S1. Technical description of the WMO stations, the UHIARC network stations, and other complementary in-situ observations

The AWS *Davis Vantage Pro2* is a popular, stable, and accurate meteorological station with a temperature measurement accuracy of 0.5 K (Bell et al., 2015) Two AWS were installed in the city center (U1) and 15 km to the west on the shore of lake

5   Imandra (R0) in 2014. In 2015, AWS were installed in the same place in the city centre and three additional AWSs were deployed in rural sites around the city. Among them, one AWS was deployed on the coast of the lake Imandra. It was collocated with the WMO R1 station (the WMO station index is 22213). Two other AWSs were deployed at the forest sites located to the north-west (R2) and the north-east (R3) of the city.

iButton temperature sensors (reported accuracy 0.5 K) are also popular for the spatially resolved urban climate studies

10 (Malevich and Klink, 2011; Ojeh et al, 2016), but they require careful radiation shielding. However, the direct solar radiation is not a problem for our study since the sun was mostly under the horizon during the winter months.

During the winter of 2013-2014, 18 *iButton* sensors were deployed at a height of 2 m above the ground in the city area. Each sensor was insulated with a rubber membrane and attached by a metal wire to extended tree branches to avoid possible radiative distortion of the temperature readings. In 2015, the *iButton* sensor network was grown to a total of 35 sensors covering both

15 the residential areas and the industrial complex to the north of the city. They were distributed around different building types: in the campus – low widely spaced buildings in the natural landscape; in the central district – middle-height administrative buildings at Lenin Square and tall buildings surrounded by paved surfaces at Geologists Square; in the city's southern part – low industrial buildings and urbanized landscapes (railway station etc.); and in the industrial zones – widely spaced low and tall structures. All sensors were set up to read the temperature at 10 minute intervals. During the short-period campaigns each

20 AWS was co-located with an *iButton* sensor to control data quality. Intercomparison of these two types of observations confirmed that the temperature records were unbiased and maintained their respective accuracy margins.

For the longer period of observations in the winter of 2015-2016, two AWSs were installed at the U1 and R1 sites. In addition, two *iButton* loggers were installed in the city (U2 and U3) and another two in rural areas to the north-east (R4) and south-east (R5) of the city. The data from the existing AWS at the R2 site were also available, but unfortunately there were many gaps in

25 the observational record. All these observations were recorded at one-hour time intervals. The data are available since 10.12.2015 up to the end of winter for all stations with exception of the R5 site, where observations ended 08.02.2016. Thus, in addition to the whole winter period (10.12.2015 – 01.03.2016), we have selected a shorter period from 10.12.2015 to 08.02.2016 for the analysis of the synchronous observations of all AWS and temperature sensors. In order to avoid possible effects of solar heating on loggers during the short periods of the sunlight, the iButton readings were processed to remove

30 times when the solar elevation was more than 2°, which covers about 5% of the selected period.

Exact location of the listed sites of observations during all three campaigns, elevation above the sea level and corresponding types of the Local Climate Zones are given in Table S1.

References for S1:

Bell, S., Cornford, D. and Bastin, L.: How good are citizen weather stations? Addressing a biased opinion, Weather, 70, 75-84, https://doi.org/10.1002/wea.2316, 2015

Malevich, S. B. and Klink, K.: Relationships between snow and the wintertime Minneapolis urban heat island, J. Appl.
5   Meteorol. Climatol., 50, 1884–1894, https://doi.org/10.1175/JAMC-D-11-05.1, 2011.

Ojeh, V., Balogun, A. and Okhimamhe, A.: Urban-Rural Temperature Differences in Lagos, Climate, 4, 29, https://doi.org/10.3390/cli4020029, 2016.

Table S1. In-situ observational network summary: data, coordinates and other metadata.

Station name / iButton ID	LCZ	Lat. °N	Lon. °E	Elevation (from ASTER)	Source of the data & remarks
Short experiment campaign in 2014 (28.01.2014 – 02.02.2014)					
R0 / ib01	LCZ B	67.6001	33.0013	128	1) AWS 2) iButton
					1) WMO station data downloaded from www.rp5.ru (6-hour resolution);
R1	LCZ 9/B	67.5508	33.3621	132	2) WMO station data with detailed 1-min resolution. This data are not publicly available.
U1 / ib02	LCZ 5	67.5683	33.4063	180	1) AWS 2) iButton
ib10	LCZ 6	67.5704	33.4031	176	iButton
ib01	LCZ 5	67.5686	33.3951	168	iButton
ib12	LCZ 2	67.5625	33.3878	151	iButton
ib13	LCZ 2/5	67.5624	33.4007	166	iButton
ib14	LCZ 5	67.5662	33.3988	159	iButton
ib15	LCZ 2/5	67.5637	33.4121	180	iButton
ib16	LCZ 5	67.5581	33.4101	171	iButton
ib17	LCZ 2/5	67.5589	33.4186	172	iButton
ib18	LCZ 5	67.5655	33.4256	182	iButton
ib19	LCZ 2	67.5562	33.4205	168	iButton
ib03	LCZ D	67.5718	33.3460	141	iButton
ib04	LCZ B	67.5665	33.3575	131	iButton
ib05	LCZ 2/5	67.5696	33.3538	137	iButton
ib06	LCZ D	67.5695	33.3660	143	iButton
ib07	LCZ 5	67.5705	33.3805	153	iButton
ib08	LCZ 2	67.5660	33.3816	150	iButton
ib09	LCZ 2/5	67.5698	33.3903	160	iButton
Short experiment campaign in 2015 (29.01.2015 – 02.02.2015)					
R1 / ib01	LCZ 9/B	67.5508	33.3621	132	1) AWS 2) iButton; 3) WMO station data downloaded from www.rp5.ru (6-hour resolution); 4) WMO station data with detailed 1-min resolution. These data are not publicly available.
U1 / ib02	LCZ 5	67.5683	33.4063	180	1) AWS 2) iButton

R2 / ib03	LCZ A	67.5804	33.3036	155	1) AWS 2) iButton
R3 / ib11	LCZ A	67.5912	33.4628	150	1) AWS 2) iButton
ib04	LCZ A	67.5977	33.4135	134	iButton
ib05	LCZ A	67.5980	33.4453	130	iButton
ib06	LCZ 2	67.5579	33.4377	182	iButton
ib07	LCZ A	67.5509	33.4005	135	iButton
ib08	LCZ A	67.5518	33.4804	169	iButton
ib09	LCZ 2/5	67.5842	33.4055	154	iButton
ib12	LCZ A	67.5686	33.3951	168	iButton
ib13	LCZ 5	67.5646	33.4286	184	iButton
ib14	LCZ B	67.5513	33.4498	149	iButton
ib15	LCZ B	67.6025	33.4295	142	iButton
ib16	LCZ 5	67.5705	33.3805	153	iButton
ib17	LCZ 2/5	67.5718	33.3461	141	iButton
ib18	LCZ 2	67.6082	33.4119	154	iButton
ib19	LCZ B	67.5892	33.4195	172	iButton
ib20	LCZ 2/5	67.5625	33.3876	151	iButton
ib21	LCZ B	67.5740	33.4523	175	iButton
ib23	LCZ 9	67.5563	33.3083	119	iButton
ib24	LCZ D	67.6041	33.4008	140	iButton
ib25	LCZ B/5	67.5779	33.4327	169	iButton
ib27	LCZ 2/5	67.5661	33.3811	149	iButton
ib28	LCZ B/D	67.5778	33.3590	138	iButton
ib29	LCZ B/D	67.5696	33.3662	144	iButton
ib31	LCZ 5	67.5622	33.4009	166	iButton
ib32	LCZ D	67.5841	33.3866	145	iButton
ib34	LCZ A	67.5742	33.3925	157	iButton
ib35	LCZ 5	67.5579	33.4101	170	iButton
ib37	LCZ D/10	67.6057	33.4101	144	iButton
ib38	LCZ D	67.5756	33.4128	137	iButton
ib39	LCZ D	67.5661	33.3568	132	iButton
ib40	LCZ B	67.5632	33.4111	179	iButton
Long-term campaign 2015-2016 (10.12.2015 – 28.02.2016)					
R1	LCZ 9/B	67.5508	33.3621	132	1) AWS 2) WMO station data downloaded from www.rp5.ru (6-hour resolution);
R2	LCZ A	67.5804	33.3036	155	AWS belonging to Polar-Alpine Botanical Garden-Institute (http://pabgi.ru/). These data are not publicly available.
U1	LCZ 5	67.5683	33.4063	180	AWS
R4 / ibR4	LCZ A	67.5884	33.4302	187	iButton
R5 / ibR5	LCZ B	67.5529	33.4499	170	iButton (observation period ends on 08.02.2016; the sensor has fallen into snow after this date)
U3 / ibU3	LCZ 5/B	67.5645	33.3998	163	iButton
U2 / ibU2	LCZ A	67.5672	33.3912	167	iButton

S2 Accuracy of the satellite remote sensing data products

The hyperspectral MODIS system performs a survey in 36 channels in the visible, near, mid, and thermal IR spectra. The LST data product uses the data of channels 31 and 32 (10.78–11.28 μm and 11.77–12.27 μm, respectively), in which the intensity of the Earth's surface thermal emission is recorded and which cover the range corresponding to the maximum intrinsic emission of the Earth (10–12 μm). The spatial resolution of images received in these channels is about 1000 m.

We recognize that the magnitude and spatio-temporal dynamics of the UHI is sensitive to the dataset used in the analysis. The meteorological observations of the surface air temperature as provided by the AWSs and iButton loggers do not necessarily refer to the same physical UHI phenomenon as the land surface temperature (LST) anomalies derived from satellite remote sensing (Christen and Vogt, 2004). Therefore we compared in-situ air temperature and MODIS LST (separately for the TERRA and AQUA platforms) for the winter of 2015/16. Figure S2 shows the rather good agreement between the LST from remote sensing and the in situ temperature observations at the R1 site. The R1 station is located on a homogenous landscape near the Imandra Lake, which is frozen in wintertime. There are somewhat weaker relationships between the LST and in-situ measurements for the temperature differences between urban and rural sites (U1 and R1) and between two different rural sites at different elevations (R4 and R1), with the determination coefficients in the range $R^2 = 0.33$-0.46 (TERRA platform) and 0.47-0.64 (AQUA platform). Nevertheless, the data confirms that a linear approximation is satifactory and the temperature dependences are monotonic. So a warmer urban temperature in the in-situ data necessarily corresponds to a positive temperature anomaly in the satellite LST data. Table S2 quantifies the mean LST for the pixels collocated with the in-situ observation sites and selected landmarks (hills) for the winter of 2015-2016 and mean LST differencess between the city (U1 site) and different rural sites. These LST differences are related to each other in the same way as the differences between mean air temperature at corresponding sites according to in situ observations (see Tables S5.1 and S5.2).

References to S2

Christen, A. and Vogt, R.: Energy and radiation balance of a central European city, Int. J. Climatol., 24, 1395–1421, https://doi.org/10.1002/joc.1074, 2004.
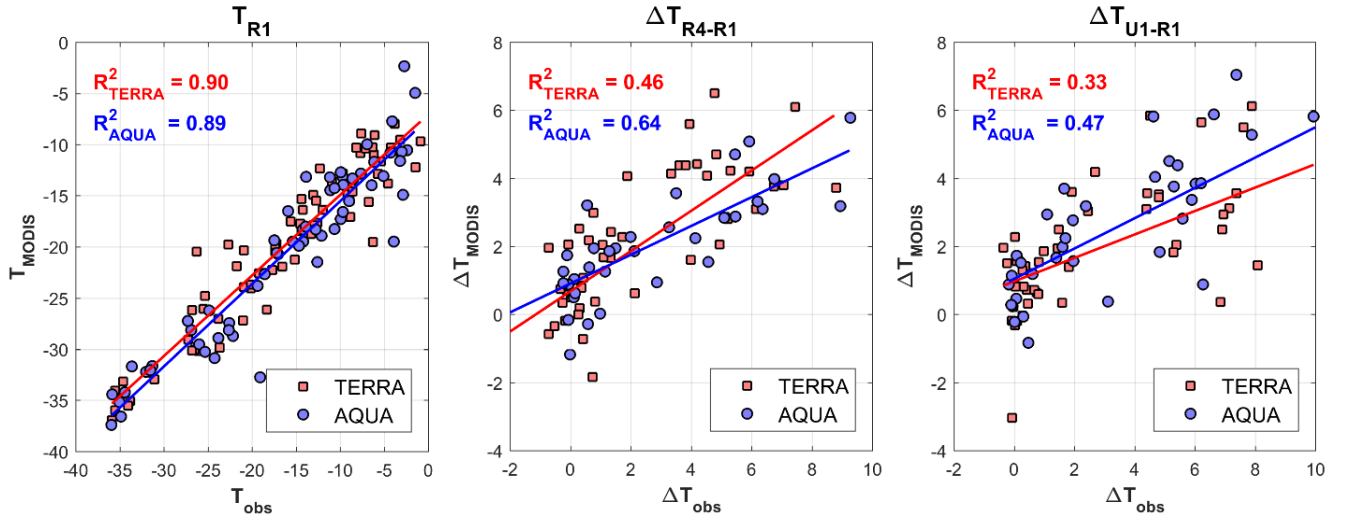
Figure S2. The relationships between the MODIS TERRA and AQUA LST (T_{MODIS}) and the in-situ surface air temperature observations (T_{obs}) during the winter of 2015-2016 for (a) data corresponding to the R1 site; (b) data corresponding to ΔT_{R1}^{R4}; and (c) data corresponding to the UHI intensity, ΔT_{R1}^{U1}.

5   **Table S2.** Statistical characteristics of the LST at the pixels collocated with the in situ observation sites and with a selection of landmarks. The mean LST is denoted as T_s (°C); the mean LST difference relative to T_s at the U1 site – as ΔT_s^{U1} (K). Mean LST difference $\Delta T_s^{U1}{}_{mean}$ is calculated for time series with filled spatial gaps (same length for each pair of points), 95%-quantile $\Delta T_s^{U1}{}_{95P}$ is calculated for time series with excluded gaps (different length for each pair of points). The analysis is given for the winter season 2015-2016.

			U1	Apatity hill	R1	R2	R4	R5	NE hill	NW hill
T_s	All days	TERRA	−17.2	−17.0	−19.6	−18.8	−17.4	−18.0	−16.8	−17.9
		AQUA	−17.7	−17.3	−20.5	−19.3	−17.9	−18.5	−17.0	−18.4
	Wf > 0.7	TERRA	−24.2	−24.1	−28.1	−26.9	−24.3	−25.6	−23.0	−25.4
		AQUA	−24.9	−24.4	−29.5	−27.5	−25.0	−25.9	−23.1	−25.6
$\Delta T_s^{U1}{}_{mean}$	All days	TERRA	0.0	−0.2	2.4	1.6	0.2	0.8	−0.4	0.7
		AQUA		−0.3	2.9	1.7	0.2	0.8	−0.6	0.7
	Wf > 0.7	TERRA		−0.1	3.9	2.7	0.1	1.4	−1.1	1.2
		AQUA		−0.4	4.6	2.6	0.1	1.0	−1.8	0.8
$\Delta T_s^{U1}{}_{95P}$	All days	TERRA	0.0	0.7	5.6	3.5	2.7	2.9	2.3	3.2
		AQUA		0.8	6.8	6.2	2.7	3.5	1.8	5.2

S3. The COSMO-CLM model and its configuration in sensitivity runs

Three nested computational domains were used to run the COSMO-CLM model with 12 km, 4 km, and 1 km horizontal meshes (Figure S3.1). All domains are centred on the city of Apatity. The model runs in the smallest domain were used to study the driving factors, extent, and sensitivity of the Apatity UHI using the TERRA-URB scheme. The surface relief used in the model is given in Figure S3.2. A comparison with the satellite digital elevation data (ASTER, availale from https://asterweb.jpl.nasa.gov/gdem.asp) shows that the key features of the landscape, including the hills surrounding the city, are present in the data used in the model.

The spectral nudging technique (von Storch et al., 2000) is utilized to ensure a more reliable binding of the internal model regime to atmospheric dynamics updated from the ERA reanalysis data. The benefits of this technique are shown in Feser and Barcikowska (2012) and Varentsov et al. (2016). Model options were set to their default values with the exception of the lake parameterization which was switched on in this study. The lower limit for the eddy viscosity coefficient of turbulent mixing was reduced in order to improve the simulations of the stable-stratified boundary layer. The scale of subgrid thermal inhomogenities was also reduced as recommended in Cerenzia et al. (2014).

The urban parameterization scheme TERRA_URB (Wouters et al., 2016 and 2017) is a rather simple bulk urban canopy scheme. It is based on a land model, TERRA, coupled to the COSMO model. The scheme corrects albedo, emissivity, surface heat capacity, surface heat conductivity and an aerodynamic roughness length for the "urban" grid cells (Demuzere et al., 2008; De Ridder et al., 2012). The scheme also includes parameterizations for a "bluff-body" thermal roughness length (Demuzere et al., 2008; Kanda et al., 2007), for the water balance of impervious urban surfaces (Wouters et al., 2015) and for anthropogenic heat flux climatology (Flanner, 2009). Geometrical effects of urban "street canyons" are accounted for in a semi-empirical urban canopy paremeterization by Wouters et al. (2016). The TERRA_URB scheme implements a tile approach which allows for a fractional urban cover in a grid cell. Anthropogenic heat flux (AHF or Q_H) is calculated according to Flanner's (2009) approach as following:

$$Q_H(t_d, t_y) = \overline{Q_H} w_d(t_d) w_y(t_y) \tag{1}$$

where $\overline{Q_H}$ is the annual-mean AHF value; w_d is fractional time of the day t_d which represents typical morning and evening peaks; w_y is fractional time of the year t_y; and the latitude is θ:

$$w_y(t_y) = 1 + A(\theta) \sin[2\pi(t_y + 0.25)] \tag{2}$$

$$A(\theta) = 1 - e^{-\frac{\theta-33}{25}} \ for \ \theta > 33 \tag{3}$$

Despite its simplicity, the TERRA_URB scheme has shown good results in several UHI studies, e.g. for Belgian cities (Wouters et al., 2016) and for Moscow (Varentsov et al., 2017; 2018). Although this scheme does not include the effects of radiative heating on vertical surfaces, which is potentially relevant for the Arctic wintertime, the very low solar angle during our experiments imply a very high aerosol optical depth and therefore small impact of this effect on the net urban heat flux. Further work and model development would be necessary to quantify the effect.

For Apatity, the urban fraction of the model grid cells was defined as it is shown in Figure S3.3. Other parameters of the urban canopy required for the model runs were defined for all urban model grid cells as follows: the building height was set to 17 meters, which is typical for the 5-storey buildings which are common in the town; the street canyon aspect ratio was set to 0.8; the building area fraction for the urban cells was set to 0.4.

5   Several approaches have been explored to choose the most realistic values of the anthropogenic heat flux (AHF) for Apatity in COSMO-CLM simulations. Global AHF databases suggest too low values for Apatity as they include information averaged over too large grid cells. For example, Flanner (2009) AHF data set has 2.5 degree spacial resolution and suggests the flux less than 1.5 Wm^{-2} for the Apatity region. Yang et al. (2017) has 1 km spatial resolution, suggesting the AHF values of 10-15 Wm^{-2} for 2010 and nearly zero AHF for the other two years included in the data set. Stewart and Oke (2012) recommended to

10   use the anuual mean AHF value of 25 Wm^{-2} for the prevaling local climate zone LCZ5 in US cities. The US winters are significantly warmer and therefore require less heating than the polar winters in Apatity. Stewart and Kennedy (2017) recommended a more realistic AHF values of 75 Wm^{-2} chosen for wintertime Moscow. This high values were successfully used in Moscow climate simulations with the COSMO-CLM by Varentsov et al. (2017; 2018).

The most comprehensive approach is proposed by a LUCY model (Allen et al., 2010, 2011). It requires specific input data

15   from local municipalities (including the population density, annual energy consumption, numer of cars per 1000 people, etc.) which is currently not availible for Apatity. The LUCY runs for European cities suggested the annual mean AHF of 50 Wm^{-2}. Survey of relevant independent estimates from scattrered literature sources generally support this AHF values too, e.g. Lindberg et al. (2013) gives the values exceeding 90 Wm^{-2} for many cities in Northern Europe, including Heldinki, Stokholm and Copenhagen.

20   We adopted the annual mean AHF of 50 Wm^{-2} for our COSMO-CLM runs. This annual mean value gives approximately 90 Wm^{-2} for the winter months in the TERRA_URB scheme (Flanner, 2009). This large AHF value agrees well with estimates based on coal consumption data (Figure 10). The daily average coal consumtion by the Apatity power plan was 2000±318 tonns per day over four winter monthes in January – February 2014 and 2015. This heat is used over the area of 4.5 km^2 in Apatity and 1.5 km^2 in Kirovsk. The specific heat of coal combustion equivalent is $2 \cdot 10^7$ J kg^{-1}. It gives an AHF estimation of

25   75 Wm^{-2}, which needs to be increased due to contribution from other anthropogenic heat sources such as transport engines, human metabolism and private heating.
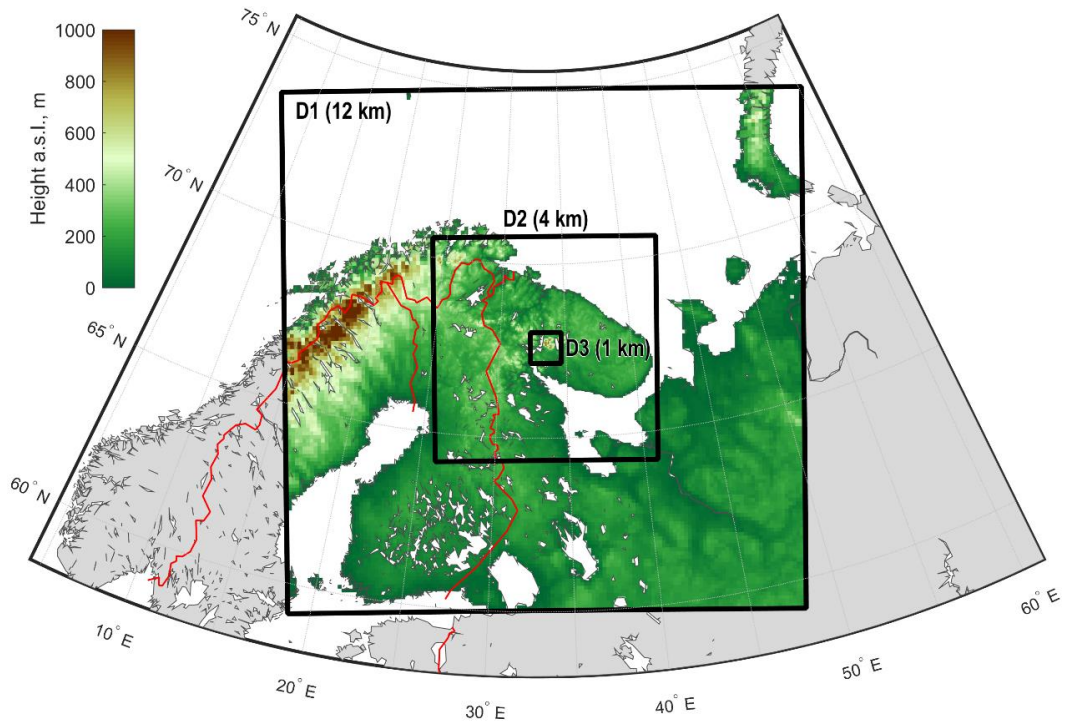
Figure S3.1. The three computational domains (D1, D2, D3) used to run the nested COSMO-CLM model in this study. The horizontal grid spacing is written in the brackets near the domain name.
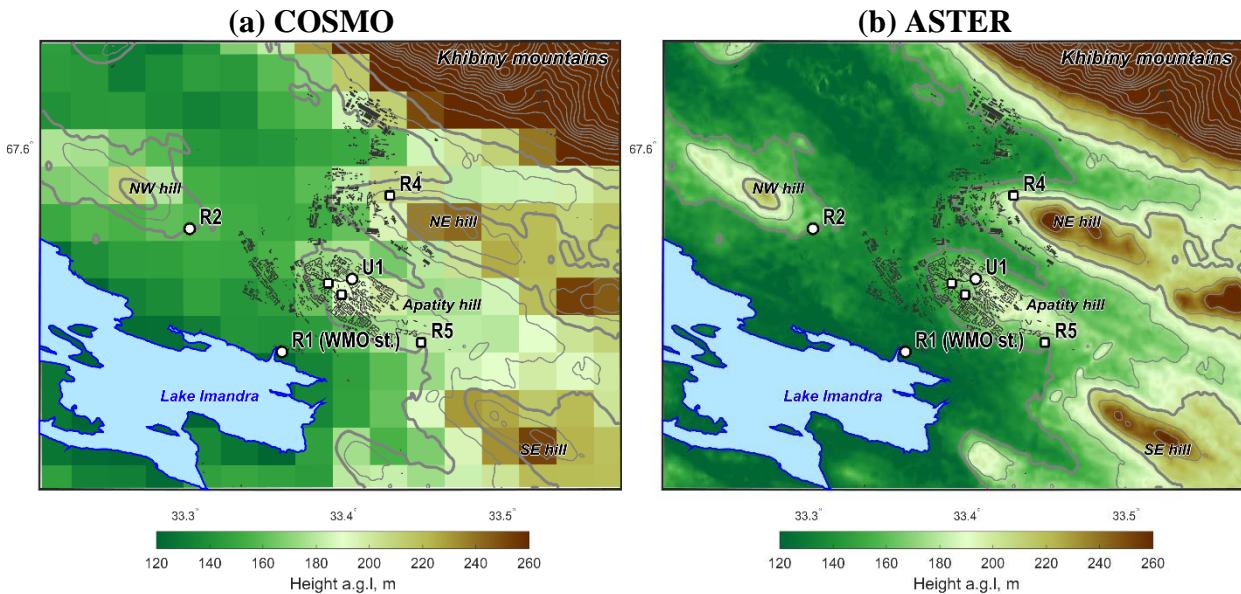
Figure S3.2. Comparison of (a) the relief used in the COSMO-CLM model runs in the smallest domain D3 and (b) the real relief according to the ASTER Digital Elevation Model. Isohypses are drawn every 25 meters according to the ASTER data in both panels.
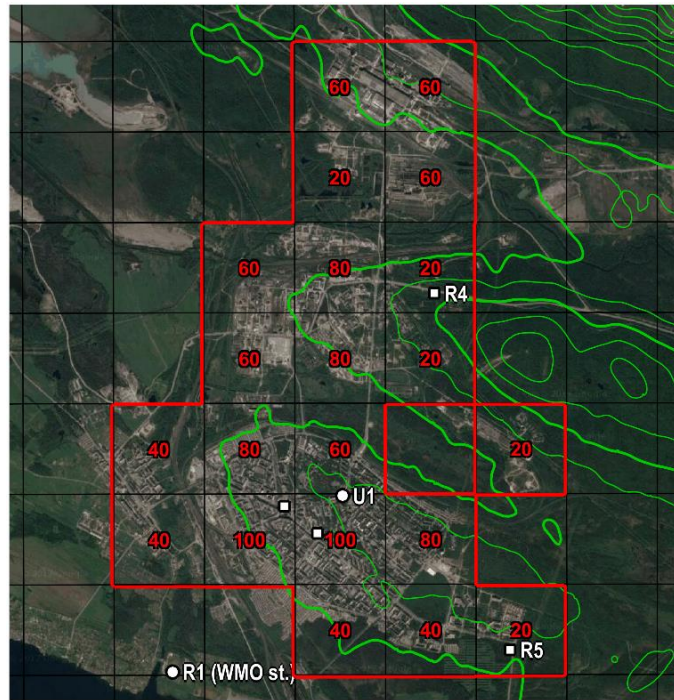
Figure S3.3. The urban surface fraction (in %, shown by red digits located at the centers of the model grid cells) as defined for the TERRA_URB scheme in the COSMO-CLM model runs. Isohypses are drawn every 25 meters according to the ASTER data.

References to S3

5   Allen, L., Beevers, S., Lindberg, F., Lamarino, M., Kitiwiron, N. and Grimond, G. S. B.: Global to City Scale Urban Anthropogenic Heat Flux: Model and variability. MEGAPOLI Sci. Report, 1-87, http://megapoli.dmi.dk/result/wp1res.html, 2011.

Allen, L., Lindberg, F. and Grimmond, C. S. B.: Global to city scale urban anthropogenic heat flux: model and variability, International Journal of Climatology, 31, 1990–2005, https://doi.org/10.1002/joc.2210, 2011.

10   Cerenzia, I., Tampieri, F. and Tesini, M. S.: Diagnosis of Turbulence Schema in Stable Atmospheric Conditions and Sensitivity Tests. COSMO Newsletter, 14, 28–36, 2014.

Demuzere, M., De Ridder, K. and van Lipzig, N. P. M. : Modeling the Energy Balance in Marseille: Sensitivity to Roughness Length Parameterizations and Thermal Admittance, J. Geophys. Res., 113, 1–19, https://doi.org/10.1029/2007JD009113, 2008.

15   Feser, F. and Barcikowska, M.: The Influence of Spectral Nudging on Typhoon Formation in Regional Climate Models, Env. Res. Lett., 7, 14024, https://doi.org/10.1088/1748-9326/7/1/014024, 2012.

Flanner, M. G.: Integrating anthropogenic heat flux with global climate models, Geophys. Res. Lett., 36, L02801, http://doi.org/10.1029/2008GL036465, 2009.

.Kanda, M., Kanega, M., Kawai, T., Moriwaki, R. and Sugawara, H.: Roughness Lengths for Momentum and Heat Derived from Outdoor Urban Scale Models, J. Applied Meteorol. Climatol., 46, 1067–1079, https://doi.org/10.1175/JAM2500.1, 2007.

Lindberg, F., Grimmond, C. S. B., Yogeswaran, N., Kotthaus, S. and Allen, L.: Impact of city changes and weather on anthropogenic heat flux in Europe 1995–2015, Urban Climate. 4, 1–15 https://doi.org/110.1016/j.uclim.2013.03.002, 2013.

De Ridder, K., Bertrand, C., Casanova, G. and Lefebvre, W.: Exploring a New Method for the Retrieval of Urban Thermophysical Properties Using Thermal Infrared Remote Sensing and Deterministic Modeling, J. Geophys. Res., 117, 1–14, https://doi.org/10.1029/2011JD017194, 2012.

Stewart, I. D. and Oke, T. R.: Local climate zones for urban temperature studies, Bull. Am. Meteorol. Soc., 93, 1879–1900, https://doi.org/10.1175/BAMS-D-11-00019.1, 2012.

Stewart, I. D. and Kennedy, C. A.: Metabolic heat production by human and animal populations in cities, Int. J. Biometeorol., 61, 1159–1171, https://doi.org/10.1007/s0048, 2017.

von Storch, H., Langenberg, H. and Feser, F.: A Spectral Nudging Technique for Dynamical Downscaling Purposes, Monthly Weather Rev., 128, 3664–3673, https://doi.org/10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2, 2000.

Varentsov M., Konstantinov, P. and Samsonov, T.: Mesoscale modelling of the summer climate response of Moscow metropolitan area to urban expansion // IOP Conf. Ser. Earth Environ. Sci., 96, 12009, https://doi.org/10.1088/1755-1315/96/1/012009, 2017.

Varentsov, M., Wouters, H., Platonov, V. and Konstantinov, P.: Megacity-Induced Mesoclimatic Effects in the Lower Atmosphere: A Modeling Study for Multiple Summers over Moscow, Russia, Atmosphere, 9, 50, https://doi.org/10.3390/atmos9020050, 2018.

Varentsov, M. I., Verezemskaya, P. S., Zabolotskih, E. V. and Repina, I. A.: Evaluation of the Quality of Polar Low Reconstruction Using Reanalysis and Regional Climate Modelling. Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa (in Russian), 4, 168–91, https://doi.org/10.21046/2070-7401-2016-13-8-168-191, 2016.

Wouters, H., Demuzere, M., De Ridder, K. and van Lipzig, N. P. M. : The Impact of Impervious Water-Storage Parametrization on Urban Climate Modelling, Urban Climate, 11, 24–50, https://doi.org/10.1016/j.uclim.2014.11.005, 2015

Wouters, H., Demuzere, M., Blahak, U., Fortuniak, K., Maiheu, B., Camps, J., Tielemans, D. and van Lipzig, N. P. M.: The efficient urban canopy dependency parametrization (SURY) v1.0 for atmospheric modelling: description and application with the COSMO-CLM model for a Belgian summer, Geosci. Model Dev., 9, 3027-3054, https://doi.org/110.5194/gmd-9-3027-2016, 2016.

Wouters, H. et al.: User Guide for TERRA URB v2.2: The Urban-Canopy Land-Surface Scheme of the COSMO Model. 2017.

Yang, W., Luan, Y., Liu, X., Yu, X., Miao, L. and Cui, X.: A new global anthropogenic heat estimation based on high-resolution nighttime light data, Scientific Data, 4, 170116, https://doi.org/10.1038/sdata.2017.116, 2017.

S4 Temperature variability during four of the selected case studies

There were four case-study periods (cases 3-6 from Figure 3) selected for a more detailed study of the UHI in the winter of 2015-2016. Figure S4 shows the observed temperature at three urban and five rural sites for those cases. Note that the temperature at the urban sites is almost always higher than the temperature at any of the rural sites. The largest differences are found for the R1, R2, and R5 sites. The R2 and R5 sites are located at about the same elevation as the urban sites. This indicates that the direct input of atmospheric temperature inversions on the observed temperature differences is smaller than other factors. The R2 station is located in dense forest, which may explain the generally higher temperature at this location. The R4 station is located at a higher elevation than the urban sites, which explains the smaller urban-rural differences with respect to this station, although they are almost always positive.
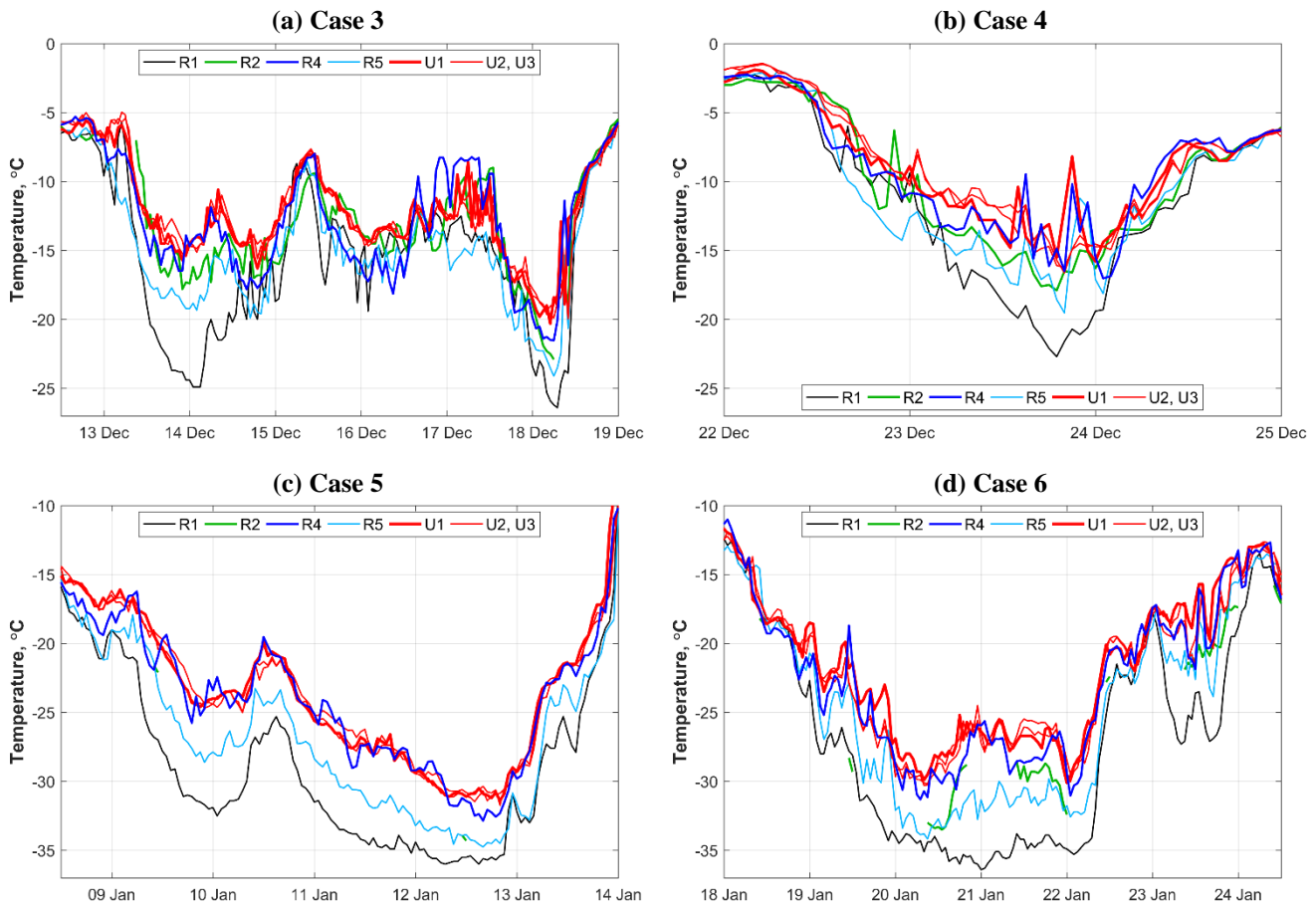
Figure S4. Temperature variability for the four selected case-studies from Figure 3: Case 3 covers the period 13-19 December 2015; Case 4: 22-25 December 2015; Case 5: 09-14 January 2016; Case 6: 18-24 January 2016. The sites U1, R1, and R2 were equipped with the AWSs, the remaining four sites were equipped with iButton temperature sensors. The urban sites are shown by the red lines.

S5. Statistical analysis of the UHIARC measurements for the winter season of 2015-2016

Detailed statistical analysis of in situ data was performed for the long-term observation campaign in 2015-2016. Separately, we calculated the statistics for the time intervals when R2 observations were available (covering about 40% of the total period of observations) and for the days with $W_f > 0.7$ (25%). Selected statistics of the UHIARC observations are provided in the Tables below. The Tables contain: mean temperatures, T_{mean}; mean temperature differences $\Delta T_{any\ site}^{U1}$ between the urban U1 site and other sites, ΔT_{mean}; and the *95%-quantile* of these differences, ΔT_{P95}. Table S5.1 contains information only for the AWS sites covering the entire winter period of observations (10 December 2015 – 1 March 2016). Table S5.2 contains information for shorter period for which the data for all iButton sensors is available (10 December 2015 – 8 February 2016), only for the cases with the solar elevation less than 2° above the horizon to avoid radiation errors of unshaded iButton sensors.

Statistical significance was estimated for ΔT_{mean} using the two-side Student test:

$$\frac{(\overline{\Delta T_j^i})}{\sqrt{\sigma^2(T_i)/n + \sigma^2(T_j)/n}} > t(p,n)$$

where T_i and T_j are the temperatures at two sites, $\overline{\Delta T_j^i}$ – their mean difference, $\sigma^2(T_i)$ and $\sigma^2(T_j)$ – their dispersions, n – sampling size, t-values $t(p,n)$ were calculated by Matlab *tinv* routine for given p-values.

Table 5.1. Selected statistics of the UHIARC observations (T_{mean}, ΔT_{mean} and ΔT_{P95}) for the AWS sites for different periods from 10 December 2015 to 1 March 2016. The ΔT_{mean} values significant at the level p < 0.01 are shown with shading.

Site		U1	R1	R2	Sampling size n
Elevation, m		185	130	153	
T_{mean}	Whole period	−10.5	−12.4	n/a	1968
	Period with R2 data available	−8.3	−10.2	−9.1	959
	Whole period with Wf > 0.7	−18.2	−23.6	n/a	420
	Period with R2 data available	−14.5	−19.8	−16.3	225
ΔT_{mean}	Whole period	0	1.9		1968
	Period with R2 data available		1.9	0.8	959
	Whole period with Wf > 0.7		5.5		420
	Period with R2 data available		5.3	1.8	225
ΔT_{P95}	Whole period	0	7.8		1968
	Period with R2 data available		8.1	3.5	959

Table S5.2. Selected statistics of the UHIARC observations (T_{mean}, ΔT_{mean} and ΔT_{P95}) for the AWS sites under the conditions that the maximum solar elevation was less than 2° above the horizon to avoid radiation errors of unshaded iButton sensors. The ΔT_{mean} values significant at the level $p < 0.01$ are shown with shading.

Site		U1	U2	U3	R1	R2	R4	R5	Sampling size n
Elevation, m		185	≈170	≈170	130	153	200	168	
T_{mean}	Whole period	−12.8	−12.8	−12.7	−15.2	n/a	−13.3	−14.5	1374
	Period with R2 data available	−10.4	−10.3	−10.2	−12.9	−11.4	−10.9	−12.1	573
	Whole period with Wf > 0.7	−19.0	−19.1	−18.8	−24.5	n/a	−19.8	−22.3	370
	Period with R2 data available	−15.4	−15.5	−15.2	−20.7	−17.2	−16.4	−18.5	184
ΔT_{mean}	Whole period	0	0.0	−0.1	**2.4**	n/a	0.4	**1.7**	1374
	Period with R2 data available		−0.1	−0.2	**2.5**	**1.0**	0.4	**1.7**	573
	Whole period with Wf > 0.7		0.1	−0.2	**5.5**	n/a	0.8	**3.3**	370
	Period with R2 data available		0.1	−0.2	**5.3**	**1.8**	1.0	**3.1**	184
ΔT_{P95}	Whole period	0	1.6	1.4	8.0	n/a	2.5	4.9	1374
	Period with R2 data available		1.7	1.4	8.6	4.2	2.6	5.0	573
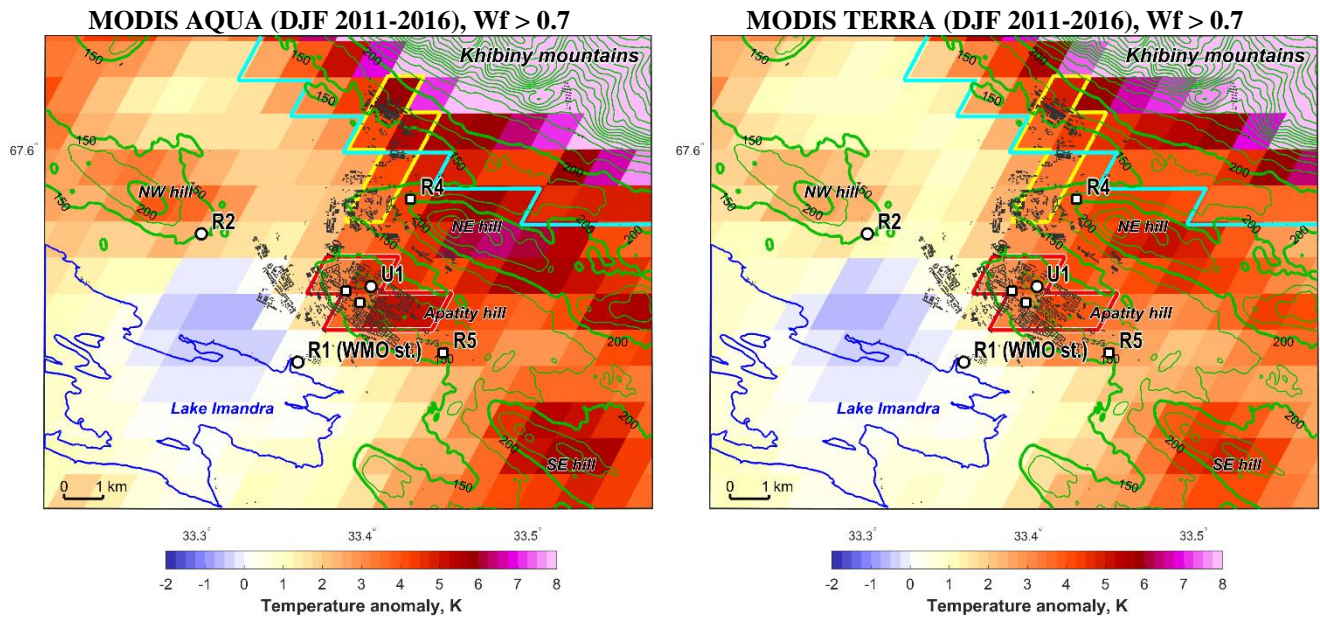
Figure S6. The MODIS LST anomalies corresponding to a large weather factor, $W_f > 0.7$ averaged over the winter months (DJF) from the years 2011-2016. A large W_f corresponds to calm anticyclonic conditions and the most intense UHIs. The anomalies are calculated as deviations from the LST at the grid cell co-located with the R1 site. The border indicated with the red line indicates grid cells classified as urban areas, yellow line – industrial areas, cyan line – grid cells related to Khibiny region.
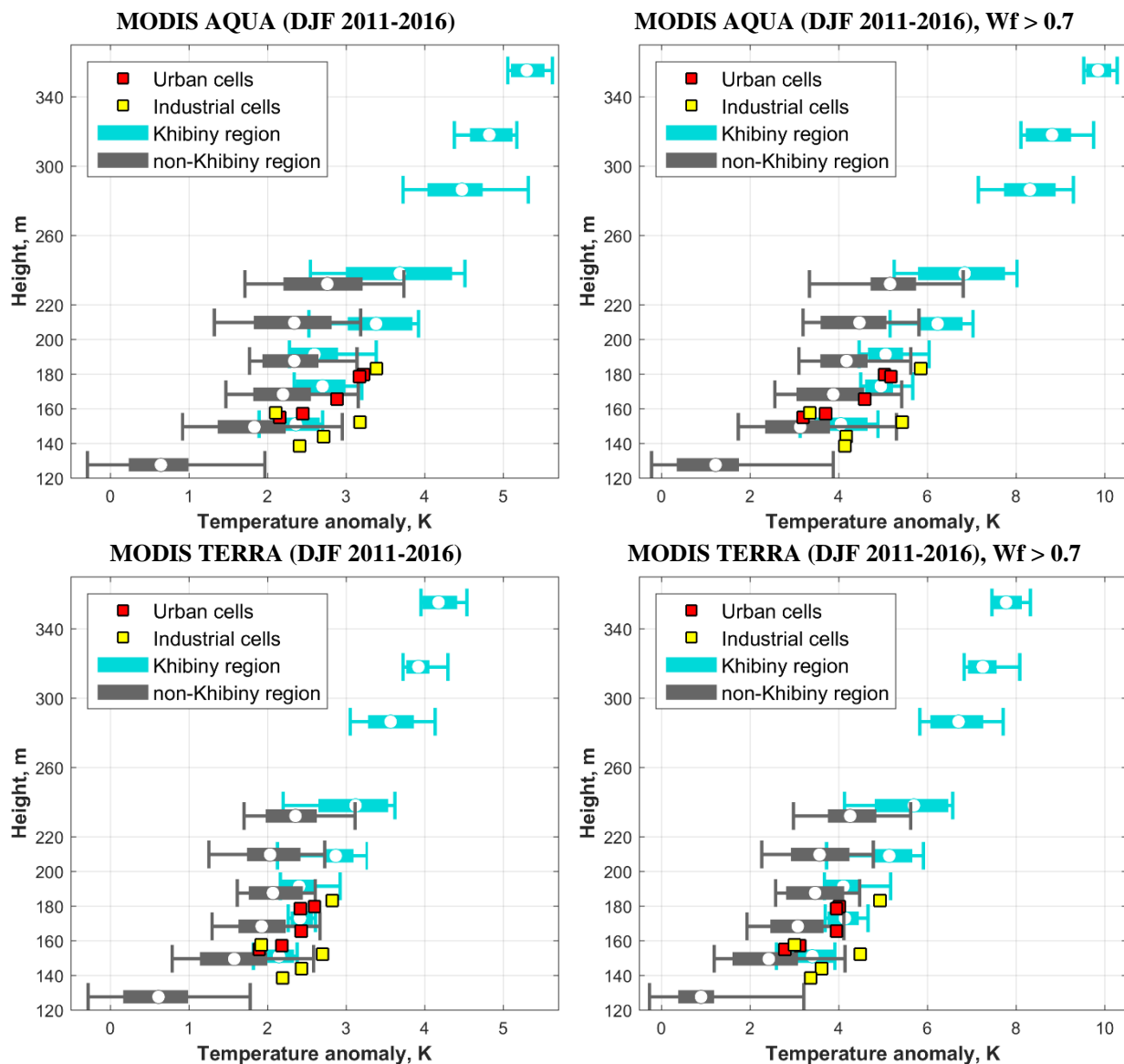
Figure S7. The relationship between the mean land surface temperature (from the MODIS LST data for DJF 2011-2016) and the surface elevation (from the ASTER Digital Elevation Model) for the winter months of 2011-2016. The right panels include all data; the left panels – the data for days with a large weather factor ($W_f > 0.7$) characterizing calm, anticyclonic conditions and the most intense UHI. All temperature anomalies are calculated as deviations from the LST in the grid cell collocated with the R1 station. Boxes-and-whiskers represents the mean values (the white circles), 25 and 75 percentiles (the filled boxes) and 5 and 95 percentiles (whiskers) of LST within the height ranges 120-140, 140-160, 160-180, 180-200, 200-220, 220-260, 260-300, 300-340 and 340-380 meters a.s.l. Data are presented separately for the non-urbanized cells in Khibiny region (cyan boxes) and for other rural areas (gray boxes).
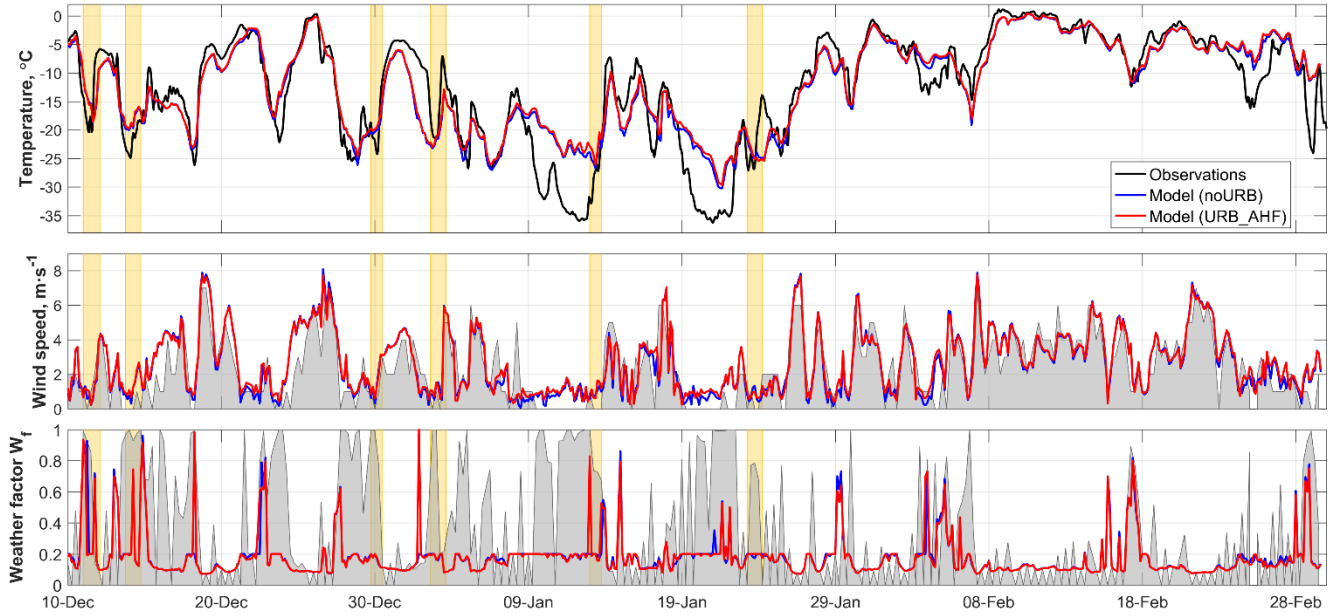
Figure S8. Intercomparison of the observed (the black line and gray shading) and modelled meteorological characteristics for R1 site: surface air temperature; wind speed at 10 m height above the ground; weather factor W_f. The observed wind speed and W_f are taken from the WMO R1 station (sampled at 6-hour intervals). The model output was sampled at 3-hour intervals. The orange shading identifies six cold periods with strong UHI which were selected for sensitivity studies.
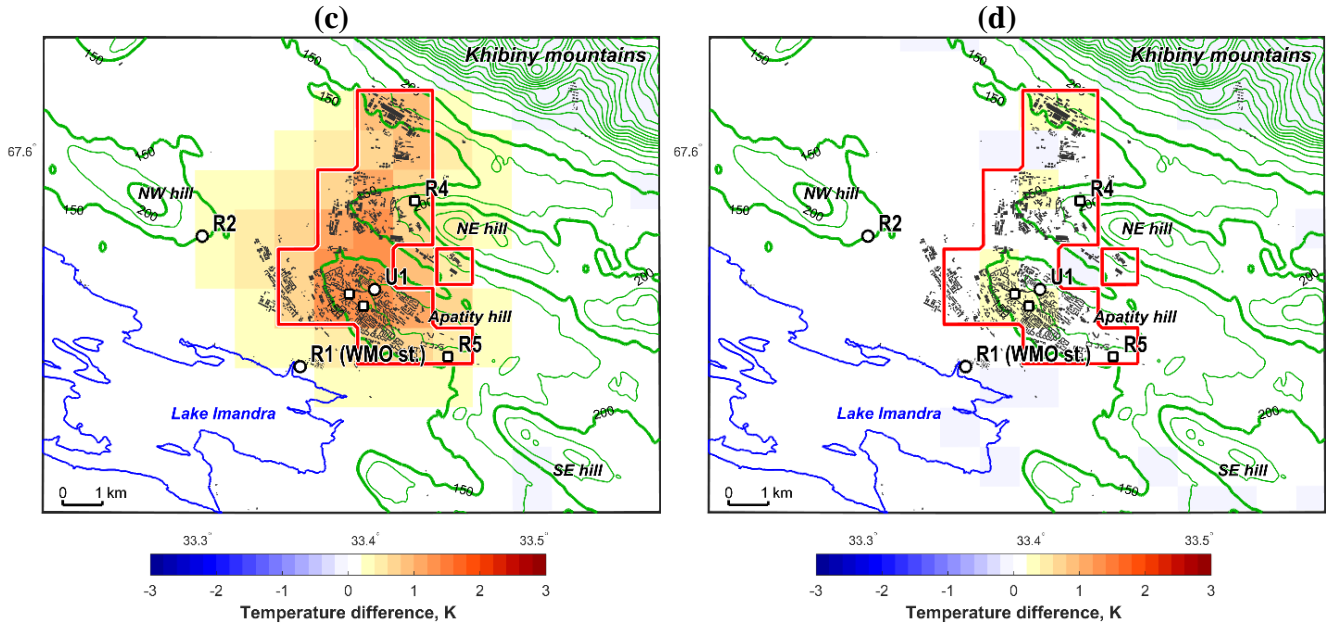
Figure S9. Maps of the simulated surface air temperature difference between the COSMO-CLM runs: URB_AHF – noURB (a), URB_noAHF – noURB (b). The differences are averaged over the whole period (from 10/12/2015 to 01/03/2016). The urban land use – land cover was introduced in the grid cells encompassed by the red lines.