K was found at midnight. On the other hand, WRF has been found (Hu et al. 2022) to model colder (daytime) near-surface temperatures than what was observed in Shanghai. In Suomi et al. (2024) it was noted that at other times of day, the model performance was significantly better during this same 2018 July heatwave. The

HARMONIE-AROME model's relatively coarse spatial resolution is a challenge in heterogeneous areas in which the temperature can vary remarkably within a small area. This pitfall could in many cases be overcome or eased by higher resolution.

Another angle towards more accurate numerical weather prediction is developing more elaborate physical models. For example, TEB is a single layer urban canopy model, and thus the model does not resolve vertical structures. Schoetter et al. (2020) points out a comprehensive list of uncertainties related to the absence of vertical resolution in SURFEX-TEB. For example, the building height being constant within a grid cell can cause inaccuracies in the wind speed prediction.

As for the TURCLIM model, a linear regression based model is deficient in detecting potential non-linear relationships between the explanatory and response variables that could be better taken into account with some non-linear modeling method (see e.g. Equere et al., 2020; Oukawa et al., 2022). Alternatively, better inclusion of those dependencies in the TURCLIM model could be performed with transformations (e.g. logarithmic transformation, square root transformation) for explanatory variables.

The TURCLIM model could also be tried to develop by including more explanatory variables in the model. This may, however, be problematic from a viewpoint of multicollinearity, as the central causes of UHI are probably mostly captured with the contemporary variables (see e.g. Shaker et al., 2019; Chen et al., 2020). Regarding the land cover variable applied, the fine-scale variability in heat storage capacity and anthropogenic heat release of the buildings in urban areas could potentially be more accurately modeled e.g., with the floor area or the volume of the buildings, or with the direct data on anthropogenic heat release (see e.g. Alonso and Renard, 2019), but as far as we authors know, this information is not as broadly, easily and freely available as the European-wide CLC dataset that forms the basis of the current land cover variable. The broad availability enhances the usability of the model also beyond the current study area, and additionally, the CLC based variable also covers the flat asphalted surfaces that have good thermal conductivity and heat storage capacity and that would not be covered by the floor area or building volume variables (see Suomi et al., 2012). Even though the urban land cover variable of the TURCLIM model captures generally well the heat of the urban-type land cover in the area, the region to the south-west of the city center forms an exception; the area that is defined as 'Port areas' in the CLC2018 classification, exists only in the surroundings of 2 out of 71 observation sites used in the model calibration. Consequently, the 'Port areas' class did not stand out from other CLC2018 classes in the preliminary correlation analyses, based on which the urban land cover variable was principally formulated, and was thus not included in the regression model.

This resulted in systematically too cold modeled temperature for these 2 observation sites that actually have urban type land cover, such as asphalted surfaces and relatively large buildings in their surroundings. This is a good example of a challenge in incorporating in the TURCLIM model all relevant environmental factors of certain observation sites that have some specific characteristics compared to the other sites. The challenge could be tried to overcome by checking e.g. from the aerial photo the actual land cover characteristics of the areas which are covered by land cover class that might have local effect on temperature but is only poorly present, and depending on the check, include the respective land cover class as part of the existing variable representing urban land cover.

6. Conclusions

This paper compares the performances of numerical weather prediction model HARMONIE-AROME and multiple linear regression based TURCLIM model in modeling the night-time temperatures and spatial temperature differences in the high-latitude coastal city of Turku, Finland, during a summer-time heatwave. The modeling accuracy was assessed against the temperature observations of the TURCLIM local climate observation network. The results indicate slightly better modeling accuracy for the TURCLIM model with an average MAE of 0.56 °C and with a warm bias of 0.02 °C against HARMONIE-AROME's average MAE of 1.47 °C and warm bias of 1.29 °C. Largest single site-specific differences between the observed and modeled temperatures occurred for the both models during the three first nights of the one-week long study period, when the weather was calm and cloudless, and the spatial temperatures differences in the study area were largest. This weather-related variability in modeling accuracy was clearer for the HARMONIE-AROME model, and moreover, under calm and cloudless circumstances, temperatures modeled by HARMONIE-AROME were systematically warmer than the observed temperatures, whereas for the TURCLIM model, both cold and warm biases occurred. Regarding spatial characteristics of model performances, the largest differences between the observed and modeled temperatures were not clearly concentrated in any certain areas but were mostly related to the site-specific characteristics of the observation sites. For both models, large differences occurred on sites in which the immediate neighborhood of the observation site deviated clearly from the principal land cover characteristics of the area, as demonstrated in the Discussion section by the cases of Mylly shopping mall and Puolalanmäki. With the

HARMONIE-AROME, the problem could be approached e.g., with higher resolution physiographic description and with more accurate physical modeling, and with the TURCLIM model e.g., by testing distance-based weighting of explanatory variables In the TURCLIM model, an insufficient incorporation of relevant environmental factors at some observation sites caused systematic cold or warm bias in those areas. This could potentially be tackled by additional explanatory variables or by merging the rare land cover classes that are relevant for spatial temperature differences to the classes that are more widely present and that are used as part of the explanatory variables.

The comparison of the modeled temperatures over the whole study area (≈8600 km2) reveals that both models predict spatial temperature variability in urban and coastal areas. The HARMONIE-AROME model predicts larger spatial temperature variability for the rural areas than the TURCLIM model. The temperatures modeled by the HARMONIE-AROME model were on average warmer than the respective temperatures modeled by the TURCLIM model. For the tiles SEA, NATURE, LAKE, and TOWN, the average differences were 0.12 °C, 0.49 °C, 0.51 °C and 1.40 °C, respectively.

Acknowledgments.

We thank the Geography Section of the University of Turku and Urban Environment Division of the City of Turku for financial support in maintaining the TURCLIM weather observation network. We thank the Finnish Meteorological Institute for open access weather data. The HARMONIE-AROME model integrations were carried out using the computing resources of the ECMWF. Financial support for all authors for this research was granted by the Academy of Finland (the consortium acronym HERCULES, decision numbers 329235 and 329241).

Data Availability Statement.

The authors are aware of the AMS data and software policies. The data used and derived in this study will be available for public use when the work is published.

APPENDIX

Observation site land cover data

Table A1. Land cover information of the observation site surroundings. The ECOCLIMAP is used as a physiographic database in the HARMONIE-AROME model. The CLC based urban land cover is used as one of the explanatory variables in the TURCLIM

model. For locations of the observation sites, see Figure 2. Common names have been translated in English, while proper nouns are in their original Finnish form.

Obs. site number	Obs. site name	The dominant ECOCLIMAP cover(s) inside a 750 m x 750 m grid cell (LCZ = Local Climate Zone)	Coverage (% of the 750 m x 750 m grid cell area)	Coverage (%) of the CLC based urban land cover inside a 700 m radius buffer
1	Market place	LCZ2: compact midrise	40	83.6
2	Puolalanmäki	LCZ2: compact midrise	40	86.6
3	Puutori	LCZ8: large low-rise	86.7	79.2
4	Betel	LCZ2: compact midrise	40	83.5
5	Virastotalo	LCZ9: sparsely built	50	67.2
6	Piispankatu	LCZ8: large low-rise	86.7	72.4
7	Railway station	LCZ8: large low-rise	60	62.8
8	University campus	LCZ8: large low-rise	40	73.3
9	Kerttuli	LCZ9: sparsely built	80	71.8
10	Mikaeli church	LCZ9: sparsely built	50	64.5
11	Sports park	LCZ9: sparsely built	50	59
12	Kähäri	LCZ8: large low-rise	60	43.9
13	Saarnitie	LCZ9: sparsely built	40	56.6
14	Sirkkala	LCZ8: large low-rise	40	70.7
15	Martti	LCZ8: large low-rise	53.8	58.1
16	Rieskalähde	LCZ9: sparsely built	86.7	20.7
17	Kupittaa park	LCZ9: sparsely built	80	48.5
18	Alfa sports center	LCZ9: sparsely built	100	25.1
19	Kakola	LCZ10: heavy industry	53.3	44.5
20	Student village	LCZ9: sparsely built	90	32.4
21	Uudenmaantie	LCZ8: large low-rise, LCZ9: sparsely built	50	36.7
22	Nummi	LCZ9: sparsely built	90.9	19.9
23	Halinen	LCZ9: sparsely built	66.7	21.4
24	Luolavuori	LCZ9: sparsely built	100	19.9
25	Turku castle	LCZ9: sparsely built	50	19.4
26	Suikkila	LCZ9: sparsely built	100	23
27	Heikkilä barracks	winter C3 crops	30	19.6
28	Ispoinen	LCZ9: sparsely built	90	21.7
29	Valkiasvuori	LCZ9: sparsely built	100	25
30	Vapaavarasto	LCZ10: heavy industry	70	17.8
31	Impivaara	LCZ9: sparsely built	70	22.6
32	Liponkuja	LCZ9: sparsely built	93.3	16.7
33	Cemetary	LCZ9: sparsely built	86.7	29.2
34	Pääskyvuori, valley	LCZ9: sparsely built	80	24.8
35	Räntämäki	LCZ9: sparsely built	64.3	15.6
36	Metsäkylä	LCZ9: sparsely built	90	24.1
37	Marjaniemi	sea and oceans	40	10.1

38	Peltola allotment	LCZ9: sparsely built	100	30.7
39	Pääskyvuori	LCZ9: sparsely built	100	14.1
40	Runosmäki	boreal needleleaf evergreen	66.7	46.5
41	Kurala	LCZ9: sparsely built	70	17.4
42	Huhkola	LCZ9: sparsely built	70	20.2
43	Vahdontie	LCZ8: large low-rise	55.6	33.8
44	Exhibition center	LCZ9: sparsely built, LCZ10: heavy industry	50	30.5
45	Katariina	winter C3 crops	40	8
46	Hirvensalo	winter C3 crops	46.7	1.5
47	Varissuo	LCZ9: sparsely built	86.7	36
48	Ruissalo	winter C3 crops	100	0.5
49	Botanical garden	winter C3 crops	53.3	0.6
50	Mylly shopping mall	LCZ8: large low-rise	100	36.3
51	Perno	LCZ9: sparsely built	80	24.7
52	Hiiriluoto, inland	winter C3 crops	50	1.7
53	Hiiriluoto, hill	winter C3 crops	50	0.3
54	Hiiriluoto, shore	temperate broadleaf deciduous	33.3	0.2
55	Rauhaniemi	lakes	55.6	8.5
56	Raisio	LCZ9: sparsely built	81.8	48.2
57	Pansio	sea and oceans	40	3.7
58	Vanhalinna	winter C3 crops	70	3.1
59	Metsämäki	winter C3 crops	50	3.4
60	Kaarina	LCZ9: sparsely built	60	24.1
61	Satava	sea and oceans	46.7	2.9
62	Camping site	boreal needleleaf evergreen	50	4.5
63	Ylijoki	winter C3 crops	100	3.7
64	Niuskala	boreal needleleaf evergreen	53.3	1.2
65	Kakskerta	winter C3 crops	40	0.4
66	Kuuva	sea and oceans	100	0
67	Kolkka	sea and oceans	91.7	4.2
68	Tuorla	LCZ9: sparsely built	53.3	11.7
69	Jäkärlä	LCZ9: sparsely built	93.3	23.9
70	Lieto	winter C3 crops	90	0.5
71	Sikilä	winter C3 crops	80	6.8
72	Kuusinen	sea and oceans	93.3	0
73	Karinainen	boreal needleleaf evergreen, winter C3 crops	33.3	0.3
74	Kirjainen	temperate needleleaf evergreen	53.8	0

REFERENCES

Alonso L, Renard F., 2019. Integrating Satellite-Derived Data as Spatial Predictors in Multiple Regression Models to Enhance the Knowledge of Air Temperature Patterns. Urban Science 3(4): 101. https://doi.org/10.3390/urbansci3040101

Bauer, P., Thorpe, A. & Brunet, G., 2015. The quiet revolution of numerical weather prediction. Nature 525, 47–55. https://doi.org/10.1038/nature14956

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., Rooy, W.C., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K., Lenderink, G., Niemelä, S., Nielsen, K.P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D.S., Subias, Á., Tijm, S., Toll, V., Yang, X., & Køltzow, M.A., 2017. The HARMONIE-AROME Model Configuration in the ALADIN-HIRLAM NWP System. Monthly Weather Review, 145, 1919-1935.

Coseo, P., Larsen, L., 2014. How factors of land use/land cover, building configuration, and adjacent heat sources and sinks explain Urban Heat Islands in Chicago. Landscape and Urban Planning 125: 117-129. https://doi.org/10.1016/j.landurbplan.2014.02.019.

Bokwa, A., Hajto, M.J., Walawender, J.P., Szymanowski, M., 2015. Influence of diversified relief on the urban heat island in the city of Kraków, Poland. Theor Appl Climatol 122: 365–382. DOI https://doi.org/10.1007/s00704-015-1577-9

Bottyán, Z., Unger, J., 2003. A multiple linear statistical model for estimating the mean maximum urban heat island. Theor Appl Climatol 75: 233–243. https://doi.org/10.1007/s00704-003-0735-7

Camilloni, I., Barrucand, M., 2012. Temporal variability of the Buenos Aires, Argentina, urban heat island. Theor Appl Climatol 107: 47–58 . https://doi.org/10.1007/s00704-011-0459-z

Chen, C. Y., Wu, C. D., Ng, U. C., & Lung, S. C. C., 2020. Characterize the spatial-temporal variations of urban heat island intensity using a land use regression approach. Paper presented at 40th Asian Conference on Remote Sensing: Progress of Remote Sensing Technology for Smart Future, ACRS 2019, Daejeon, Korea, Republic of.

City of Turku, 2022. Turku map service, accessed 29 Oct 2022. https://opaskartta.turku.fi

CNRM, 2017. Surfex V8_1, accessed 17 Aug 2023. https://www.umr-cnrm.fr/surfex/spip.php?rubrique151

CNRM, 2018. ECOCLIMAP-SG, accessed 17 Aug 2023. https://opensource.umr-cnrm.fr/projects/ecoclimap-sg/

Dai, X., Guldmann, J.-M., Hu, Y. 2018. Spatial regression models of park and land-use impacts on the urban heat island in central Beijing. Science of The Total Environment 626: 1136-1147. https://doi.org/10.1016/j.scitotenv.2018.01.165.

Earl, N., Simmonds, I., Tapper, N., 2016. Weekly cycles in peak time temperatures and urban heat island intensity. Environ. Res. Lett. 11 074003. https://doi.org/10.1088/1748-9326/11/7/074003

Eliasson, I., 1992. Infrared thermography and urban temperature patterns. Int J Remote Sens 13: 869–879. https://doi.org/10.1080/01431169208904160

Equere, V., Mirzaei, P.A., Riffat, S., 2020. Definition of a new morphological parameter to improve prediction of urban heat island. Sustainable Cities and Society 56: 102021. https://doi.org/10.1016/j.scs.2020.102021

FMI, 2022. Ilmasto-opas.fi, Finnish Meteorological Institute. Accessed 7 Oct 2022. https://ilmasto-opas.fi

Foissard, X., Dubreuil, V., Quénol, H., 2019. Defining scales of the land use effect to map the urban heat island in a mid-size European city: Rennes (France). Urban Climate 29: 100490. https://doi.org/10.1016/j.uclim.2019.100490.

Guerri, G., Crisci, A., Morabito, M. 2023. Urban microclimate simulations based on GIS data to mitigate thermal hot-spots: Tree design scenarios in an industrial area of Florence. Building and Environment 245: 110854. https://doi.org/10.1016/j.buildenv.2023.110854.

Hajat, S., Kosatky, T., 2010. Heat-related mortality: a review and exploration of heterogeneity. Journal of Epidemiology & Community Health 64: 753-760. http://dx.doi.org/10.1136/jech.2009.087999

Hamdi, R., 2010. Estimating Urban Heat Island Effects on the Temperature Series of Uccle (Brussels, Belgium) Using Remote Sensing Data and a Land Surface Scheme. Remote Sensing 2, no. 12: 2773-2784. https://doi.org/10.3390/rs2122773

Hamdi, R., Degrauwe, D., Termonia, P., 2012. "Coupling the Town Energy Balance (TEB) Scheme to an Operational Limited-Area NWP Model: Evaluation for a Highly Urbanized

Area in Belgium", Weather and Forecasting 27, 2 : 323-344, accessed Dec 14, 2022, https://doi.org/10.1175/WAF-D-11-00064.1

Hamdi, R., Kusaka, H., Doan, QV. et al., 2020 The State-of-the-Art of Urban Climate Change Modeling and Observations. Earth Syst Environ 4, 631–646. https://doi.org/10.1007/s41748-020-00193-3

Hu, Y., Tan, J., Grimmond, S., Ao, X., Yan, Y., Liu, D., 2022. Observed and Modeled Urban Heat Island and Sea-Breeze Circulation Interactions: A Shanghai Case Study. J. Appl. Meteor. Climatol., 61, 239–259. https://doi.org/10.1175/JAMC-D-20-0246.1.

Hürzeler, A., Hollósi, B., Burger, M., Gubler, M., Brönnimann, S., 2022. Performance analysis of the urban climate model MUKLIMO_3 for three extreme heatwave events in Bern. City and Environment Interactions, Volume 16: 100090. https://doi.org/10.1016/j.cacint.2022.100090.

Jokinen, P., Pirinen, P., Kaukoranta, J.-P., Kangas, A., Alenius, P., Eriksson, P. Johansson, M., Wilkman, S., 2021. Tilastoja Suomen ilmastosta ja merestä 1991-2020 (Climatological and oceanographic statistics of Finland 1991–2020). Raportteja 2021:8. Accessed Nov 13, 2022. https://doi.org/10.35614/isbn.9789523361485

Kalverla, P., Steeneveld, G.-J., Ronda, R., Holtslag, A.A.M., (2019). Evaluation of three mainstream numerical weather prediction models with observations from meteorological mast IJmuiden at the North Sea. Wind Energy 22, 34–48. https://doi.org/10.1002/we.2267

Ketterer, C., Matzarakis, A., 2015. Comparison of different methods for the assessment of the urban heat island in Stuttgart, Germany. Int J Biometeorol 59: 1299–1309. https://doi.org/10.1007/s00484-014-0940-3.

Kim, Y.-H., Baik, J.-J., 2002. Maximum Urban Heat Island Intensity in Seoul. Journal of Applied Meteorology (1988-2005) 41(6): 651–659. http://www.jstor.org/stable/26185006

Kivimäki, M., Batty, D.G., Pentti, J., Suomi, J., Nyberg, S.T., Merikanto, J., Nordling, K., Ervasti, J., Suominen, S.B., Partanen, A.-I., Stenholm, S., Käyhkö, J., Vahtera, J., 2023. Climate Change, Summer Temperature, and Heat-related Mortality in Finland: Multicohort Study with Projections for a Sustainable Versus Fossil-fueled Future to 2050. Environmental Health Perspectives 131: 12. https://doi.org/10.1289/EHP12080

Kollanus, V., Tiittanen, P., Lanki, T., 2021. Mortality risk related to heatwaves in Finland - Factors affecting vulnerability. Environmental Research 201: 111503. https://doi.org/10.1016/j.envres.2021.111503.

Masson, V., 2000. A Physically-Based Scheme For The Urban Energy Budget In Atmospheric Models. Bound.-Layer Meteorol. 94, 357–397. https://doi.org/10.1023/A:1002463829265.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., Voldoire, A., 2013. The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geosci. Model Dev. 6, 929–960. https://doi.org/10.5194/gmd-6-929-2013.

National Land Survey of Finland, 2019. Elevation model 2019, 10 m x 10 m. CSC – IT Center for Science. http://urn.fi/urn:nbn:fi:csc-kata00001000000000000622

National Land Survey of Finland, 2020. SLICES 2010, 10 m x 10 m, generalized raster, ETRS-TM35FIN. CSC - IT Center for Science Ltd. http://urn.fi/urn:nbn:fi:csc-kata00001000000000000262

Noilhan, J. and Planton, S., 1989. "A Simple Parameterization of Land Surface Processes for Meteorological Models", Monthly Weather Review 117, 3: 536-549, https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2

Oke T.R. and Maxwell G.B., 1975. Urban heat island dynamics in Montreal and Vancouver. Atmospheric Environment 9: 191–200.

Oke T.R., 1987. Boundary Layer Climates. 2nd edition. 435 s. Routledge, London.

Oke T.R. (2006) Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites. Instruments and Observing Methods Report No. 81. World Meteorological Organization, Geneva.

Oukawa, G.Y., Krecl, P., Targino, A.C., 2022. Fine-scale modeling of the urban heat island: A comparison of multiple linear regression and random forest approaches. Science of The Total Environment 815: 152836. https://doi.org/10.1016/j.scitotenv.2021.152836.

Porangaba, G.F.O., Teixeira, D.C.F., Amorim, M., da Silva, M.H.S., Dubreuil, V., 2021. Modeling the urban heat island at a winter event in Três Lagoas, Brazil. Urban Climate 37: 100853. https://doi.org/10.1016/j.uclim.2021.100853

Ronda, R. J., Steeneveld, G. J., Heusinkveld, B. G., Attema, J. J., Holtslag, A. A. M., 2017. Urban Finescale Forecasting Reveals Weather Conditions with Unprecedented Detail. Bull. Amer. Meteor. Soc., 98, 2675–2688, https://doi.org/10.1175/BAMS-D-16-0297.1.

Ruuhela, R., Votsis, A., Kukkonen, J., Jylhä, K., Kankaanpää, S., Perrels, A., 2021. Temperature-Related Mortality in Helsinki Compared to Its Surrounding Region Over Two Decades, with Special Emphasis on Intensive Heatwaves. Atmosphere 12 (1): 46. https://doi.org/10.3390/atmos12010046.

Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., and Balsamo, G., 2013. Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models?. J. Adv. Model. Earth Syst. 5, 117–133. https://doi.org/10.1002/jame.20013.

Santamouris, M., 2014. On the energy impact of urban heat island and global warming on buildings. Energy and Buildings 82: 100-113. https://doi.org/10.1016/j.enbuild.2014.07.022.

Schatz, J., Kucharik, C.J., 2016. Urban heat island effects on growing seasons and heating and cooling degree days in Madison, Wisconsin USA. Int. J. Climatol., 36: 4873-4884. https://doi.org/10.1002/joc.4675

Schoetter, R., Kwok, Y. T., de Munck, C., Lau, K. K. L., Wong, W. K., Masson, V., 2020. Multi-layer coupling between SURFEX-TEB-v9.0 and Meso-NH-v5.3 for modelling the urban climate of high-rise cities, Geosci. Model Dev., 13, 5609–5643. https://doi.org/10.5194/gmd-13-5609-2020.

Sekula, P., Bokwa, A., Bochenek, B., Zimnoch, M., 2019. Prediction of Air Temperature in the Polish Western Carpathian Mountains with the ALADIN-HIRLAM Numerical Weather Prediction System. Atmosphere 10, 186. https://doi.org/10.3390/atmos10040186.

Shaker, R.R., Altman, Y., Deng, C., Vaz, E., Forsythe, K.W., 2019. Investigating urban heat island through spatial analysis of New York City streetscapes. Journal of Cleaner Production 233: 972-992. https://doi.org/10.1016/j.jclepro.2019.05.389.

Stewart, I. D., Oke, T. R., 2012. Local Climate Zones for Urban Temperature Studies. Bull. Amer. Meteor. Soc., 93, 1879–1900, https://doi.org/10.1175/BAMS-D-11-00019.1.

Straub, A., Berger, K., Breitner, S., Cyrys, J., Geruschkat, U., Jacobeit, J., Kühlbach, B., Kusch, T., Philipp, A., Schneider, A., Umminger, R., Wolf, K., Beck, C., 2019. Statistical modeling of spatial patterns of the urban heat island intensity in the urban environment of Augsburg, Germany. Urban Climate 29, 100491. https://doi.org/10.1016/j.uclim.2019.100491.

Suomi, J., 2018. Extreme temperature differences in the city of Lahti, southern Finland: Intensity, seasonality and environmental drivers. Weather and Climate Extremes 19: 20–28. DOI: 10.1016/j.wace.2017.12.001

Suomi, J., Hjort, J., Käyhkö, J., 2012. Effects of scale on modelling the urban heat island in Turku, SW Finland. Climate Research 55: 121-136. https://doi.org/10.3354/cr01123

Suomi, J., Käyhkö, J., 2012. The impact of environmental factors on urban temperature variability in the coastal city of Turku, SW Finland. International Journal of Climatology 32: 451–463. DOI: 10.1002/joc.2277

Suomi, J., Saranko, O., Partanen, A.-I., Fortelius, C., Gonzales-Inca, C., Käyhkö, J., 2024. Evaluation of surface air temperature in the HARMONIE-AROME weather model during a heatwave in the coastal city of Turku, Finland. Urban Climate, Volume 53, 101811. https://doi.org/10.1016/j.uclim.2024.101811

Szymanowski, M., Kryza, M., 2009. GIS-based techniques for urban heat island spatialization. Climate Research 38(2), 171–187. http://www.jstor.org/stable/24870407

SYKE, 2023. Corine maanpeite 2018. < https://ckan.ymparisto.fi/en/dataset/corine-maanpeite-2018> Accessed 17.8.2023.

Unger, J., 2006. Modelling of the annual mean maximum urban heat island using 2D and 3D surface parameters. Climate Research 30(3): 215–226. http://www.jstor.org/stable/24869249

Wicki A, Parlow E, Feigenwinter C., 2018. Evaluation and Modeling of Urban Heat Island Intensity in Basel, Switzerland. Climate 6(3):55. https://doi.org/10.3390/cli6030055

WMO, 2023. Guidance on Measuring, Modelling and Monitoring the Canopy Layer Urban Heat Island (CL-UHI). 2023 edition. WMO-No. 1292. World Meteorological Organization, Geneva. https://library.wmo.int/idurl/4/58410

Yang, X., Peng, L.H.H., Jiang., Chen, Y., Yao, L., He, Y., Xu, T., 2020. Impact of urban heat island on energy demand in buildings: Local climate zones in Nanjing. Applied Energy 260, 114279. https://doi.org/10.1016/j.apenergy.2019.114279