Vol. 55: 105–118, 2012
doi: 10.3354/cr01123

CLIMATE RESEARCH
Clim Res

Published online November 29

OPEN ACCESS

Effects of scale on modelling the urban heat island in Turku, SW Finland

Juuso Suomi[1,*], Jan Hjort[2], Jukka Käyhkö[1]

[1]Department of Geography and Geology, University of Turku, Turku 20014, Finland
[2]Department of Geography, University of Oulu, 90014 Oulu, Finland

ABSTRACT: Scale has a central role in the modelling of an urban heat island (UHI). In the present study, the effects of the spatial extent of explanatory variables and modelling approach were explored in Turku (179 000 inhabitants), SW Finland, based on the average temperatures recorded by 50 temperature loggers over a 2 yr period. First, the optimal monthly circle of influence around temperature measurement points was determined for 6 urban and 6 non-urban explanatory variables. Second, the independent effect of the intercorrelated urban variables was tested using hierarchical partitioning (HP), a method for tackling multicollinearity. Third, the relative importance of the strongest urban variable was explored with 6 non-urban variables utilizing HP and multiple linear regression (LR). The mode of the optimal monthly buffer size was 1000 m for the urban variables and 2000 m for the non-urban variables. The optimal buffer size of the urban variables was largest in spring and smallest in autumn and winter. A clear seasonal trend was not observed with the non-urban variables. Based on HP, the strongest urban variable had a more independent effect on temperatures than the non-urban variables. The importance of the 'best' urban variable was also revealed in LR. Of the non-urban variables, the importance of the Landsat ETM+ derived 'greenness' variable was highlighted in HP and 'water bodies' in LR. In conclusion, the circle of influence of the explanatory variables varied substantially. Scale should be considered before doing multivariate analyses in UHI studies. Methodologically, the HP technique complements traditional UHI studies in multivariate model settings.

KEY WORDS: Urban heat island · Circle of influence · Scale · Explanatory variable · Spatial modelling · Multicollinearity · Hierarchical partitioning · Multiple linear regression

1. INTRODUCTION

A typical feature of urban climate is the urban heat island (UHI), in which city areas are warmer than surrounding rural areas. There are several factors affecting the development of UHI, but the key determinants are differences in solar heat storage, anthropogenic heat release and evaporation between urban and rural areas (e.g. Landsberg 1981, Oke 1987, Cotton & Pielke 1995). When assessing the thermal effect of environmental variables on temperatures, the influence of scale should be considered. In urban climatology, variables have often been calculated inside certain buffers around temperature measurement points (e.g. Wen et al. 2011). Oke (2006) concluded that the 'circle of influence', i.e. the effects of the surroundings on the weather measurement site, depends on instrument height, building density and boundary layer conditions, but it is typically around 500 m. At low instrument heights and among high-rise buildings, the circle of influence is smaller than, for example, in measurements high above open fields (Stewart & Oke 2009).

Giridharan et al. (2008) used a buffer size as small as 15 to 17 m when investigating the effects of vegetation on urban temperatures. In Kolokotroni & Giridharan (2008), a 50 m radius buffer was used in a UHI study in London. Giridharan et al. (2007) differenti-

*Email: juusuo@utu.fi

ated on-site variables such as sky view factor from off-site variables such as proximity to the sea. Giridharan & Kolokotroni (2009) concluded that the buffer size of on-site variables in winter could be around 25 m. At that resolution, the effects of anthropogenic factors could be captured and the effects of local wind patterns could be negated. In summer, however, they concluded that the buffer size should be larger. In contrast, the scale of the off-site variables should be based on the sizes of natural elements like hills and water bodies. In less densely built environments, the maximum buffer size of the off-site variables may have a radius of 1 km, but in densely built areas it is around 300 m (Giridharan & Kolokotroni 2009). Relatively large buffer size has been adopted by e.g. Wen et al. (2011), who used a 1500 m radius buffer in a UHI study in Guangzhou, southern China.

Houet & Pigeon (2011) tested how well the urban climate zones classification by Oke (2006) reflects climatically different areas on a city scale. They used circles with a radius of 100, 250 and 500 m in computing different environmental variables used in Oke's (2006) classification. Houet & Pigeon (2011) did not adopt a strong stance towards the optimal buffer size of the environmental variables, but they concluded that, when creating a map representing urban climate zones, the grid size should be much smaller than the 500 m circle of influence suggested by Oke (2006).

The shape of the area that has been used in the calculations has sometimes deviated from a circle, or the observation point has not been in the centre of the circle. Wong et al. (2011) used 4 partially overlapping 50 m radius buffers for predicting outdoor air temperature for buildings in Singapore. As a result, a little less than 100 m around the buildings, on average, was taken into account in the calculations. Sakakibara & Matsui (2005) used 250 m grid cells in assessing the effects of urban canopy characteristics and city size indices on UHI intensity, while Costa et al. (2007) used a buffer with a 350 m radius to the predominant wind direction and 150 m radius to the downwind direction.

In modelling a UHI, a common methodological problem is the significant intercorrelation of explanatory variables in the analyses. Multicollinearity among the environmental variables might result in the exclusion of other causal variables from multivariate models. Hierarchical partitioning (HP) (Chevan & Sutherland 1991) is an efficient approach for tackling multicollinearity. HP is a quantitative statistical method that can provide novel insights into UHI studies by decomposing the variation in response variables into independent components (cf. MacNally 2000, Heikkinen et al. 2004).

Considering the increasing interest in spatial analysis of UHI and the relatively large variation in employed buffer sizes, there are surprisingly few studies on scale effects and optimal buffer sizes in urban climatology. Moreover, we are not aware of any UHI studies using partitioning methods for tackling multicollinearity problems. In the present study, the optimal buffer size for 6 urban and 6 non-urban explanatory variables was assessed, and the variables with determined buffers were used in modelling the observed temperature differences in the coastal city of Turku, SW Finland, in 2006–2007. Our aim was to explore: (1) the optimal buffer size for 12 explanatory variables for monthly mean temperatures, (2) how the optimal buffer size varies between seasons, (3) which of the 6 highly intercorrelated urban variables has the greatest independent effect on temperatures in HP and (4) the relative role and explanatory power of the 'best' urban variable and the 6 non-urban (i.e. land use and environmental) variables in 2 different modelling settings, HP and linear regression (LR).

2. STUDY AREA

The study area consisted of the mid-sized (179 000 inhabitants) coastal city Turku (city centre: 60° 27′ N, 22° 16′ E) and parts of its neighbouring municipalities. The area is located in SW Finland at the mouth of the River Aura. The Baltic Sea coastline in the region is very indented, exhibiting a large archipelago to the southwest of the city (Fig. 1a,b). Due to the effects of nearby islands and the large archipelago, the climate of Turku is a mixture of coastal and inland types. Depending on the location and movements of large weather systems, either continental or marine characteristics can dominate (Alalammi 1987). At Turku airport, 7 km north of the city centre, the annual average temperature for the period 1981 to 2010 was 5.5°C. February is typically the coldest month, while July is the warmest, with average temperatures of –5.2 and 17.5°C, respectively. The highest temperature in the period 1981 to 2010 was 32.1°C (2010) and the coldest was –34.8°C (1987). Mean annual precipitation is 723 mm, of which 30% falls as snow. The average duration of permanent snow cover is approximately 90 d starting from 24 December. The wettest month is typically August, while April is the driest, with monthly rainfalls of 80 and 32 mm, respectively. Winds are highly variable in speed (average 3.4 m s^{-1}; highest in November, December and January: 3.6 m s^{-1}; and lowest in August: 3.1 m s^{-1}) and direction (dominant SW with

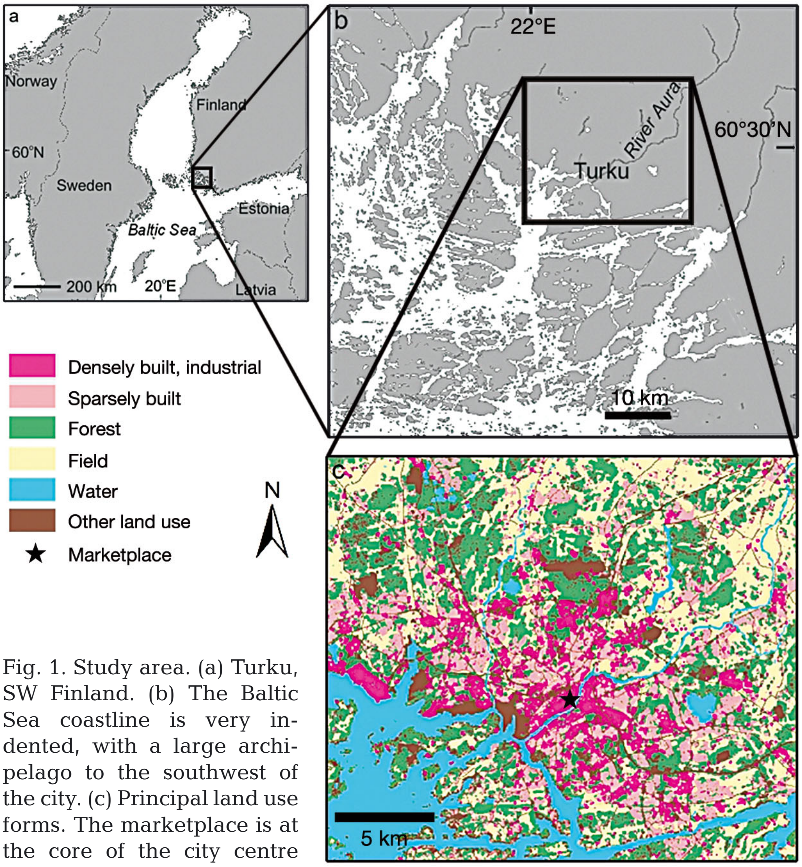

Fig. 1. Study area. (a) Turku, SW Finland. (b) The Baltic Sea coastline is very indented, with a large archipelago to the southwest of the city. (c) Principal land use forms. The marketplace is at the core of the city centre

17% proportion of all winds) due to cyclonic activity (FMI 2011, 2012). A strong seasonal factor in the climate of Turku is the spatiotemporal extent of sea ice cover, which typically lasts from 2 January to 1 April, but can have substantial year to year variation (Seinä & Peltola 1991, Seinä et al. 2006). Continuous sea ice eliminates heat and water vapour transfer from the sea and effectively makes the city more continental. In Köppen's climate classification (Peel et al. 2007), Turku belongs to the hemiboreal and humid continental 'Dfb' class together with southern parts of the Scandinavian Peninsula, the Baltic countries and much of Eastern Europe. The same climate type also exists around the Great Lakes region in the USA and in the Russian Far East around the Amur River basin and northern parts of Sakhalin Island.

The grid plan area of the Turku city centre is approximately rectangular in form, with a spatial extent of 4 km (SW–NE) by 1.5 km (SE–NW). The streets in the grid plan area are covered with tarmac or cobble stones. The marketplace (Fig. 1c) is the core area of the city centre, located in the middle of the grid plan and surrounded by commercial buildings. Residential areas and parks are the most common land use forms elsewhere in the grid plan area. The majority of the buildings in the grid plan area are 6- to 8-storey stone houses with varying surface colours. River Aura flows across the city centre, where its width varies between 50 and 100 m (City of Turku 2001). The majority of the urban parks are on the southeastern side of the river. Industrial areas, consisting of a dockyard and various light industry, lie scattered to the west and north of the city centre. Elsewhere outside the grid plan area, the land cover is a mosaic of suburbs, forests and fields. The largest sparsely populated areas are found towards the north of the city centre (Fig. 1c)

Topographically, the grid plan area consists of 30 to 50 m high bedrock outcrops with interleaving flat terrain 5 to 10 m above sea level. The area is old sea bottom, exposed due to post-glacial isostatic rebound during the last millennia. Therefore, flat areas consist of marine clay deposits whereas bedrock outcrops have been washed of fine sediments by wave action. Many of the hills are partially covered with buildings, while some of them host parks in their summit areas. Outside the grid plan area, hills alternate with flat areas and, in general, ground elevation rises gently towards the inland, reaching a maximum of 67 m above sea level in the northeast (Fig. 2).

3. DATA AND METHODS

Temperature data consisted of observations from 50 Hobo H8 Pro temperature loggers during the 2 yr period 2006 to 2007 (Fig. 2, Table 1). The logging interval was 30 min. According to the manufacturer, the accuracy of this instrument is ±0.2°C at 0 to 50°C, while the resolution is 0.02°C. Below 0°C, the accuracy is slightly worse, being about ±0.38°C at –20°C. Loggers were placed inside radiation shields at 3 m elevation above the immediately surrounding ground. The observations were part of the Turku Urban Climate Research Project TURCLIM (initiated in 2001) of the Department of Geography and Geology at the University of Turku. Calculated monthly average temperatures were used as the response

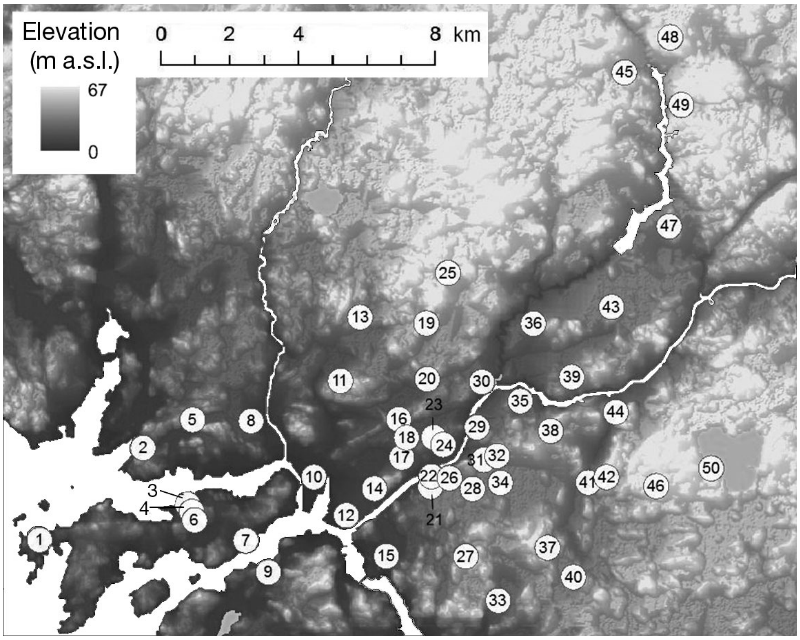

Fig. 2. Topography of the study area. Numbers refer to temperature logger sites (see Table 1)

variables in the statistical analyses. A month was considered the proper time span, as it is short enough to reflect seasonal variation in independent effects and explanatory powers of the variables. On the other hand, a month is long enough to smooth the variation caused by short-term weather phenomena.

Altogether, 12 explanatory variables were included in the first step of the analysis in which optimal buffer sizes were determined (Tables 2 & 3). The land use and environmental variables included 6 urban and 6 non-urban variables. The division into urban and non-urban variables was made mainly based on thermal characteristics of the land use forms. The urban variables described densely built areas with a strong effect on UHI. One of the urban variables, namely the building floor area, was derived from the City of Turku (2010) database. The other 5 urban variables were obtained from the SLICES (2011) land use classification. SLICES is a national mapping project initiated by the Ministry of Agriculture and Forestry and headed and processed by the National Land Survey of Finland. It consists of 45 land use classes, including the following main ones: A = residential and leisure areas; B = business, administrative and industrial areas; C = supporting activity areas; D = rock and soil extraction areas; E = agricultural land; F = forestry land; G = other land; and H = water areas.

Of the 6 non-urban variables, 'industrial and warehouse areas', 'detached houses' and 'agricultural lands' were derived from SLICES; 'water bodies' from SLICES and the CORINE land cover database (EEA 2006); 'relative elevation' from a digital elevation model (DEM; NLS 2009); and 'greenness' of tasseled cap transformation from Landsat ETM+ satellite images (cf. Crist & Cicone 1984, Hjort & Luoto 2006). In tasseled cap transformation, the spectral bands' pixel values are converted to weighted sums of the original values of a set of bands. One of these weighted sums, 'greenness', indicates the amount of green (i.e. photosynthetically active) vegetation, and it was selected as one variable instead of the more commonly used normalized difference vegetation index (NDVI). The selection was based on a preliminary analysis in which 'greenness' showed a higher correlation with the temperature variables (using Spearman's rank correlation r_S and optimal buffer sizes, the mean r_S of monthly temperatures with 'greenness' and 'NDVI' were −0.81 and −0.79, respectively). The spatial resolution of SLICES is 10 × 10 m, whereas the DEM and CORINE have a resolution of 25 × 25 m, and Landsat ETM+ imaging 30 × 30 m for bands used in the computations. The values of the explanatory variables were computed using ArcGIS 9.3 and MS Excel 2003 software.

The theoretical maximum buffer sizes were determined for each explanatory variable based on the literature and estimated thermal effects of the variable. In the literature, the radius of the buffers used or recommended varied typically between 17 and 1500 m, whereas the UHI itself has been reported to extend up to 30 km in large towns (e.g. Hicks et al. 2010). Oke (1973) made comparisons between city size and maximum UHI intensity. As UHI intensity strengthens between 100 000 and 1 000 000 inhabitants, the circle of influence is supposed to be >500 m, and in some cases even >1000 m. Stewart & Oke (2009) have also concluded that the circle of influence is larger in an open environment compared to that of built surroundings. Based on these facts, the theoretical maximum buffer size of most of the urban variables was set at 2 km, and for most of the non-urban variables at 5 km (Table 3, Fig. 3). The theoretical maximum buffer size for 'floor area' was 300 m because the data were available only from Turku and Kaarina municipalities, and larger buffers crossed their borders. For 'water bodies', a relatively large

Table 1. Temperature logger sites (see Fig. 2 for locations). Distance from city centre is measured from the centre point of the marketplace. Land cover includes the 2 most common land use forms inside a 100 m radius buffer based on the SLICES land use classification

Logger no.	Regional character	Elevation (m a.s.l.)	Distance from city centre (km)	Land cover in the surroundings
1	Rural	2.4	9.8	Forest, gravel road
2	Semi-urban	4.9	7.1	Forest, field
3	Rural	0.4	6.2	Forest, reeds
4	Rural	34.7	6.2	Forest, gravel road
5	Semi-urban	9.3	6.0	Gravel field, 2-storey building
6	Rural	15.1	6.2	Forest, gravel road
7	Rural	1.5	5.2	Field, meadow
8	Semi-urban	5.8	4.6	Parking lot, 2-storey building
9	Rural	0.6	5.1	Field, forest
10	Semi-urban	0.0	3.2	Gravel field, high grassland
11	Semi-urban	21.2	2.9	Forest, block of flats
12	Urban	4.6	2.8	Grassland, different-sized buildings
13	Rural	24.9	3.7	Forest, agricultural buildings
14	Urban	18.5	1.9	Block of flats, gravel field
15	Semi-urban	5.8	2.9	Forest, asphalt road
16	Semi-urban	19.9	1.3	Detached houses, park
17	Urban	10.3	1.0	Park, asphalt road
18	Urban	10.0	0.9	Railway, row houses
19	Semi-urban	26.2	3.0	Parking lot, field
20	Semi-urban	30.0	1.7	Park, 2-storey building
21	Semi-urban	25.0	1.0	Forest, detached houses
22	Urban	7.4	0.7	Asphalt road, block of flats
23	Urban	27.4	0.4	Park, public buildings
24	Urban (marketplace)	8.6	0.1	Asphalt road, block of flats
25	Semi-urban	36.8	4.2	Block of flats, 2-storey building
26	Urban	30.8	0.7	Block of flats, park
27	Rural	41.2	2.6	Forest, landfill
28	Semi-urban	20.0	1.2	Detached houses, gravel road
29	Urban	9.6	0.9	Row houses, roads
30	Semi-urban	19.9	1.8	Industrial, asphalt road
31	Urban	22.9	1.0	Block of flats, asphalt road
32	Urban	20.7	1.3	Park, asphalt road
33	Semi-urban	15.0	3.8	Allotment garden houses, forest
34	Urban park	20.0	1.6	Park, leisure-time areas
35	Semi-urban	14.8	2.1	Detached houses, roads
36	Semi-urban	20.1	3.6	Public buildings, detached houses
37	Semi-urban	25.1	3.4	Forest, graveyard
38	Semi-urban	30.0	2.6	Detached houses, roads
39	Semi-urban	23.6	3.5	Detached houses, forest
40	Rural	19.9	4.3	Field, roads
41	Semi-urban	15.8	3.5	Forest, meadow
42	Semi-urban	49.9	3.9	Block of flats, row houses
43	Rural	20.0	5.2	Landfill, forest
44	Rural	12.7	4.2	Field, forest
45	Rural	34.6	9.9	Field, forest
46	Semi-urban	45.2	5.1	Public buildings, block of flats
47	Rural	21.5	7.5	Field, forest
48	Semi-urban	56.4	11.1	Blocks of flats, asphalt road
49	Rural	32.4	9.9	Forest, meadow
50	Semi-urban	38.7	6.3	Detached houses, lake

maximum buffer size is justifiable, as the sea stores heat in a thick layer and releases latent heat with evaporation (see Hjort et al. 2011, Suomi & Käyhkö 2012). Taking into account the characteristics of the study area and the extent of the observation network, the theoretical maximum buffer size for 'water bodies' was determined to be 20 km. A theoretical maximum buffer size for 'relative elevation' was set to 1 km due to the present topographic features whose horizontal extent is often <1 km (Table 3).

Spearman's rank order correlation coefficients were calculated between monthly average temperatures and explanatory variables to determine the optimal buffer sizes. The correlation coefficients were

Table 2. Main steps of the analyses

Step	Data	Method
1. Buffer size optimization	12 variables, see Table 3	Spearman's rank correlation
2. Independent effects of urban variables	6 urban variables, see Table 3	Hierarchical partitioning
3a. Independent effects of 1 urban and 9 other variables	'Best urban variable' in Step 2, 6 non-urban and 3 spatial variables	Hierarchical partitioning
3b. Multivariate modelling	Same as in Step 3a	Linear regression

Table 3. Potential buffer sizes (i.e. the circle of influence) of the explanatory variables. The largest potential buffer size of the variable is the theoretical maximum buffer size of that variable. DEM = digital elevation model

Explanatory variables	Potential buffer sizes (radius, m)	Source of data
Urban		
Urban land use (Ul)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Large and urban buildings (Lb)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Floor area (Fl)	100, 200, 300	City of Turku (2010)
Block of flats (Bf)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Business and administrative areas (Ba)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Traffic areas (Tr)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Non-urban		
Industrial and warehouse areas (In)	100, 200, 300, 500, 1000, 2000	SLICES (2011)
Detached houses (Dh)	100, 200, 300, 500, 1000, 2000, 5000	SLICES (2011)
Agricultural lands (Ag)	100, 200, 300, 500, 1000, 2000, 5000	SLICES (2011)
Water bodies (Wa)	100, 500, 1000, 1500, 2000, 5000, 10000, 20000	SLICES (2011), CORINE (EEA 2006)
Relative elevation (Re)	100, 200, 300, 500, 1000	DEM (NLS 2009)
Tasseled cap greenness (Gr)	100, 200, 300, 500, 1000, 2000, 5000	Landsat ETM+ images (Crist & Cicone 1984, Hjort & Luoto 2006)

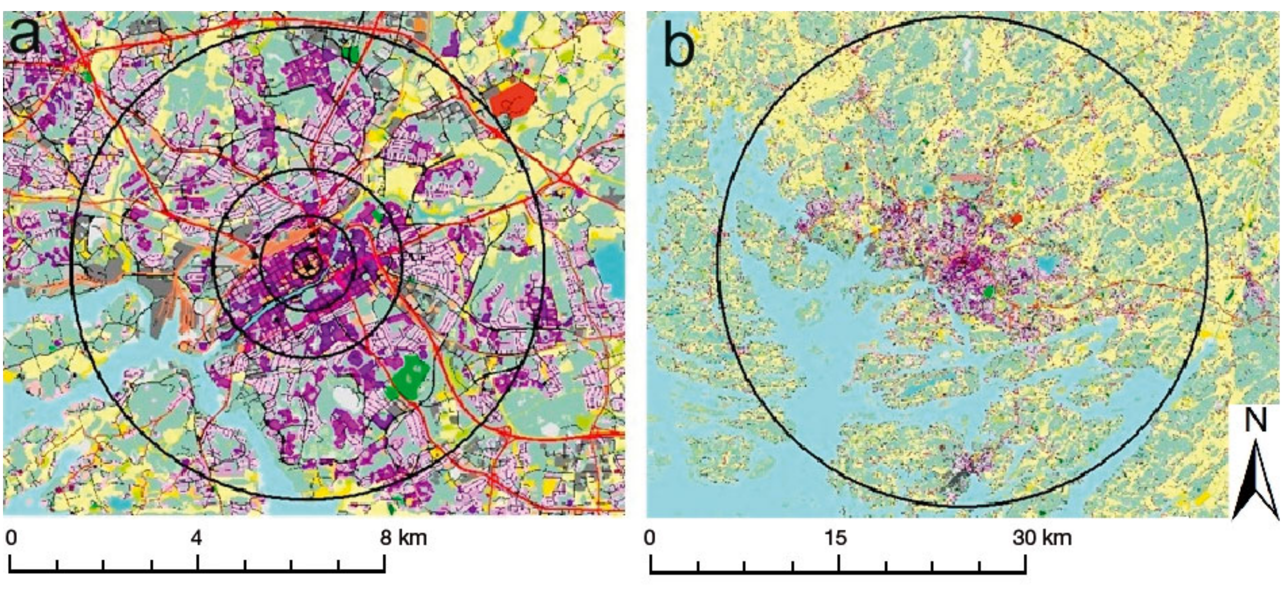

Fig. 3. Examples of theoretical maximum buffer sizes of the explanatory variables in relation to the study area. The centre of the buffers is the marketplace. Radius of the buffers: (a) 300, 1000, 2000 and 5000 m, (b) 20 000 m

calculated for each of the potential buffer sizes (Table 3). For most of the variables, selection of the optimal buffer size was based on the highest correlation coefficients. However, if a clear curve saturation or levelling off effect in the curve of the coefficients was detected, the point of levelling off rather than the highest value was considered the optimum buffer size. The correlations were calculated with SPSS 19 software.

After determining the optimal buffer size for each explanatory variable, the relative independent effect (hereinafter referred to as 'independent effect') of the variables was investigated in HP (Steps 2 and 3a in Table 2). The HP method is designed to overcome multicollinearity problems (i.e. intercorrelation among explanatory variables) by using a mathematical hierarchical theorem by which the explanatory capacities of a set of explanatory variables can be estimated (Chevan & Sutherland 1991, MacNally 1996). HP employs goodness-of-fit measures for each of 2^k possible models for k explanatory variables. In HP, the variances are partitioned so that the total independent contribution of a given explanatory variable can be estimated (Chevan & Sutherland 1991). For example, the independent contribution of the 'floor area' variable is estimated by comparing goodness-of-fit measures for all models including the 'floor area' variable with their reduced version (i.e. the exact same model but without 'floor area') within each hierarchical level; e.g. with 3 explanatory variables (2^3), there are 8 possible models at 4 hierarchical levels (models with only intercept, 1, 2 and 3 variables). The average improvement in fit for each hierarchical level that includes 'floor area' is then averaged across all hierarchies, giving the independent contribution of the 'floor area' variable. Thus, HP allows one to differentiate between those explanatory variables whose independent, as distinct from partial, correlation with a response variable may be important, from variables that have little independent effect on the studied phenomena. In the present study, HP was conducted using the 'hier.part' package version 1.0-3 in statistical software R version 2.11.1 (http://r-project.org). The goodness-of-fit measure we used was root-mean-square prediction error (RMSPE) (e.g. MacNally 2000). For every response variable, several regression models were generated (64 in Step 2, and 1024 in Step 3a) to reveal the most meaningful explanatory variables in determining monthly average temperatures.

In order to find the urban variable with the strongest independent effect, the 6 intercorrelated urban variables were compared with each other in HP (Step 2 in Table 2). Next, the selected urban variable of each month was taken into the further analysis, where the urban variable was modelled with 9 other explanatory variables in HP (Step 3a in Table 2). In this step, spatial variables, namely eastern (X) and northern (Y) map coordinates (Finland Uniform Coordinate System) and 'distance from city centre', were included in the HP in addition to the 6 land use and environmental variables. The spatial variables were employed to cover the potential spatial structures (i.e. trend and spatial autocorrelation) in the responses not covered by any other explanatory variables in the analyses. Consequently, the main focus was on urban and non-urban variables, whereas the spatial variables served only as a control in this study.

As the second part of the final step, a multivariate LR (for LR, see Sokal & Rohlf 1995) was employed to model the observed temperature differences in the study area (Step 3b in Table 2). The calibration of the LR model was performed utilizing the standard *lm* function in R. The LR model optimization was based on a step-wise approach and Akaike's information criterion (AIC) (Akaike 1974, Burnham & Anderson 1998). The normality of the distribution of variables was tested before the HP and LR analyses. For cases whose frequency-size distribution was not normal, different transformations were tested. If the variable could not be transformed to meet the normality assumption, the original variable was used (Sokal & Rohlf 1995). In general, LR has been used in several UHI studies (e.g. Unger et al. 2001, Kim & Baik 2002, Unger 2006, Szymanowski & Kryza 2009, 2011, Yokobori & Ohta 2009), but to our knowledge, HP is a new method in the UHI context. HP has, however, been successfully utilized in the biosciences (e.g. MacNally 2000, Heikkinen et al. 2004) and geosciences (e.g. Hjort et al. 2007, Hjort & Luoto 2009).

4. RESULTS

4.1. Optimal buffer sizes

For urban variables, the radius of the optimal buffer sizes varied between 300 and 1000 m, the larger one being the most common value (Table 4). Among the other variables, the optimal buffer size varied between 100 and 20 000 m, the mode being 2000 m. The average optimum buffer size of urban variables was ca. 700 m, and for non-urban variables it was 2500 m. Of all the variables, 'water bodies' had the largest average buffer size, 6100 m, and 'relative elevation' the smallest one, 200 m.

Table 4. Determined optimal buffer sizes (m) and corresponding Spearman's rank order correlation coefficients (*italics*) of the explanatory variables. Average (avg.) values (rounded to the nearest hundred) are calculated only for the buffer sizes. Significance: *p < 0.05, **p < 0.005. See Table 3 for abbreviations

Variable	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Avg.
Ul	300	300	1000	1000	1000	300	300	300	300	300	300	300	500
	0.481**	0.544**	0.839**	0.836**	0.838**	0.831**	0.787**	0.764**	0.544**	0.421**	0.391**	0.489**	
Lb	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	0.548**	0.637**	0.860**	0.881**	0.879**	0.879**	0.849**	0.822**	0.589**	0.492**	0.454**	0.527**	
Fl	300	300	300	300	300	300	300	300	300	300	300	300	300
	0.478**	0.560**	0.876**	0.897**	0.903**	0.913**	0.852**	0.832**	0.570**	0.408**	0.368**	0.458**	
Bf	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	0.403**	0.509**	0.798**	0.800**	0.801**	0.789**	0.716**	0.696**	0.429**	0.329*	0.298*	0.387**	
Ba	300	300	1000	1000	1000	1000	1000	1000	300	300	300	300	700
	0.495**	0.558**	0.832**	0.844**	0.863**	0.861**	0.796**	0.767**	0.580**	0.466**	0.417**	0.498**	
Tr	300	300	1000	1000	1000	1000	1000	1000	300	300	300	300	700
	0.494**	0.542**	0.834**	0.889**	0.849**	0.852**	0.797**	0.760**	0.536**	0.426**	0.423**	0.517**	
Avg. urban	500	500	900	900	900	800	800	800	500	500	500	500	700
In	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
	0.408**	0.481**	0.575**	0.585**	0.575**	0.562**	0.623**	0.612**	0.477**	0.402**	0.380**	0.391**	
Dh	1000	1000	5000	5000	5000	5000	5000	5000	1000	1000	1000	1000	3000
	−0.406**	−0.333*	0.525**	0.571**	0.542**	0.546**	0.443**	0.420**	−0.357*	−0.443**	−0.424**	−0.374**	
Ag	5000	5000	500	500	500	500	500	500	5000	5000	5000	5000	2800
	−0.787**	−0.815**	−0.695**	−0.614**	−0.692**	−0.707**	−0.671**	−0.709**	−0.701**	−0.758**	−0.772**	−0.729**	
Wa	5000	5000	500	500	500	2000	20000	20000	5000	5000	5000	5000	6100
	0.589**	0.578**	−0.151	−0.131	−0.116	−0.110	0.146	0.233	0.510**	0.645**	0.664**	0.559**	
Re	100	100	500	300	300	300	100	100	100	100	100	100	200
	0.251	0.262	0.411**	0.311*	0.358*	0.342*	0.312*	0.349*	0.256	0.184	0.170	0.184	
Gr	2000	2000	200	200	200	200	200	200	2000	2000	2000	2000	1100
	−0.872**	−0.879**	−0.755**	−0.758**	−0.786**	−0.818**	−0.798**	−0.788**	−0.830**	−0.832**	−0.810**	−0.807**	
Avg. non-urban	2500	2500	1500	1400	1400	1700	4600	4600	2500	2500	2500	2500	2500
Avg. all	1500	1500	1200	1200	1200	1200	2700	2700	1500	1500	1500	1500	1600

On a monthly basis, the average buffer size of all the variables was largest in July and August (2700 m) and smallest in April and May (1200 m). For the urban variables, the average buffer size was largest (900 m) from March to May, and smallest (500 m) from September to February. For non-urban variables, the average buffer size was largest (4600 m) in July and August, and smallest (1400 m) in April and May (Table 4). It should be noted that the average values of the optimal buffer sizes are used here only to illustrate the differences between variables and between variable groups, as well as seasonal differences of the variable groups.

4.2. Independent effects of urban variables in HP

Based on HP, the 'large and urban buildings' variable had the highest independent explanatory power of the 6 urban variables, the average proportion being 24.9% (Table 5). On average, the weakest urban variable, with an independent effect of 12.2%, was 'urban land use'. Seasonally, 'floor area' had the largest independent effect in spring and early summer (from March to June), whereas in all other months, 'large and urban buildings' was the

Table 5. Average independent effects of urban variables based on hierarchical partitioning analysis. See Table 3 for abbreviations

Urban variable	Independent effect (%)
Ul	12.2
Lb	24.9
Fl	19.6
Bf	12.3
Ba	15.5
Tr	15.5

strongest variable (hereinafter the term 'best urban variable' is used for the strongest variable based on HP analysis). The highest and lowest single independent effects (32.1 and 9.5%, respectively) existed in November, the lowest being that of the 'urban land use' variable (Fig. 4).

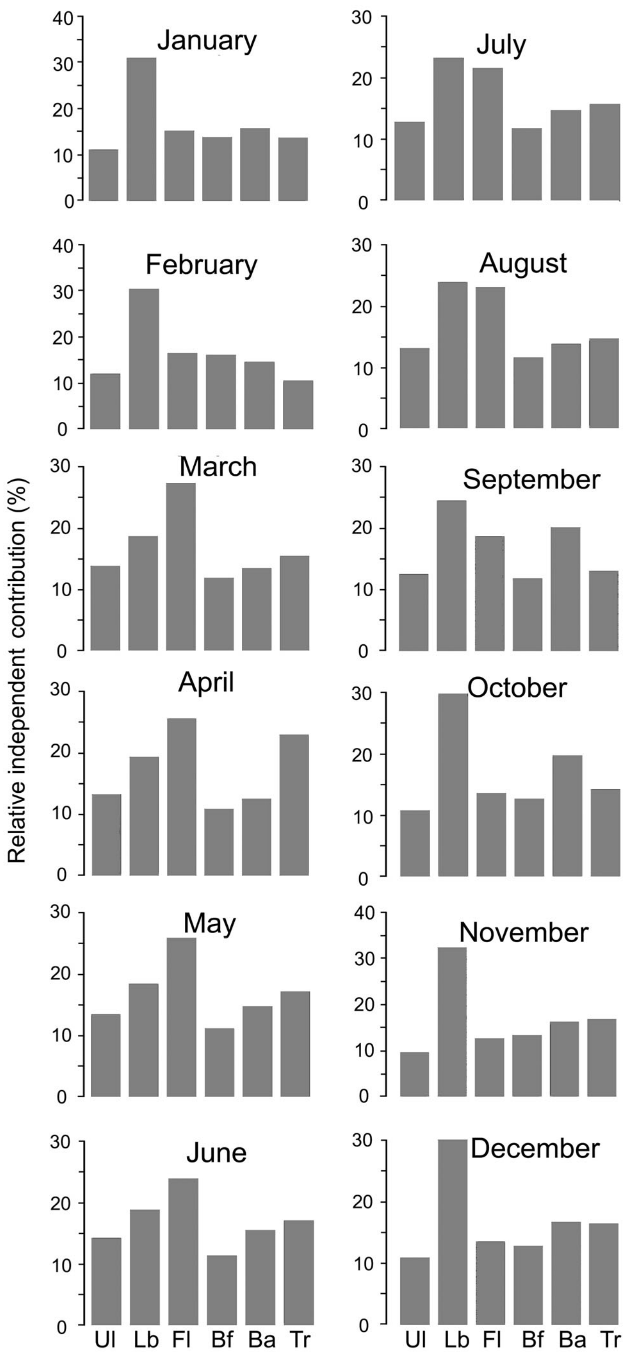

Fig. 4. Independent effects (% of total independently explained variance) of the urban variables in explaining temperature variability by month, as estimated from hierarchical partitioning. See Table 3 for abbreviations

4.3. Independent effects of the best urban variable and non-urban variables in HP

The HP results indicated that the 'best urban variable' had on average the highest (21.2%) independent effect on temperatures. The effect of 'greenness' was almost as large (21.0%). The effect of 'relative elevation' was on average the weakest (1.1%, Table 6). On a monthly basis, the 'greenness' variable had the highest independent explanatory power in the 7 mo from August to February, whereas in other months, the 'best urban variable' was dominant. The highest single independent effect (30.7%) existed for 'floor area' in March and the lowest (0.3%) for 'relative elevation' in April (Fig. 5).

4.4. Linear regression

The results of LR modelling are presented in Table 7. According to the residual plots, the assumption of normal errors was appropriate. The adjusted R^2 values of the models varied between 0.71 and 0.91, being highest in March and lowest in September. The 'best urban variable' was the most important variable during 5 months, whereas the 'water bodies' variable dominated in 4 months. Seasonally, the urban variables dominated mostly in spring, and 'water bodies' in late autumn and early winter. 'Greenness' was the most important variable in July and August, and 'distance from city centre' in May. The impact of 'best urban' and 'water body' variables on temperature was constantly positive, whereas the 'greenness' and 'distance from city centre' variables had cooling effects.

Table 6. Average independent effects of the 'best urban variable' (Ur) and non-urban variables. Distance from city centre (Di) is measured from the centrepoint of the marketplace. X = eastern map coordinate, Y = northern map coordinate. See Table 3 for other abbreviations

Variable	Independent effect (%)
Ur	21.2
In	5.4
Dh	5.8
Ag	12.1
Wa	5.1
Re	1.1
Gr	21.0
X	5.8
Y	5.5
Di	17.0

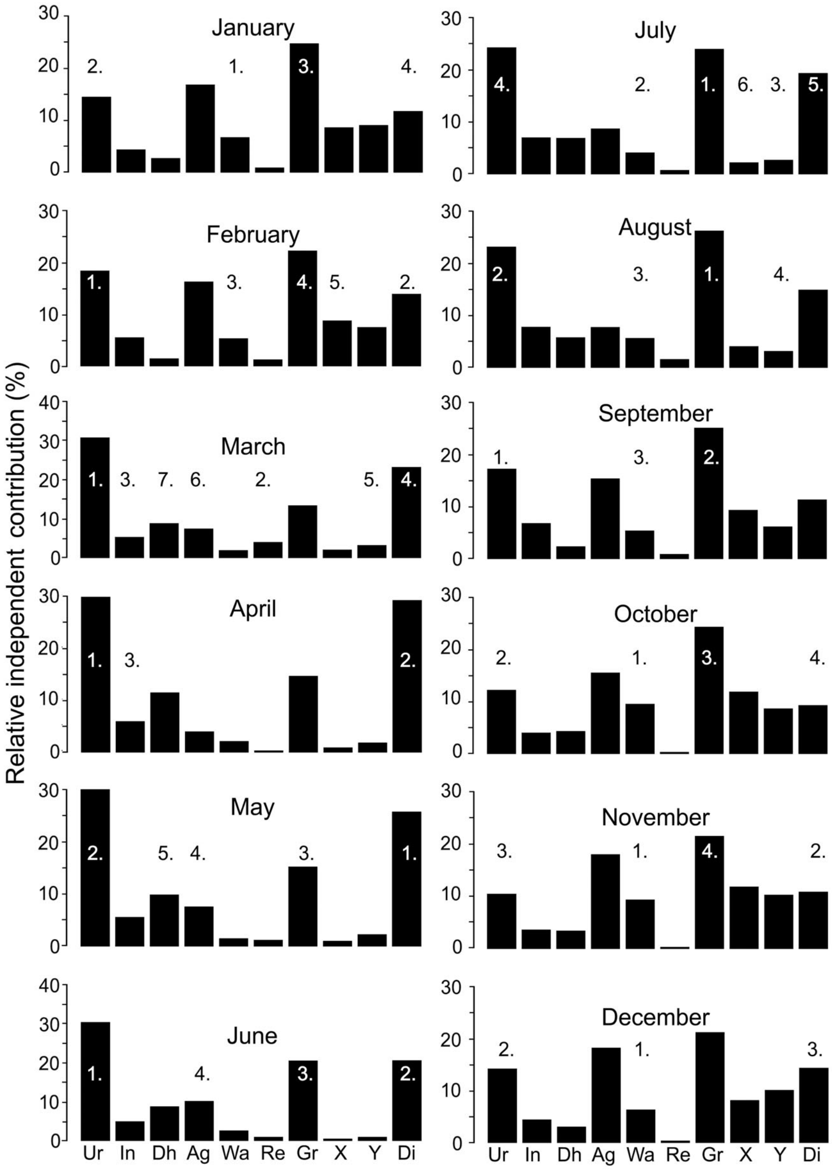

Fig. 5. Independent effects (% of total independently explained variance) of the 'best urban variable' ('large and urban buildings' or 'floor area') and 9 other explanatory variables in explaining temperature variability by month, as estimated from hierarchical partitioning. Numbered ranks indicate the relative importance of the variables in the final linear regression models (see Table 7). See Tables 3 and 6 for abbreviations

5. DISCUSSION

The results of this study clearly indicate that when possible, it is important to explore the optimum buffer sizes of explanatory variables before further analyses of UHI. This is in contrast with the common practice in urban climate studies, where the employed buffer sizes have often been based on the literature or site-specific consideration. The results also demonstrate that the optimal buffer size of a certain variable can have large seasonal variation, which is important to take into account in UHI studies.

The determined optimal buffer sizes of urban variables were either 300 or 1000 m and thus settling under the theoretical maximum buffer size set before the analyses. The observed buffer sizes were consistent with Oke's (2006) estimation of 500 m for the circle of influence, the size that has been used by e.g. Eliasson & Svensson (2003) and Suomi & Käyhkö (2012). The determined buffer sizes were, however, clearly larger than the 25 to 50 m estimated by Giridharan & Kolokotroni (2009). The difference in optimum buffer sizes can, however, be partly due to different explanatory variables used in the studies. The mode and average size of the non-urban variables' optimal buffers were larger than those of urban variables. This observation is in line with Stewart & Oke's (2009) and Giridharan & Kolokotroni's (2009) conclusions that the circle of influence is larger in open and sparsely built areas than in a densely built environment.

The optimal buffer sizes of urban variables were, on average, largest in spring, and smallest in autumn and winter. This finding is analogous to that of Giridharan & Kolokotroni (2009), who concluded that the optimal buffer size of on-site variables was smaller in winter than in summer. For non-urban variables, seasonal variation in optimal buffer sizes differed from the urban variables in that the maximum sizes occurred in July and August and the smallest in spring. This largely reflects the variation of the optimal buffer size of 'water bodies' between spring and summer, i.e. 500 m in spring to 20 000 m in July and August (cf. Table 4). In variable-specific comparisons, the optimal buffer size varied seasonally more uniformly among urban variables than among non-urban variables; the optimal buffer size of 'detached houses' was largest (5000 m) from March to September, whereas at the same time the optimal buffer size of 'greenness' and 'agricultural lands' was smallest

Table 7. Linear regression (LR). The relative importance (rank) of the explanatory variable is also indicated (variables in the final LR models were ranked based on Akaike's information criterion). + and –: direction of the effect of the explanatory variables. See Tables 3 & 6 for abbreviations

	Adjusted R^2	Ur	In	Dh	Ag	Wa	Re	Gr	X	Y	Di
Jan	0.82	2,+				1,+		3,–			4,–
Feb	0.85	1,+				3,+		4,–	5,–		2,–
Mar	0.91	1,+	3,+	7,–	6,–		2,+			5,–	4,–
Apr	0.88	1,+	3,+								2,–
May	0.87	2,+		5,–	4,–			3,–			1,–
Jun	0.86	1,+			4,–			3,–			2,–
Jul	0.84	4,+				2,+		1,–	6,+	3,+	5,–
Aug	0.86	2,+				3,+		1,–		4,+	
Sep	0.71	1,+				3,+		2,–			
Oct	0.80	2,+				1,+		3,–			4,–
Nov	0.82	3,+				1,+		4,–			2,–
Dec	0.74	2,+				1,+					3,–

(200 and 500 m, respectively). Such a contrasting pattern was not observed among the urban variables. The more heterogeneous optimal buffer size variation among the non-urban variables is probably due to the more versatile spectrum of factors that they reflect. In addition, for the urban variables, the correlation between them and temperatures is in every case positive, whereas for the non-urban variables, the direction of the correlation varies both between and within variables.

Based on HP results for the urban variables, 'large and urban buildings' and 'floor area' showed the highest independent effect on temperatures, ca. 20 to 30 % depending on the month. 'Floor area' dominated in spring and early summer, whereas 'large and urban buildings' dominated in other seasons. As the latter variable is a combination of blocks of flats, administrative and business areas and industrial and warehouse areas, it mostly reflects the warming impact of buildings, similar to the 'floor area' variable. Logically, the variables consisting of buildings and the variable consisting mainly of asphalt surfaces, namely 'traffic areas', showed the greatest difference in winter and late autumn, when the solar heat storage of streets is minimal. Somewhat surprising, however, was the finding that 'traffic areas' contributed in UHI only a little also in summer. This can be partly due to the fact that 'traffic areas' included also relatively open suburban and rural areas, where released radiation was not trapped in nearby buildings, as was the case in the city centre. The larger impact of building-based variables compared to traffic areas was also detected by Hart & Sailor (2009) in daytime UHI modelling in June in Portland, Oregon. They found clear differences between weekdays and weekends, with the average heat island intensity near the main arterial roads being 1.9°C on weekdays and 0.6°C on weekends. This indicates a relatively large effect of traffic compared to solar heat storage in road areas. In Turku, the anthropogenic component of the warming effect of traffic areas is relatively low on smaller suburban and rural roads, which is one reason for the weak contribution of the 'traffic areas' variable on UHI. In addition, traffic has a relatively minor impact at night, and the heat flux from horizontal surfaces decreases towards the end of the night (see e.g. Kuttler et al. 1996), which also diminishes the warming effect of traffic areas on average temperatures.

In a comparison between the 'best urban variable' and the non-urban variables, the urban variable had on average the highest independent impact based on HP. In LR, the 'best urban variable' was most often the strongest, (in 5 out of 12 months) and it was included in the model every month. Based on both methods, the dominance of the 'best urban variable' was most obvious in spring and early summer. This may be due to relatively cold sea areas during spring, which highlights the heating effect of the urban variables. In Gothenburg, the variables reflecting urban effects were also strongest compared to other variables in spring nights (Eliasson & Svensson 2003).

Among the non-urban variables, 'greenness' and 'water bodies' dominated, although with some differences, depending on the employed method; 'greenness' dominated in HP and 'water bodies' in LR. One should not directly compare the results of LR to HP, or consider either of these more correct, as these are rather different modelling approaches. HP is used to reveal independently the most important variables among correlated explanatory variables, whereas the output of LR is a multivariate model (Chevan & Sutherland 1991, Sokal & Rohlf 1995). In the present study, the explanatory power of 'greenness' was partly or mostly covered by other variable(s) in the final LR models (cf. MacNally 2000). Compared with the other rural variable, namely 'agricultural lands', 'greenness' clearly had stronger explanatory power. This may be due to the fact that forests were not included in the 'agricultural lands' variable, but their effect was seen in 'greenness' values; in forested and vegetated areas, evapotranspiration is relatively high, resulting in lower temperatures. Also, 'greenness' was the most important variable in LR in mid- and late summer (July and August), when the abundance of vegetation is highest.

In addition to the strong effect of 'greenness' in summer, when using HP, its importance was also revealed in autumn and winter, when the verdancy of vegetation is low. This finding is believed to originate at least partially in the low amount of anthropogenic heat sources in green areas (Crist & Cicone 1984). Moreover, 'greenness' as a continuous variable was able to better reflect the lack of heat sources than the categorical variables (e.g. agricultural areas). NDVI has been commonly used in UHI studies (see e.g. Chen et al. 2006, Zhang et al. 2009, Szymanowski & Kryza 2011), but in our analyses, 'greenness' outperformed NDVI in the preliminary correlation analysis. Consequently, the use of a greenness index is worth considering in urban climate studies.

Methodologically, traditional regression techniques may distort inferences about the relative importance of explanatory variables (Chevan & Sutherland 1991) because they lack sensitivity to detect intercorrelation between the explanatory variables (Sokal & Rohlf 1995). The HP method used in the present study provided a successful separate measure of the amount of variation explained independently by 2 or more explanatory variables and, therefore, it helped to make deductions on the important environmental determinants (MacNally 1996, 2000). However, one should be aware of a couple of constraints when using this approach. Firstly, HP does not produce a model and, secondly, the importance of polynomial variables (i.e. nonlinear relationships between response and explanatory variables) cannot be assessed. Consequently, HP is a suitable method for exploring causal variables, but it is not suited for e.g. predicting temperatures. However, the relationships between temperature and environmental variables in UHI studies have often been reported to be linear, and thus the difficulty of investigating nonlinear responses is potentially a minor shortcoming in the UHI context (e.g. Eliasson & Svensson 2003). Finally, rather than considering LR and HP as exclusionary methods, it is advisable to combine them in multivariate analyses, especially if explanatory variables are highly intercorrelated (MacNally 2000, Heikkinen et al. 2004, Hjort & Luoto 2009).

6. CONCLUSIONS

Based on our results, we can draw 4 main conclusions. (1) The comparison of the 6 urban and 6 non-urban explanatory variables on different scales revealed that the optimal buffer size radius of the urban variables varied seasonally, being either 300 or 1000 m. For the non-urban variables, the optimal monthly buffer size showed larger variation; between 100 and 20 000 m, the mode being 2000 m.

(2) The optimal buffer size of the urban variables was largest in spring and smallest in winter and autumn. For the non-urban variables, no clear seasonal trend was observed.

(3) Of the 6 intercorrelated urban variables, 'large and urban buildings' and 'floor area' had the largest independent impact on temperatures. This indicates that in our high-latitude study area, heat released from buildings dominates in UHI formation throughout the year over heat released from horizontal urban surfaces.

(4) In the comparison between the 'best urban variable' and 6 non-urban variables, HP revealed the high importance of the 'urban' and Landsat ETM+ based 'greenness' variables, whereas LR highlighted the importance of the 'urban' and 'water body' variables.

Acknowledgements. The Finnish Cultural Foundation's Varsinais-Suomi Regional Fund, Emil Aaltonen Foundation and Turku University Foundation provided financial support in the form of personal scholarships (J.S.). The TURCLIM project collaborates with the Turku Environment and City Planning Department, whose assistance is of great value. The long-term reference weather data were provided by the Finnish Meteorological Institute.

LITERATURE CITED

Akaike H (1974) A new look at statistical model identification. IEEE Trans Automat Control 19:716–723

Alalammi P (1987) Atlas of Finland: 131, Climate. National Board of Survey, Geographical Society of Finland, Helsinki

Burnham KP, Anderson DR (1998) Model selection and inference: a practical information-theoretic approach. Springer, New York, NY

Chen XL, Zhao HM, Li PX, Yin ZY (2006) Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens Environ 104:133–146

Chevan A, Sutherland M (1991) Hierarchical partitioning. Am Stat 45:90–96

City of Turku (2001) General plan 2020. Environmental and City Planning Department, Plan Office, Turku

City of Turku (2010) Floor area of the buildings. Real Estate Department, Turku

Costa A, Labaki L, Araujo V (2007) A methodology to study the urban distribution of air temperature in fixed points. In: Santamouris M, Wouters P (eds) 2nd PALENC and 28th AIVC Conference: Building Low Energy Cooling and Advanced Ventilation Technologies in the 21st Century, 27–29 Sep 2007, Crete, p 227–230

Cotton WR, Pielke RA (1995) Human impacts on weather and climate. Cambridge University Press, Cambridge

Crist EP, Cicone RC (1984) Application of the tasseled cap concept to simulated Thematic Mapper data. Photogramm Eng Remote Sens 50:343–352

EEA (European Environment Agency) (2006) CORINE land cover. EEA, Denmark

Eliasson I, Svensson MK (2003) Spatial air temperature variations and urban land use — a statistical approach. Meteorol Appl 10:135–149

FMI (Finnish Meteorological Institute) (2011) Ilmasto-opas. http://ilmasto-opas.fi/fi/

FMI (Finnish Meteorological Institute) (2012) Tilastoja Suomen ilmastosta 1981–2020. www.ilmatieteenlaitos.fi

Giridharan R, Kolokotroni M (2009) Urban heat island characteristics in London during winter. Sol Energy 83: 1668–1682

Giridharan R, Lau SSY, Ganesan S, Givoni B (2007) Urban design factors influencing heat island intensity in high-rise high-density environments of Hong Kong. Build Environ 42:3669–3684

Giridharan R, Lau SSY, Ganesan S, Givoni B (2008) Lowering the outdoor temperature in high-rise high-density residential developments of coastal Hong Kong: vegetation influence. Build Environ 43:1583–1595

Hart MA, Sailor DJ (2009) Quantifying the influence of landuse and surface characteristics on spatial variability in the urban heat island. Theor Appl Climatol 95:397–406

Heikkinen RK, Luoto M, Virkkala R, Rainio K (2004) Effects of habitat cover, landscape structure and spatial variables on the abundance of birds in an agricultural–forest mosaic. J Appl Ecol 41:824–835

Hicks BB, Callahan WJ, Hoekzema MA (2010) On the heat islands of Washington, DC, and New York City, NY. Boundary-Layer Meteorol 135:291–300

Hjort J, Luoto M (2006) Modelling patterned ground distribution in Finnish Lapland: an integration of topographical, ground and remote sensing information. Geogr Ann A 88:19–29

Hjort J, Luoto M (2009) Interaction of geomorphic and ecologic features across altitudinal zones in a subarctic landscape. Geomorphology 112:324–333

Hjort J, Luoto M, Seppälä M (2007) Landscape scale determinants of periglacial features in subarctic Finland: a grid-based modelling approach. Permafrost Periglac Process 18:115–127

Hjort J, Suomi J, Käyhkö J (2011) Spatial prediction of urban-rural temperatures using statistical methods. Theor Appl Climatol 106:139–152

Houet T, Pigeon G (2011) Mapping urban climate zones and quantifying climate behavior — an application on Toulouse urban area (France). Environ Pollut 159: 2180–2192

Kim Y, Baik J (2002) Maximum urban heat island intensity in Seoul. J Appl Meteorol 41:651–659

Kolokotroni M, Giridharan R (2008) Urban heat island intensity in London: an investigation of the impact of physical characteristics on changes in outdoor air temperature during summer. Sol Energy 82:986–998

Kuttler W, Barlag AB, Rossmann F (1996) Study of the thermal structure of a town in a narrow valley. Atmos Environ 30:365–378

Landsberg HE (1981) The urban climate. Academic Press, London

MacNally R (1996) Hierarchical partitioning as an interpretative tool in multivariate inference. Aust J Ecol 21: 224–228

MacNally R (2000) Regression and model-building in conservation biology, biogeography and ecology: the distinction between — and reconciliation of — 'predictive' and 'explanatory' models. Biodivers Conserv 9:655–671

NLS (National Land Survey of Finland) (2009) Digital elevation model. NLS, Helsinki

Oke TR (1973) City size and the urban heat island. Atmos Environ 7:769–779

Oke TR (1987) Boundary layer climates, 2nd edn. Routledge, London

Oke TR (2006) Initial guidance to obtain representative meteorological observations at urban sites. Instruments and Observing Methods Report No. 81. World Meteorological Organization, Geneva

Peel MC, Finlayson BM, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. Hydrol Earth Syst Sci 11:1633–1644

Sakakibara Y, Matsui E (2005) Relation between heat island intensity and city size indices/urban canopy characteristics in settlements of Nagano basin, Japan. Geogr Rev Jpn 78:812–824

Seinä A, Peltola J (1991) Duration of the ice season and statistics of fast ice thickness along the Finnish coast 1961–1990. Finn Mar Res 258:1–46

Seinä A, Eriksson P, Kalliosaari S, Vainio J (2006) Ice seasons 2001-2005 in Finnish sea areas. Report Series No. 57. Finnish Institute of Marine Research, Helsinki

SLICES (2011) SLICES Maankäyttö. National Land Survey of Finland. www.maanmittauslaitos.fi/digituotteet/slices-maankaytto

Sokal RR, Rohlf FJ (1995) Biometry. WH Freeman, New York, NY

Stewart ID, Oke TR (2009) Classifying urban climate field sites by 'local climate zones': the case of Nagano, Japan. In: Preprints, 7th Int Conf Urban Climate, 29 Jun–3 Jul, Yokohama. www.ide.titech.ac.jp/~icuc7/extended_abstracts/pdf/385055-1-090515165722-002.pdf

Suomi J, Käyhkö J (2012) The impact of environmental factors on urban temperature variability in the coastal city of Turku, SW Finland. Int J Climatol 32:451–463

Szymanowski M, Kryza M (2009) GIS-based techniques for urban heat island spatialization. Clim Res 38:171–187

Szymanowski M, Kryza M (2011) Application of geographically weighted regression for modelling the spatial structure of urban heat island in the city of Wroclaw (SW Poland). Proc Environ Sci 3:87–92

Unger J (2006) Modelling of the annual mean maximum urban heat island using 2D and 3D surface parameters. Clim Res 30:215–226

Unger J, Sümeghy Z, Gulyás Á, Bottanyán Z, Mucsi L (2001) Land-use and meteorological aspects of the urban heat island. Meteorol Appl 8:189–194

Wen X, Yang X, Hu G (2011) Relationship between land cover ratio and urban heat island from remote sensing and automatic weather stations data. J Indian Soc

Remote Sens 39:193–201
Wong NH, Jusuf SK, Syafii NI, Chen Y, Hajadi N, Sathyanarayanan H, Manickawasagam YV (2011) Evaluation of the impact of the surrounding urban morphology on building energy consumption. Sol Energy 85:57–71
Yokobori T, Ohta S (2009) Effect of land cover on air temperatures involved in the development of an intra-urban heat island. Clim Res 39:61–73
Zhang Y, Odeh IOA, Han C (2009) Bi-temporal characterization of land surface temperature in relation to impervious surface area, NDVI and NDBI, using a sub-pixel image analysis. Int J Appl Earth Obs Geoinf 11:256–264

Editorial responsibility: Helmut Mayer, Freiburg, Germany

Submitted: May 10, 2012; Accepted: July 10, 2012
Proofs received from author(s): November 12, 2012