Supplementary Material for the paper "Quantifying Local and Mesoscale Drivers of the Urban Heat Island of Moscow with Reference and Crowdsourced Observations"

Mikhail Varentsov, Daniel Fenner, Fred Meier, Timofey Samsonov, Matthias Demuzere

Supplementary S1. Reference weather stations

Table S1. Information about reference weather stations, selected for determining the mean background air temperature. WMO: World Meteorological Organization, LCZ: Local Climate Zone.

Station name	**WMO ID**	**Longitude (° E)**	**Latitude (°N)**	**Distance from city center (km)**	**Dominant LCZ class**
Serpukhov	27618	37.46556	54.9225	91.9	D – low plants
Maloyaroslavets	27606	36.48583	55.01694	108.4	6 – open low-rise
Kolomna	27625	38.7325	55.14222	96.5	6 – open low-rise
Naro-Fominsk	27611	36.70111	55.38722	70.6	D – low plants
Pavlovsky Posad	27523	38.6925	55.77167	66.4	6 – open low-rise
Novo-Iyerusalim	27511	36.825	55.90639	53.3	6 – open low-rise
Klin	27417	36.74972	56.35	86.5	8 – large low-rise
Dmitrov	27419	37.55722	56.3575	68.1	6 – open low-rise
Alexandrov	27428	38.75056	56.4	100.5	6 – open low-rise

Supplementary S2. Evaluation of the CWS quality control algorithm

In order to evaluate QC results, a comparison between CWS and reference data was performed for those CWS that are located in the vicinity from reference stations, with a distance of 2 km for rural weather stations (located beyond 25-km distance from Moscow city center) and 1 km for all other reference sites. Figure S2.1 demonstrates that the proposed QC algorithm decreases the root-mean-square error from about 5 °C for L0 data to about 1 °C in winter and 1.5 °C in summer for L5 data, which is always lower than the scores obtained for data that was filtered with the CrowdQC algorithm (O1). It should be noted that CrowdQC algorithm rejects less data than the last level (L5) of the proposed QC algorithm, with rejected data fraction of 33% and 34% for winter and summer, respectively (should be compared with $\mathbf{100\%} - \boldsymbol{f}$ from Figure 2 from the paper).

We further compare errors for contrasting conditions of rural areas (beyond 25 km distance from city center) and central part of the city (within 5 km distance from city center). In general, errors are lower for urban sites than for rural sites (Figure S2.1). This finding is coherent considering that most CWSs are installed in built-up areas and near buildings, more resembling reference stations in urban areas than in rural areas.

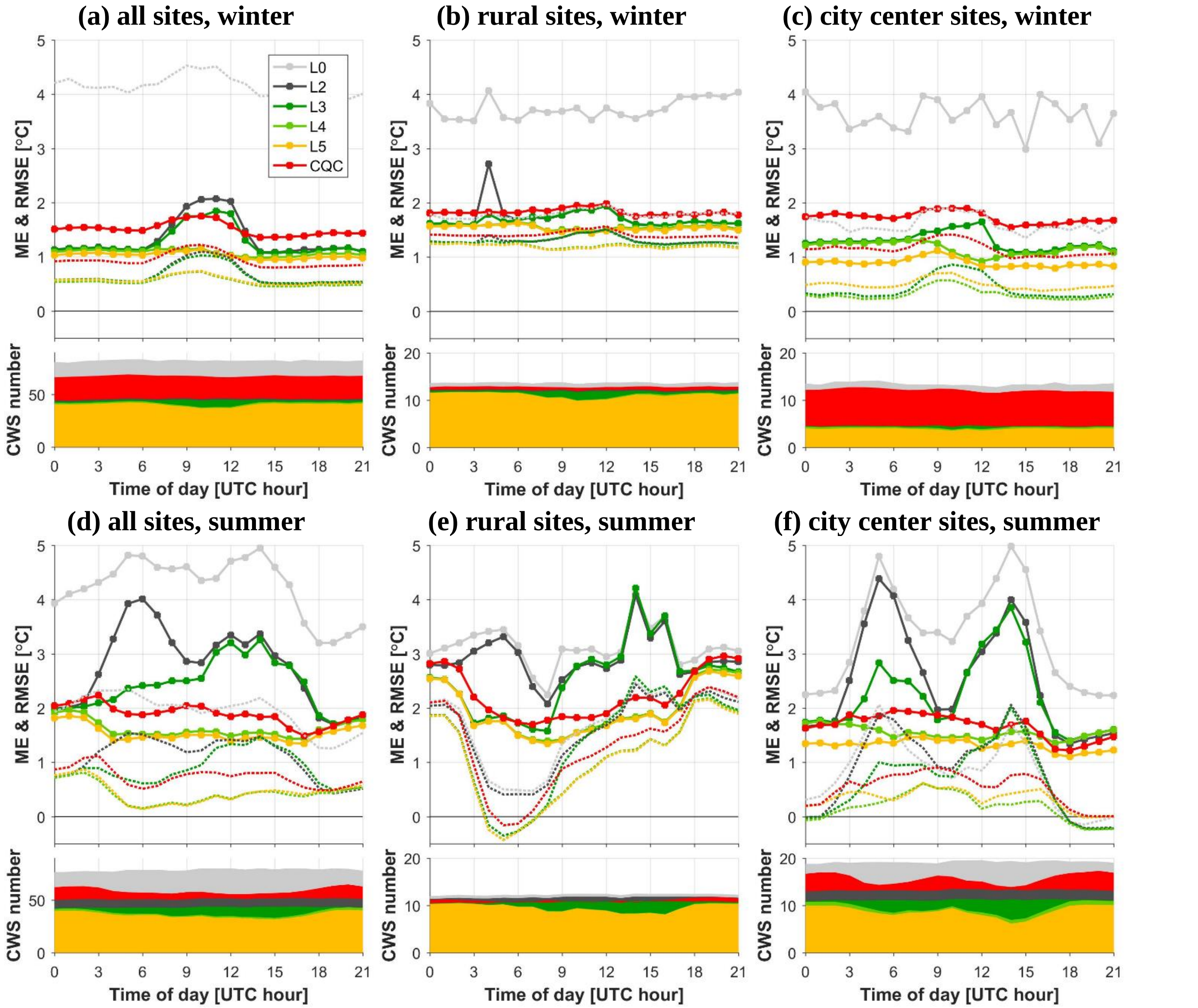

Figure S2.1. Mean (ME, dotted lines) and root-mean-square (RMSE, solid lines) errors for the CWS temperature observations with different QC levels for winter (a-c) and summer (d-e) periods, calculated with respect to the reference temperature observations for all reference WSs and AAQSs (a, d), rural and suburban WS in a distance of more than 25 km from the city center (b, e), WSs and AAWSs in the in a distance of less than 5 kmfrom the city center (c, f). CQC refers to the CWS data controlled using Crowd QC after level O1. In each subfigure, lower plots indicate the number of valid CWS temperature observations. Note that ranges differ in these plots.

Supplementary S3. Weather conditions for sampled cases

The near-surface wind speed during the sampled cased does not exceed 3 m/s according to observations of the MSU meteorological observatory as well as ERA5 reanalysis (for a grid cell covering Moscow). Mean wind speed over the sampled cases according to MSU observations (ERA5) is 0.9 (2.0) and 1.3 (2.1) m/s for summer and winter, respectively. Mean low cloud-cover fraction according to MSU observations is 0.14 and 0.48 for summer and winter, respectively. These values are significantly lower in comparison to the means over the whole selected periods. Wind direction during the sampled cases is not homogeneous but is still quite diverse (Figure S3.1), so we deem it acceptable for a coarse quasi-climatic approximation.

(a) winter, ERA5 (b) winter, MSU (c) summer, ERA5 (d) summer, MSU

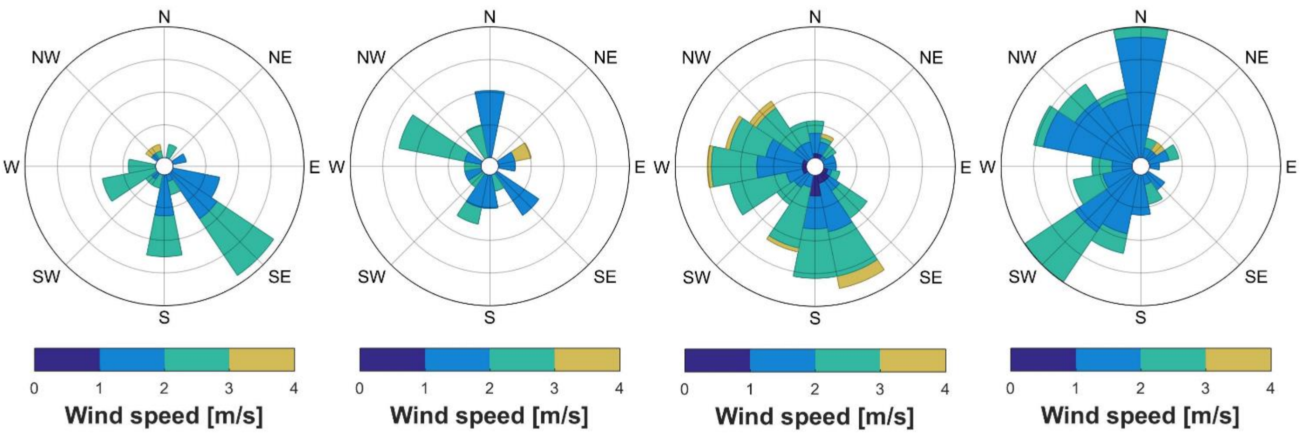

Figure S3.1.Wind roses showing the distribution of the near-surface wind speed and direction during the sampled winter (a, b) and summer (c, d) cases with intense UHII according to 10-m wind from ERA5 reanalysis (a, c) and observations at 15-m height at the MSU site (b, d). For observations, only cases with wind speed higher than lower sensitivity limit (1 m/s) are used.

Supplementary S4. Dependence between UHI intensity and distance from city center

Table S4. Quantitative parameters of the dependence between UHI intensity and distance from city center for specific LCZs and LCZ groups used in Figure 7, calculated for different seasons and data types (REF/CWS), where n denotes number of observations sites available for specific LCZs or LCZ groups (considering conditions described in the Section 2.5), R denotes the Spearman correlation coefficient, and R^2 the coefficient of determination. R and R^2 are estimated only for $n > 3$. Gray shading indicates LCZs or LCZ groups with $R^2 > 0.5$.

LCZs	REF, winter			CWS, winter			REF, summer			CWS, summer		
	n	R	R2	n	R	R2	n	R	R2	n	R	R2
1	0	-	-	1	-	-	0	-	-	1	-	-
2	3	-	-	2	-	-	3	-	-	8	0.76	0.58
3	0	-	-	1	-	-	0	-	-	1	-	-
4	35	-0.80	0.64	167	-0.44	0.20	32	-0.62	0.38	179	-0.52	0.27
5	4	-0.40	0.16	24	-0.53	0.28	3	-	-	31	-0.54	0.29
6	9	-0.90	0.81	167	-0.36	0.13	8	-0.88	0.78	153	-0.33	0.11
7	0	-	-	0	-	-	0	-	-	0	-	-
8	1	-	-	6	-0.94	0.89	1	-	-	3	-0.50	0.25
9	1	-	-	41	-0.63	0.40	1	-	-	44	-0.57	0.33
10	0	-	-	0	-	-	0	-	-	0	-	-
A	2	-	-	12	-0.46	0.21	2	-	-	10	-0.59	0.35
B	2	-	-	0	-	-	2	-	-	1	-	-
C	0	-	-	0	-	-	0	-	-	0	-	-
D	4	-1.00	1.00	8	-0.62	0.38	4	-0.80	0.64	9	-0.35	0.12
E	0	-	-	1	-	-	0	-	-	0	-	-
F	0	-	-	0	-	-	0	-	-	0	-	-
G	0	-	-	1	-	-	0	-	-	1	-	-
4 & 5	39	-0.76	0.59	191	-0.46	0.21	35	-0.51	0.26	210	-0.51	0.26
8 & 10	1	-	-	6	-0.94	0.89	1	-	-	3	-0.50	0.25
6 & 9	10	-0.93	0.86	208	-0.42	0.17	9	-0.90	0.81	197	-0.39	0.15
A-D	8	-0.60	0.35	20	-0.58	0.33	8	0.10	0.01	20	-0.47	0.22

Supplementary S5. Estimation of the length scale of areas with homogeneous LCZ cover

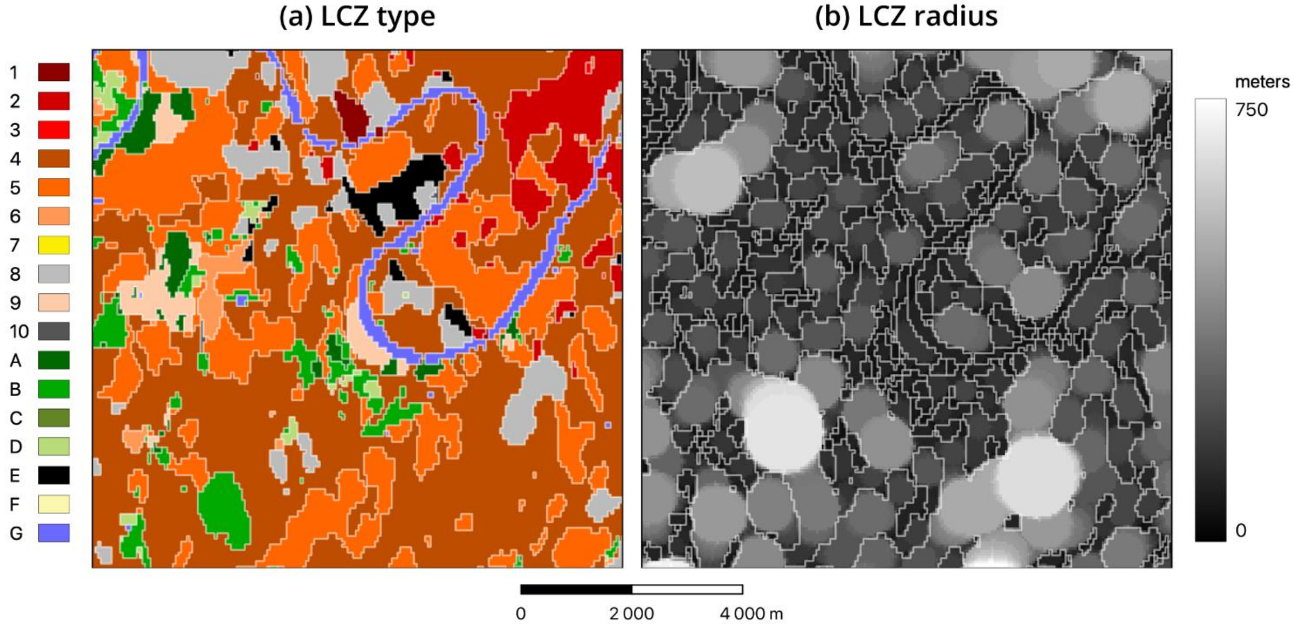

Figure S5.1. Example of application of the regional width estimate method (Samsonov et al., 2019) for a section of the LCZ map for Moscow (a), with resultant length scale, defined as inscribed circle-radius, for the same area (b).

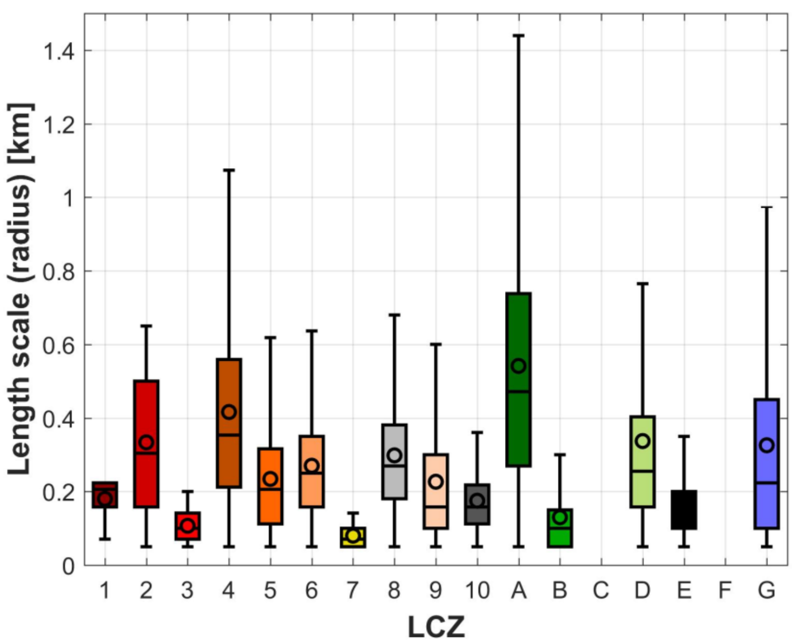

Figure S5.2. Boxplots representing the distribution of length scale of areas with homogeneous LCZ type, expressed via the inscribed circle radius (Figure S4.1b), for different LCZ classes within the study area.