High-Resolution Temperature Mapping by Geostatistical Kriging with External Drift from Large-Eddy Simulations

MIKHAIL VARENTSOV

*Research Computing Center/Faculty of Geography, Moscow State University, and
A. M. Obukhov Institute of Atmospheric Physics, Moscow, Russia*

IGOR ESAU AND TOBIAS WOLF

Nansen Environmental and Remote Sensing Centre, Bergen, Norway

(Manuscript received 13 June 2019, in final form 14 November 2019)

ABSTRACT

Detailed temperature maps are required in various applications. Any temperature interpolation over complex terrain must account for differences in land cover and elevation. Local circulations and other small-scale factors can also perturb the temperature. This study considers the surface air temperature T mapping with geostatistical kriging. The kriging methods are implemented for both T and temperature anomalies ΔT , defined as difference between T at a given location and T at the same elevation in the free atmosphere. The study explores the parallelized atmospheric large-eddy simulation (LES) model (PALM) as a source for variogram and external drift in the kriging methods. Ten kriging methods for the temperature mapping have been considered: ordinary kriging (OK) of T and ΔT with variogram derived from the observations (methods 1 and 2, correspondingly); OK of T and ΔT with variogram derived from LES data (3 and 4); universal kriging with external drift (KED) that utilizes the LES data (5 and 6); a weighted combination of KED of T and ΔT (method 7); and methods 5, 6, and 7 enhanced with additional “virtual” points in remote areas (methods 8, 9, and 10). These 10 methods are evaluated for eight typical weather situations observed in Bergen, Norway. Our results demonstrate considerable added value of the LES information for the detailed meteorological temperature mapping. The LES data improve both the variogram and the resulting temperature maps, especially in the remote mountain parts of the domain and along the coast.

1. Introduction

Detailed temperature maps are required to develop environmental and infrastructural projects in complex and heterogeneous domains. At larger scales, geographical temperature variations are reasonably accounted for through optimal spatial interpolation and retrospective modeling analysis (reanalysis). At local scales, however, meteorological observations are often too sparse

and nonrepresentative. Sophisticated spatial interpolation is needed to satisfy the interests of stakeholders (Taheri-Shahraiyini and Sodoudi 2017; Beck et al. 2018). This complication is rather unfortunate because the lack of knowledge about local temperature variations might impede identification of ecosystem refugia (e.g., Gregg et al. 2003; Dobrowski 2011; Miles and Esau 2016; Bois et al. 2018) as well as of environmental and health risks (e.g., Buscail et al. 2012; Hjort et al. 2016; Varentsov et al. 2018; Mironova et al. 2019). Detailed temperature maps are also required to optimize urban green space and urban comfort (e.g., Fernández et al. 2015; Zhang et al. 2017) and, more broadly, to improve urban management (e.g., Shao et al. 1997; Szymanowski and Kryza 2009; Zhang et al. 2011). Local and usually unobserved temperature anomalies may constitute adverse factors both for agriculture (e.g., frost damages in local relief

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/MWR-D-19-0196.s1>.

Corresponding author: Mikhail Varentsov, mvar91@gmail.com

DOI: 10.1175/MWR-D-19-0196.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

depressions) and for humans (e.g., stress from too-hot or too-cold temperatures). As an extreme example of such anomalies, we would refer to the recent discovery of nearly -100°C temperatures in Antarctica. This temperature minimum was attributed to a shallow relief depression where no direct measurements were conducted (Scambos et al. 2018).

Detailed temperature maps are not easy to obtain. The strongest control factor for the surface air temperature T is a place's elevation. Reliance on the mean or standard vertical temperature gradient is widespread in temperature mapping (Ishida and Kawashima 1993). High-resolution surface topography is readily available for almost any inhabited area. Such a simplified temperature mapping, however, provides unsatisfactory results (e.g., Stahl et al. 2006; Benavides et al. 2007; Hofstra et al. 2008). Atmospheric temperature is also sensitive to thermal properties of the underlying surface and land-use–land-cover types (e.g., Smoliak et al. 2015; Esau and Miles 2018). In mountainous coastal areas, T is controlled by distance from the coast not less than by the vertical temperature gradient (e.g., Stahl et al. 2006; Ho et al. 2016). In addition, T is controlled by a multitude of local physical processes, such as radiation, cloudiness, and local atmospheric dynamics (Courault and Monestiez 1999; Ho et al. 2016; Wolf-Grosse et al. 2017).

The reviewed literature suggests that geostatistical interpolation (mapping) with kriging—a kind of best linear unbiased estimator (BLUE) methods—provides more realistic and physically more adequate temperature maps than other considered interpolation methods. Statistical mapping at the local scales requires a dense meteorological observational network. Such a network is usually unavailable. Despite the fact that the total number and density of amateur, irregular temperature observations have dramatically increased in the recent years (Meier et al. 2017), the certified, high-quality, regular observations remain sparse and scarce even in urban areas. In the case of a sparse observational network, universal global kriging with external drift has demonstrated certain advantages as compared to other methods (Lapen and Hayhoe 2003; Benavides et al. 2007; Hofstra et al. 2008). The kriging with external drift (KED) is a kind of a bivariate regression accounting for correlation between an observed (dependent) variable and additional information (independent variable). This independent variable is called external drift, and it is available both at sampling and interpolated site locations. For the temperature mapping, the elevation is frequently specified as the external drift (Hudson and Wackernagel 1994).

Increased computational capacity opens an opportunity to apply the kriging methods to the high-resolution temperature mapping, for example, to map urban heat islands. Szymanowski and Kryza (2009) applied 5 methods to 7 different cases of the urban heat island in Wroclaw (Poland) demonstrating advantages of the residual kriging methods with additional information from land-use datasets and satellite (Landsat) thematic images. Hengl et al. (2011) found that the local space–time regressions obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature products at 1 km resolution can explain up to 84% of the temperature variations in the Croatian mountains. The KED application with external drift from elevation increases the explained part to 91% of the total temperature variability. Smoliak et al. (2015) used kriging to estimate the magnitude and variability of the urban heat island in Minneapolis. Yao et al. (2013) found superior quality of the kriging methods applied to soil moisture interpolation within a domain with complex relief and fragmented land use.

Spatial interpolation with kriging methods is a popular approach. An extensive review by Li and Heap (2011, 2014) compares several spatial interpolation methods for environmental applications. The review concludes that, generally, the kriging methods produce results of high quality in terms of the data variation and the maximum absolute errors. If relevant additional information, for example, elevation or land-cover types, is available, the KED methods return superior results as compared to other interpolation methods. Courault and Monestiez (1999) showed that including information on the atmospheric dynamics, specifically on the local circulation patterns, can improve the kriging results even more. However, their work explores only the ordinary kriging aiming to create temperature maps for 10 regional circulation patterns, and the patterns were prescribed. Wackernagel et al. (2004) proposed a way to use advantages of dynamic model simulations. The model output is used as external drift in mapping of ozone concentrations in Paris. Independently, an interesting idea to utilize a meteorological model appeared in the DeGaetano and Belcher (2007) study. They interpolated modeled temperatures into locations with observations, and then obtained spatial corrections for model output data. But they did not use the KED methods that could naturally include the modeled temperatures.

In this study, we further explore applications of the high-resolution models to temperature mapping. We utilize a parallelized atmospheric large-eddy simulation model (PALM) (Maronga et al. 2015; Wolf-Grosse et al. 2017;

FIG. 1. Geographical location of the case study area of Bergen, Norway. The AWSs are shown by circles in the right panel. The relief of the area is shown according to ASTER DEM with the contours drawn every 50 m. The first (bold) contour shows 5 m MSL in DEM, which is the closest level to the real shoreline. (Images from Google Maps are used as the background.)

Wolf et al. 2020). The main advantage of using PALM is that the modeled temperature is delivered on a regular, fine-grained grid. Thus, areas with missing observations are covered by the model simulations. The model mesh has the horizontal grid spacing of 27 m. The modeled temperature is used in our study for calculations of variogram and for the external drift in the KED.

We evaluate 10 kriging methods applied to 8 demonstration cases in Bergen, Norway. This presentation has the following structure. Section 2 describes the study area, the observational network and the meteorological conditions of the case studies. Section 3 describes the PALM model and its simulation setup. Section 4 provides a description of the kriging methods and the temperature mapping. Section 5 provides the methods' evaluation using a jackknifing approach. Unfortunately, many reviewed publications lack a constructive and precise description of the applied kriging methods. Although our description of the kriging methods is necessarily brief and schematic, we believe that the description provided in this study is sufficient for implementation of all methods by other groups familiar with their theoretical foundations as described in, for example,

Wackernagel (2003) or Bivand et al. (2013). Section 6 outlines our conclusions and recommendations.

2. Datasets and the case study description

a. The case study location and relief

We evaluate the proposed temperature mapping in the complex and heterogonous area of Bergen (60.4°N, 5.3°E), which is found at the western coast of Norway. The study area has the size of about 40 km². The area includes numerous narrow sea inlets (fjords), narrow valleys, steep hills of 200–640 m height, elevated plateaus and rocky islands. Figure 1 shows the geographical location of the area.

To demonstrate universal applicability of the methodology, we utilize a global digital elevation model (DEM) from an Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). [The ASTER DEM is available from <https://asterweb.jpl.nasa.gov/gdem.asp> (Tachikawa et al. 2011a).] The spatial (horizontal) resolution of the ASTER DEM is 1 arcsec, which is equal to the grid steps of about $\Delta x = 15$ m and $\Delta y = 30$ m at the latitude of Bergen. The ASTER DEM is not perfect (Tachikawa et al. 2011b). For example, significant discrepancies are found between the ASTER DEM readings

TABLE 1. Locations of the automatic weather stations in Bergen, used in the case study.

No.	Station name	Lon (°N)	Lat (°E)	Sensor elevation (m MSL)	Sensor elevation (m AGL)	Surface elevation MSL from ASTER (m)
1	Ulriken	60.377	5.379	605	4	533
2	Ulriksbakken	60.379	5.373	408	2.5	332
3	Haukeland	60.380	5.362	64	10	47
4	Florida	60.384	5.332	48	30	10
5	Lovestakken	60.360	5.321	472	2.5	426
6	Strandafjellet	60.372	5.323	303	2.5	242
7	Nordnes	60.400	5.304	48	30	10

and the true elevations of our weather stations (see Table 1). We used the true elevation at every location where it is known to us.

b. Local meteorological data and local climate

Bergen is relatively well covered with meteorological observations. Several environmental institutions and the Geophysical Department of the University in Bergen (GFI) operate diverse meteorological instruments. Together, they constitute a Bergen meteorological testbed. Seven calibrated automatic meteorological stations (AWSs) provide data for this study (Table 1). The stations represent different microclimatic conditions (Valved 2012). The Nordnes AWS is located on a peninsula surrounded by the sea. The Florida and Haukeland AWSs are located in the lower, urban part of the Bergen valley. The Strandafjellet and Ulriksbakken AWSs are found on steep northern and western hill slopes correspondingly, whereas the Lovestakken and Ulriken AWSs are placed on summits at altitudes 472 and 605 m above sea level (MSL).

In addition to the AWS network, we utilize temperature scans from a passive angular scanning microwave radiometer MTP-5HE (MTP). The instrument is placed at the roof of the GFI at 45 m MSL. It scans the atmospheric temperature in the direction of the southern Bergen valley between Strandafjellet and Ulriksbakken between 45 and 1045 m MSL. (The MTP data are available from <https://veret.gfi.uib.no/?action=mtp>.) A number of studies (Kadygrov and Pick 1998; Esau et al. 2013; Wolf et al. 2014) describe the dataset and its quality. The entire unprocessed MTP dataset can be requested from the Nansen Environmental and Remote Sensing Centre (NERSC). This study does not include observations from uncertified and irregular meteorological stations, which are also numerous in the area.

We use hourly averaged AWS and MTP data. Their temperature readings were found in a reasonable agreement by a long-term intercomparison study (Esau et al. 2013). The readings were particularly close under windy and convective atmospheric conditions with no precipitation. The relationships between the AWS and

MTP data are further discussed in the online supplemental material (section S1). Both datasets were intensively used in several observational and modeling studies (Jonassen et al. 2012; Wolf et al. 2014; Wolf-Grosse et al. 2017). Unfortunately, the stations are concentrated in central Bergen and at the closest hills. It makes the observational network geographically lumpy. Moreover, the stations are installed at different elevations on the hillsides and at different distances from the Byfjorden sea inlet. This configuration renders irrelevant any unweighted interpolation methods.

Local temperature climate is characterized by a significant contrast between the sea and land surface temperatures. The land surface is usually warmer than the sea surface in late spring and summer, but it becomes colder in autumn and winter. These differences drive local circulations and intensify turbulent mixing in some areas, while air in the other areas can stagnate, aggravating the local temperature differences (Wolf-Grosse et al. 2017; Davy et al. 2017; Davy and Esau 2016). Relatively strong winds are frequently observed in Bergen (Jonassen et al. 2012). The surface wind directions are aligned with the valley axis revealing weak sensitivity to the direction of geostrophic winds above the mountains. These local wind systems are captured neither with the sparse observational networks nor with coarse-resolution models. Other physical factors may further exacerbate spatial temperature differences. The most important factor among them is differential solar heating of the differently oriented steep hill slopes surrounding the central valley.

c. The case studies

We consider 8 selected daytime cases for the demonstration and evaluation of the kriging methods. All cases were observed under different wind conditions in May–July 2011 (Table 2). The cases are sufficiently diverse to be relatively easy projected to other periods, areas and observational network's configurations. We focus the presentation on the northwesterly wind (NW) case. The information on the other cases can be found in the supplemental material (sections S3 and S4).

TABLE 2. Time and meteorological conditions of the selected cases. The conditions are related to the end of 1-h averaging periods. The NW case is used for demonstration in the main text, while other cases are given in the online supplemental material.

Case name	Date and time	Wind at Ulriken		Mean vertical temperature gradient $\nabla_z T^V$ (at Florida) (K km^{-1})	Incoming solar radiation at Florida (W m^{-1})
		Direction ($^\circ$)	Speed (m s^{-1})		
NN	1700 CET 21 Jul 2011	15	5.9	0.78	500
NE	1100 CET 21 Jul 2011	26	4.1	0.84	300
EE	1600 CET 17 Jul 2011	98	12.9	0.82	200
SE	1300 CET 18 Jul 2011	128	3.0	0.94	850
SS	1400 CET 16 Jun 2011	178	5.8	0.95	450
SW	1500 CET 12 Jun 2011	231	3.9	0.98	600
WW	1200 CET 21 Jun 2011	276	2.8	0.95	900
NW	1200 CET 27 May 2011	304	2.9	0.91	250

Temperature and wind observed during all 8 cases are shown in Fig. 2.

The selected cases are characterized by moderate wind ($2.8\text{--}12.9\text{ m s}^{-1}$) and insignificant cloudiness. Somewhat more clouds were observed only on 17 July in the EE case (see Table 3). The lower atmosphere was convectively stratified in all cases. The mean vertical temperature gradient obtained from the MTP-5HE data varied from 7.8 to 9.8 K km^{-1} , which corresponds to nearly dry adiabatic conditions. Relative humidity at 2 m above the ground was 70%–95%. Intensive summer solar radiation with the maximum diurnal flux of $250\text{--}900\text{ W m}^{-1}$ (according to the measurements at Florida, GFI) amplified the local temperature differences. All cases were observed between 1100 and 1700 of the central European time (CET) when maximum radiation heating is received by the southern and southeastern slopes of the hills. Due to very long days in May–July, the solar angle above the horizon does not change much between noon and 1700. The surface air temperature at Florida varied between $+11^\circ$ and $+16^\circ\text{C}$ with exception of 18 July ($+21^\circ\text{C}$). The sea surface temperature at Nordnes was between $+8^\circ$ and $+14^\circ\text{C}$ (<https://www.seatemperature.org/europe/norway/>).

3. The PALM model and simulated data

An essentially new element of this study is a high-resolution, large-eddy simulation (LES) model. We use the LES model as an additional source of information for the interpolation methods. But why do we not substitute the statistical interpolation directly with the modeled data? The model, as any code of computational fluid dynamics, reasonably computes fluid flows and turbulent mixing over a complex surface geometry. However, we lack observational input data to initialize the simulations properly, and we lack knowledge about boundary conditions at scales of the model's grid. There are other technological challenges too. One must spin up and run the LES model in a relatively small domain that is not

conjugated with the smallest domain of meteorological models (Muñoz-Esparza et al. 2017). In result, without proper initialization, lateral and surface boundary conditions, and a proper canopy module, the LES model produces rather idealized simulations. Nevertheless, it is likely reproducing the main features of the orographically stirred flows and the vertical turbulent exchanges, and therefore, spatial variability of atmospheric dynamics over the realistic relief that is consistent with the applied large-scale forcing (Salveti et al. 2011; Bosveld et al. 2014; Schalkwijk et al. 2015). We seek for this consistency between the modeled winds and the prescribed surface features.

The PALM (Maronga et al. 2015) serves as the LES model in this study. The PALM was frequently and successfully applied to study atmospheric dynamics over complex and physically heterogeneous surfaces (e.g., Gronemeier et al. 2017; Letzel et al. 2008; Wolf-Grosse et al. 2017; Resler et al. 2017). We run PALM version 3.10 (revision 1424). Table 2 lists the simulated cases. Since there is no robust procedure for the data assimilation and model initialization working on turbulent scales, a “trial-and-fail” approach must be used. Each case was first simulated by several (3 to 6) lower-resolution PALM runs with slightly different initial and boundary settings. The challenge is to counteract overmixing in the model simulations. We look for such initial conditions that will provide the observed case profiles of wind and temperature after the spinup and local flow development periods in the simulations. The best simulations were rerun at high-resolution grid using 2048 computation cores at the Hexagon—the Norwegian national supercomputing infrastructure. In total, we used almost 500 000 processor hours.

PALM is a finite-difference numerical model. It solves filtered Navier–Stokes equations in the Boussinesq approximation on a Cartesian grid. We used a fifth-order advection scheme (Wicker and Skamarock 2002), and a third-order Runge–Kutta time step scheme (Williamson 1980). The 1.5-order flux-gradient subgrid closure solves a prognostic equation for the turbulence kinetic energy

FIG. 2. Observed speed and direction of local wind (arrows) and the difference between the AWS and MTP temperatures, ΔT (circles), for eight selected cases. Color shading shows ΔT magnitude. The temperature difference is defined in section 4a. The AWSs are numbered according to Table 1.

(Deardorff et al. 1980). The model grid of the Arakawa C type is regular with 1024×1024 grid points at each of 128 vertical levels. The model horizontal resolution is 27 m. The total domain size is $27\,648 \times 27\,648 \text{ m}^2$ (764.4 km^2) including 1000 m wide tapering zones at the eastern and northern lateral boundaries to match the domain edges. We used periodic lateral boundary conditions. The model vertical levels are spaced by 10 m at the bottom and stretching by 5% per level above 1000 m. The domain height is 1613 m. The ASTER DEM relief is approximated at the model grid by the bottom-mounted cuboids. Such a simplified approximation introduces certain errors but greatly accelerates the runs. More details on the PALM setup are given in the supplemental material (section S2).

We run PALM with fixed external forcing for each considered case. The temperature field in the model is initialized with corresponding idealized temperature profiles from the MTP observations. Wind is initialized with observations from the Ulriken AWS. A constant geostrophic wind forcing is applied in the entire domain. Table 3 provides information about the initial conditions and surface fluxes in the model. The sea surface in the model domain has a prescribed kinematic sensible heat flux of -0.03 K m s^{-1} , which corresponds to the

downward heat flux of about -40 W m^{-2} . The heat flux at each facet of the surface cubes is set to its own value as it approximates the solar insolation at the hill slope of different orientation with respect to the sun. The model spinup period is set to two hours. The model output over the simulated last 30 min is averaged and passed to the kriging methods as the LES data.

4. Methods and results of the temperature mapping

a. Initial data processing

Our task is to create a detailed temperature map over a complex terrain. Let $T(\mathbf{x})$ define the surface air temperature at the location given by the coordinates $\mathbf{x} = [x, y] \in \mathbf{X}$ where x and y are aligned with zonal (along the latitude circle) and meridional directions correspondingly. The ASTER DEM, which has been described in section 2a, gives the surface elevation, $z(\mathbf{x})$, above the sea level. We utilize the original ASTER grid $\mathbf{X}$ for the temperature mapping in this study, so that interpolation of the ASTER DEM data is not required. Seven specific locations, where the AWS observations are available, are identified as $\mathbf{x}_i$, $i = 1, \dots, 7$ (see Table 1).

TABLE 3. Initial conditions of the PALM runs. The simulations used the constant geostrophic wind speed in the entire domain. The vertical temperature gradient is set in three layers: 0–300 m, 300–1000 m, and above 1000 m. The surface sensible heat flux is prescribed at the top, east, west, south, and north facets of the surface cubic grid.

Case	Zonal wind (m s^{-1})	Meridional wind (m s^{-1})	Land surface temperature (K)	Temperature gradient [K (100 m)^{-1}]	Surface sensible heat flux (K m s^{-1})
NN	−1.5	−4.0	285.0	0.0, 0.25, 0.6	0.1, 0.1, −0.1, 0.1, −0.1
NE	−1.8	−3.7	287.0	0.0, 0.35, 0.6	0.1, 0.0, 0.1, 0.1, 0.0
EE	−12.7	1.9	290.0	0.15, 0.25, 0.6	0.1, 0.1, −0.01, 0.1, −0.01
SE	−2.4	−1.9	293.0	0.0, 0.97, 0.97	0.1, 0.0, 0.0, 0.1, −0.01
SS	−0.2	5.8	290.0	0.0, 0.25, 0.6	0.1, 0.01, 0.01, 0.1, −0.01
SW	3.0	2.4	288.0	0.0, 0.25, 0.6	0.1, 0.0, 0.0, 0.1, −0.1
WW	2.8	−0.3	285.0	0.0, 0.25, 0.6	0.1, 0.0, 0.0, 0.1, −0.01
NW	2.4	−1.6	285.0	0.0, 0.25, 0.6	0.1, 0.0, 0.05, 0.1, −0.01

The MTP temperature profiles in the free atmosphere are available at 20 equally spaced levels, z_n , $n = 1, \dots, 20$, between 45 (the top of the GFI) and 1045 m. The MTP instrument is installed at $\mathbf{x}_4$ (Florida), and has a field of view over the Bergen valley. This one-dimensional temperature profile is denoted as $T^V(z_n)$. A simple linear interpolation is used to obtain the free atmospheric temperature at any elevation z between z_n and z_{n+1} :

$$T^V = T^V(z) = T^V(z_n) + \frac{z - z_n}{z_{n+1} - z_n} [T^V(z_{n+1}) - T^V(z_n)]. \quad (1)$$

The temperature observed by an AWS is denoted as $T^{\text{AWS}}(\mathbf{x}_i)$. Hereafter, for brevity, we assume that the true elevation of any location $\mathbf{x}$ is $z(\mathbf{x})$. Since the vertical temperature change is the most significant factor of the spatial temperature variations, ad hoc correction for the ASTER DEM errors is needed. We implemented a heuristic correction as

$$T(\mathbf{x}_i) = T^{\text{AWS}}(\mathbf{x}_i) + T^V[z(\mathbf{x}_i)] - T^V(z_i), \quad (2)$$

where z_i is the true elevation of the location $\mathbf{x}_i$ from Table 1; $z(\mathbf{x}_i)$ is the elevation from the ASTER DEM.

We also introduce $T_0(\mathbf{x}) = T^V[z(\mathbf{x})]$ as the temperature on the two-dimensional surface $z(\mathbf{x})$, which has been set equal to $T^V[z(\mathbf{x})]$ at the height of the location $\mathbf{x}$. It is convenient to consider $T_0(\mathbf{x})$ —the elevation-corrected temperature map—as a result of a basic temperature mapping method. The other methods in the study will be compared with this method. Having the observed temperature profile T^V available, we can map the temperature more accurately than it is typically done with the mean vertical temperature gradient from climate data. Thus, it could be beneficial to interpolate only a correction to $T_0(\mathbf{x})$, which is given as

$$\Delta T(\mathbf{x}) = T(\mathbf{x}) - T_0(\mathbf{x}). \quad (3)$$

Such an approach would emphasize the impact of microphysical and dynamical factors.

The PALM model provides three-dimensional temperature fields $T^{3D,M}(\mathbf{x}^M, z^M)$ on a regular horizontal grid $\mathbf{x}^M \in \mathbf{X}^M$ and vertical levels z^M . We denote quantities related to the model data with a superscript M . In addition to the averaging over the last 30 min of simulations, we smooth $T^{3D,M}(\mathbf{x}^M, z^M)$ with a spatial running average using a square window of 150 m by 150 m to further suppress small-scale temperature fluctuations related to the approximation of the complex relief on our model grid. To obtain the two-dimensional surface temperature field $T^M(\mathbf{x}^M)$, we correct the model output $T^{3D,M}(\mathbf{x}^M, z)$ to eliminate a systematic model bias with respect to the observed temperature profile T^V at the MTP location $\mathbf{x}_4$ as

$$\Delta_T = \frac{1}{n} \sum_{i=1}^n [T^{3D,M}(\mathbf{x}_4, z_i) - T^V(z_i)],$$

$$T^M(\mathbf{x}^M) = T^{3D,M}[\mathbf{x}^M, z^{Ms}(\mathbf{x})] - \Delta_T. \quad (4)$$

Here, the surface, $z^{Ms}(\mathbf{x}) = z(\mathbf{x}) + \Delta z/2$, is the first model level over the ASTER DEM surface. The modeled temperature field $T^M(\mathbf{x})$ is obtained by a simple linear interpolation of $T^M(\mathbf{x}^M)$ from the model grid $\mathbf{X}^M$ onto the ASTER DEM grid $\mathbf{X}$. The local temperature deviations from $T^M(\mathbf{x})$ with respect to the simulated vertical temperature profile are given by

$$\Delta T^M(\mathbf{x}) = T^M(\mathbf{x}) - \{T^{3D,M}[\mathbf{x}_4, z^{Ms}(\mathbf{x})] - \Delta_T\}. \quad (5)$$

b. The kriging methods: Definitions and the main elements

“Kriging” is a family of probabilistic spatial prediction methods. The kriging methods use data-dependent covariance or variogram model rather than an a priori interpolating function (Chilès and Delfiner 2012). The kriging methods minimize a difference between the predicted (interpolated) value and its mathematical

TABLE 4. The methods of the temperature mapping.

Method	Data	Kriging	Variogram uses	LES data used for	Brief presentation
0	T^V	No	No	No	$T_0(\mathbf{x}) = T^V[z(\mathbf{x})]$
1	$T(\mathbf{x}_i)$	OK	$T^G(\mathbf{x})$	No	$T_{1OK}(\mathbf{x}) = OK\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^G(\mathbf{x})]\}$
2	$\Delta T(\mathbf{x}_i)$	OK	$\Delta T^G(\mathbf{x})$	No	$T_{2OK}(\mathbf{x}) = T_0(\mathbf{x}) + OK\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^G(\mathbf{x})]\}$
3	$T(\mathbf{x}_i)$	OKLES	$T^M(\mathbf{x})$	γ^e	$T_{1OKLES}(\mathbf{x}) = OK\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^M(\mathbf{x})]\}$
4	$\Delta T(\mathbf{x}_i)$	OKLES	$\Delta T^M(\mathbf{x})$	γ^e	$T_{2OKLES}(\mathbf{x}) = T_0(\mathbf{x}) + OK\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^M(\mathbf{x})]\}$
5	$T(\mathbf{x}_i)$	KED	$T^M(\mathbf{x})$	γ^e , drift	$T_{1KED}(\mathbf{x}) = KED\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^M(\mathbf{x})], T^M(\mathbf{x})\}$
6	$\Delta T(\mathbf{x}_i)$	KED	$\Delta T^M(\mathbf{x})$	γ^e , drift	$T_{2KED}(\mathbf{x}) = T_0(\mathbf{x}) + KED\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^M(\mathbf{x})], \Delta T^M(\mathbf{x})\}$
7	Weighted combination of methods 5 and 6				
	$T_{12KED}(\mathbf{x}) = [w_{1KED}(\mathbf{x})\Delta T_{1KED}(\mathbf{x}) + w_{2KED}(\mathbf{x})\Delta T_{2KED}(\mathbf{x})]/[w_{1KED}(\mathbf{x}) + w_{2KED}(\mathbf{x})]$				
8	The same as method 5 but with artificial points				
9	The same as method 6 but with artificial points				
10	The same as method 7 but with artificial points				

expectation. Simple kriging assumes that the mean value of the interpolated variable is known. Ordinary kriging (OK) assumes that the mean value is unknown. Universal kriging assumes that the mean value is unknown, but there are additional predictors to be used for interpolation (Bivand et al. 2013). A special case of universal kriging is given by KED. The KED method involves only one additional predictor.

We compare 10 OK and KED methods as applied to high-resolution temperature mapping. Table 4 summarizes short specifications of these 10 methods. Typically, the kriging methods are presented both in terms of covariance and variogram, but for brevity we provide here only the presentation in terms of variogram. Methods that interpolate the temperature field T itself are referred to with the subscript index 1; methods that interpolate ΔT are referred to with the subscript index 2. We introduce the methods for T , silently assuming, if nothing else is written, that ΔT can be used similarly.

Since all interpolation methods are realized with standard computing library packages, we outline only the formulation of the mathematical problem for the interpolation, and then we lump the description into a simplified notation. For further details, we refer the reader to the vast disciplinary literature and textbooks (e.g., Wackernagel 2003; Chilès and Delfiner 2012; Bivand et al. 2013). This study utilizes MATLAB software with open code geostatistical packages Gstat (Pebesma and Wesseling 1998; Bivand et al. 2013; Gräler et al. 2016; available at <http://github.com/edzer/gstat>) and mGstat (<http://mgstat.sourceforge.net/>). Our simplified notations consist of an identifier of the method with a list of arguments. The list may include identifiers of essential components of the method. The components may be realized and described as separate routines. For example, a notation for the OK method is written as

$$T_{OK}(\mathbf{x}) = OK\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T(\mathbf{x}_i)]\}. \quad (6)$$

Here, T is the observed temperature at the set of locations $\mathbf{x}_i$. The procedure interpolates the observations to a set of locations $\mathbf{x} \in \mathbf{X}$. The kriging procedure requires fitting of a theoretical variogram, which is done in a procedure designated as $\gamma[T(\mathbf{x}_i)]$. Its argument $T(\mathbf{x}_i)$ tells that the observed temperature at $\mathbf{x}_i$ is used to fit the variogram.

c. Variogram

Variogram measures the mean dissimilarity between $T(\mathbf{x}_i)$ and $T(\mathbf{x}_i + \rho)$ —two temperature readings separated by some distance ρ . Variogram is calculated as half the mean squared difference between the temperature pairs

$$\gamma_1^e(\rho_{ij}) = \frac{1}{2}[T(\mathbf{x}_i) - T(\mathbf{x}_j)]^2. \quad (7)$$

In this equation, the upper index e refers to the empirical (discrete) variogram; $\rho_{ij} = (\mathbf{x}_i, \mathbf{x}_j)$ is the horizontal distance between $\mathbf{x}_i$ and $\mathbf{x}_j$. The empirical variogram γ^e is defined for a small number of pairs. Therefore, it is useful to fit a theoretical variogram, $\gamma(\rho)$, $\rho \in \mathbb{R}$. We utilized three theoretical functions from Cressie (1991): exponential, $\gamma(\rho) = c[1 - \exp(-\rho/a)]$; power law, $\gamma(\rho) = c\rho^a$; Gaussian, $\gamma(\rho) = c[1 - \exp[-(\rho/a)^2]]$. The Gaussian function is used only as a part of weighted sum with the exponential function (with weights 0.9 and 0.1 correspondingly). The fitting coefficients, a and c , were obtained through minimization of weighted squares. All three functions are always fitted and the best fitted variogram is returned from $[T(\mathbf{x})]$, and used as the theoretical variogram for kriging.

By meteorological standards, the 7 urban AWS constitute a rather dense observational network for such a small area as we found it in Bergen. For geostatistical interpolation, however, this network is not satisfactory, and the theoretical variogram is poorly fitted (Fig. 3a). We cannot increase the number of locations where we have observations, but we still can improve variogram smoothness and fitting through an intermediate interpolation step. We need to eliminate lumpiness of the

FIG. 3. (a) Empirical variograms, γ^e , based on $\Delta T(\mathbf{x}_i)$ (circles) and $\Delta T^G(\mathbf{x})$ (gray curve), and the theoretical variogram, γ , based on $\Delta T^G(\mathbf{x})$ (dashed black curve); (b) the gridded field $\Delta T^G(\mathbf{x})$ within the convex polygon (triangle) given by the observational network of the AWSs; and (c) the map of the temperature deviation $\Delta T_{OK}(\mathbf{x})$, obtained with method 2—ordinary kriging—after Eq. (11). The interpolated data are shown with color shading. The observed values $\Delta T(\mathbf{x}_i)$ are shown by the colored circles. The data are shown for the NW case.

observational network. To do this, we consider two ideas. The first idea exploits a simpler interpolation method realized through grid data $[T(\mathbf{x}_i), \mathbf{x}_i, \mathbf{X}]$ —a built-in data gridding routine of the MATLAB software, that uses biharmonic spline interpolation (the “v4” option of the grid data function). It returns the temperatures $T^G(\mathbf{x})$ at a dense and regular grid after spline interpolation of the original sparse and lumped $T(\mathbf{x}_i)$. We retain only the data within the interpolation triangle (see Fig. 3b). This intermediate interpolation step is a pure engineering solution to smooth and to improve fitting on the empirical variogram (Fig. 3a). Still, the statistical spatial regularization of the observational network remains unsatisfactory in a sense that it does not produce a variogram that accounts for temperature variability in different physical and geographical conditions at larger distances from the observation sites.

Access to significant computational resources opens an opportunity to exploit the second idea. It proposes to construct the variogram from the high-resolution model results. The simulated temperature $T^M(\mathbf{x})$ is not only regular and dense but also includes additional local dynamical and physical information. It makes $T^M(\mathbf{x})$ preferable over $T^G(\mathbf{x})$ in the OK methods.

d. Presentation of 10 kriging methods

We introduce the kriging methods from the simplest to more sophisticated. Each description includes the

temperature map obtained by the presented method, so that the reader could easily consider effects of the subsequent sophistication of the methods.

1) METHODS 1 AND 2: OK

The OK methods minimize the following Lagrangian (the cost function)

$$-\sum_{i=1}^n \sum_{j=1}^n \omega_i(\mathbf{x}) \omega_j(\mathbf{x}) \gamma(\rho_{ij}) + 2 \sum_{i=1}^n \omega_i(\mathbf{x}) \gamma(\rho_{ix}) + \sigma_T^2(\mathbf{x}) + 2\mu(\mathbf{x}) \left[\sum_{i=1}^n \omega_i(\mathbf{x}) - 1 \right] = L(\mathbf{x}), \quad (8)$$

where $n = 7$ is the total number of the sampling locations, $\mathbf{x}_i$. The Lagrangian coefficients, $\mu(\mathbf{x})$, are unknown in the OK method and must be determined by a minimization procedure. The theoretical temperature variance $\sigma_T^2(\mathbf{x})$ is also unknown but depends only on the location and disappears in the minimization of Eq. (8). Solving the minimization problem in Eq. (8), one arrives to the following system of linear equations

$$\begin{cases} \sum_{j=1}^n \omega_j(\mathbf{x}) \gamma(\rho_{ij}) - \mu(\mathbf{x}) = \gamma(\rho_{ix}), \\ \sum_{i=1}^n \omega_i(\mathbf{x}) = 1. \end{cases} \quad (9)$$

FIG. 4. The temperature maps obtained by the OK methods: (a) T_{1OK} by method 1 and (b) T_{2OK} by method 2. (c) The temperature map is obtained by the linear interpolation of the MTP temperature profile T^V after Eq. (1). The data are shown for the NW case.

The obtained weights, $\omega_i(\mathbf{x})$, are based on the fitted theoretical variograms $\gamma(\rho_{ij})$. In turn, $\omega_i(\mathbf{x})$ are used to reconstruct the temperature map through the weighted interpolation of $T(\mathbf{x}_i)$ given at the location $\mathbf{x}_i$ to any given location $\mathbf{x} \in \mathbf{X}$ as

$$T_{1OK}(\mathbf{x}) = \sum_{i=1}^n \omega_i(\mathbf{x}) \Delta T(\mathbf{x}_i). \quad (10)$$

Equations (8)–(10) constitute method 1. The description of method 1 in the simplified notation read as

$$T_{1OK}(\mathbf{x}) = \text{OK}\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^G(\mathbf{x}_i)]\}. \quad (11)$$

Now, using the simplified notation in Eq. (11), method 2 reconstructs the temperature map using the temperature deviations $\Delta T(\mathbf{x}_i)$ as

$$T_{2OK}(\mathbf{x}) = T_0(\mathbf{x}) + \Delta T_{OK}, \quad (12)$$

$$\Delta T_{OK} = \text{OK}\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^G(\mathbf{x}_i)]\}.$$

Figure 3c shows ΔT_{OK} . As expected, this map is similar to the gridded ΔT^G map (Fig. 3b).

The direct implementation of the OK in method 1 after Eq. (11) provides unsatisfactory results. The temperature field $T_{1OK}(\mathbf{x})$ is unrealistic everywhere except a small area in central Bergen where most of the AWS are concentrated. Figure 4 reveals that kriging of the temperature deviations $\Delta T(\mathbf{x}_i)$ in method

2 improves the realism. The temperature map now resembles the map given by the vertical temperature profile T^V , but it captures significantly lower temperatures over the sea surface as represented by the observations at Nordnes. It also captures local wind cooling at the top of Ulriken and wind shading and direct solar warming of the southern slopes between Florida and Ulriken.

The temperature maps $T_{2OK}(\mathbf{x})$ still expose several physically unrealistic features arising from the lumpiness of the observation network. The northern elevated plateau (about 150 m MSL) is too cold, whereas the southern part of the Bergen valley is likely to be too warm. The map does not capture lower temperatures over the sea inlets in the central and southern Bergen. As there are no representative observations in those areas, these unphysical features could be corrected only through incorporation of additional information from a relief-resolving meteorological model. We emphasize that the modeling information cannot substitute the geostatistical temperature mapping because there is insufficient amount of high-resolution data to set up and run such a model. For instance, only insufficient or biased information is available for the local vertical wind profile, the local cloud distribution, and the local physical surface properties. The local surface fluxes must be obtained from the surface temperature map, which needs to be constructed. It would imply a costly iteration processes that are out of scope here.

FIG. 5. The temperature deviation maps obtained in the NW case for: (a) ΔT^M , (b) ΔT_{OKLES} , and (c) ΔT_{KED} .

2) METHODS 3 AND 4: ORDINARY KRIGING WITH LES VARIOGRAM (OKLES)

The OK methods 1 and 2 could be modified to include more adequate information about spatial temperature variability from the LES model runs. One modification is to exploit the simulated spatial temperature variability to fit the variogram. In this case, we substitute $\gamma[T^G(\mathbf{x}_i)]$ with $\gamma[T^M(\mathbf{x})]$ as

$$T_{1OKLES}(\mathbf{x}) = \text{OK}\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^M(\mathbf{x})]\}, \quad \text{and} \quad (13)$$

$$T_{2OKLES}(\mathbf{x}) = T_0(\mathbf{x}) + \Delta T_{2OKLES},$$

$$\Delta T_{2OKLES} = \text{OK}\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^M(\mathbf{x})]\}. \quad (14)$$

The temperature deviations ΔT^M and ΔT_{2OKLES} are shown in Figs. 5a and 5b. The ΔT_{2OKLES} is mostly similar to ΔT_{2OK} in Fig. 3c. The largest differences are found at larger distances from the AWSs (e.g., in the northeastern and southwestern corners) where the differences of the theoretical variograms become significant. Variogram plots for all cases can be found in supplemental material (section S3).

3) METHODS 5 AND 6: UNIVERSAL KRIGING WITH EXTERNAL DRIFT FROM THE LES (KED)

Universal kriging is more complicated than the OK. We consider the KED with the external drift from the LES results. The system to be solved in the KED methods reads

$$\begin{cases} \sum_{j=1}^n \omega_j(\mathbf{x}) \gamma(\rho_{ij}) - \mu_0(\mathbf{x}) - \mu_1(\mathbf{x}) T^M(\mathbf{x}) = \gamma(\rho_{ix}), & i = 1, \dots, n, \\ \sum_{i=1}^n \omega_i(\mathbf{x}) = 1, & \sum_{i=1}^n \omega_i(\mathbf{x}) T^M(\mathbf{x}) = T^M(\mathbf{x}). \end{cases} \quad (15)$$

The functions $\mu_0(\mathbf{x})$, $\mu_1(\mathbf{x})$ are to be determined solving this system of equations. The external drift is given by $T^M(\mathbf{x})$ as the last argument of the KED notation in

$$T_{1KED}(\mathbf{x}) = \text{KED}\{T(\mathbf{x}_i), \mathbf{X}, \gamma[T^M(\mathbf{x})], T^M(\mathbf{x})\}, \quad (16)$$

$$T_{2KED}(\mathbf{x}) = T_0(\mathbf{x}) + \Delta T_{KED}(\mathbf{x}), \quad \text{and}$$

$$\Delta T_{KED}(\mathbf{x}) = \text{KED}\{\Delta T(\mathbf{x}_i), \mathbf{X}, \gamma[\Delta T^M(\mathbf{x})], \Delta T^M(\mathbf{x})\}. \quad (17)$$

The temperature deviation obtained for the KED methods is shown in Fig. 5c. The temperature maps themselves are shown in Fig. 6. These maps are to be compared with the map for T_{2OKLES} , which is also

FIG. 6. The temperature maps in the NW case obtained for: (a) T_{2OKLES} , (b) T_{1KED} , and (c) T_{2KED} .

shown in Fig. 6a. By contrast to the OK methods, the KED methods demonstrate reduced differences between T_{1KED} and T_{2KED} .

4) METHOD 7: WEIGHTED COMBINATIONS OF THE TEMPERATURE MAPS

We observe that the interpolation methods behave differently in some parts of the domain. We know that their theoretical variogram models are different. So, the expected variances of spatial temperature prediction, which is also called kriging variance and is minimized in the interpolation algorithm, are also dissimilar in those methods. It is tempting to combine two or more kriging methods in a way to give the larger local weight to the method, which has the lower kriging variance at any specified location. The kriging variance, for example, $\sigma_{T_{1KED}}^2(\mathbf{x})$ in method 5, is given by

$$\sigma_{T_{1KED}}^2 = \sum_{i=1}^n \omega_i(\mathbf{x}) \gamma(\rho_{ix}) - \mu_0(\mathbf{x}) - \mu_1(\mathbf{x}) T^M(\mathbf{x}). \quad (18)$$

It is returned by the kriging routines. For $\sigma_{T_{2KED}}^2(\mathbf{x})$, $\Delta T^M(\mathbf{x})$ appears in Eq. (18) instead of $T^M(\mathbf{x})$. Similarly, we obtain the variance from the other kriging methods. The variance is zero at $\mathbf{x}_i$ where the observations are available; and grows with increasing ρ_{ix} . The growth rate depends on the adopted variogram model and spatial covariance of the observations. We have noted in the variogram presentation that the theoretical variogram for each method is chosen to achieve the best fitting.

In the NW case, $\gamma[T^M]$ for method 5, which returns T_{1KED} , is based on the Gaussian function, whereas $\gamma[\Delta T^M]$ for method 6, which returns T_{2KED} , is based on the power-law functions. Because of such differences, the growth rate on shorter distances, ρ_{ix} , is faster for $\sigma_{T_{2KED}}^2(\mathbf{x})$ than for $\sigma_{T_{1KED}}^2(\mathbf{x})$. Contrary, the growth rate of $\sigma_{T_{1KED}}^2(\mathbf{x})$ is faster at longer distances. These variograms and variances are shown in Figs. 7a–c.

The differences in the local kriging variance open an opportunity to create a weighted combination of the temperature maps with more geographically homogeneous quality. The weighted combination of the temperature maps is constructed with the local weights taken to be inversely proportional to $\sigma_{T_{1KED}}^2(\mathbf{x})$ and $\sigma_{T_{2KED}}^2(\mathbf{x})$ as

$$w_{1KED}(\mathbf{x}) = \frac{1}{\sigma_{T_{1KED}}^2(\mathbf{x})}, \quad w_{2KED}(\mathbf{x}) = \frac{1}{\sigma_{T_{2KED}}^2(\mathbf{x})}. \quad (19)$$

The weights $w_{1KED}(\mathbf{x})$ are shown in Fig. 7d. In method 7, the temperature map $T_{12KED}(\mathbf{x})$ is obtained as

$$T_{12KED}(\mathbf{x}) = \frac{w_{1KED}(\mathbf{x}) \Delta T_{1KED}(\mathbf{x}) + w_{2KED}(\mathbf{x}) \Delta T_{2KED}(\mathbf{x})}{w_{1KED}(\mathbf{x}) + w_{2KED}(\mathbf{x})}. \quad (20)$$

Although our evaluation later in this study will demonstrate that method 7 improves the mapping quality, a visual inspection of the temperature map for $T_{12KED}(\mathbf{x})$ shows only minor differences as compared with the two other KED methods.

FIG. 7. The kriging variances (a) $\sigma_{T1_KED}^2(\mathbf{x})$ and (b) $\sigma_{T2_KED}^2(\mathbf{x})$ in the NW case. (c) The empirical and theoretical variograms for method 5, which returns T_{1KED} , method 6, which returns T_{2KED} , are shown. (d) The map of the weights $w_{1KED}(\mathbf{x})$ in method 7, which returns T_{12KED} , is shown.

FIG. 8. The temperature deviations (a) ΔT^M , (b) ΔT_{KED} , and (c) ΔT_{KEDAP} in the NE case. Here, the Nordnes AWS (marked with the white cross) is withheld from the procedure. The KED method 6 reverse the sign of the temperature deviations in this setup. Artificial observations (marked with the black triangles) return the correct sign of the deviations in the KEDAP method 8.

5) METHODS 8, 9, AND 10: THE KED METHODS WITH ARTIFICIAL OBSERVATIONS (KEDAP)

In the process of implementing the KED methods 5, 6, and 7, we observed that the methods are sensitive to the data from remote AWSs, which show poor correlations with the rest of the observational network. Nordnes at $\mathbf{x}_7$ is such an AWS in our demonstration network. The Nordnes AWS is located at an extended peninsula where a large temperature contrast between the marine- and land-locked air is found. Due to this large marine influence, $T(\mathbf{x}_7)$ is poorly correlated with the other AWS in the network. Removal of the Nordnes AWS from the dataset inverts the sign of temperature deviations in the KED methods as shown in Fig. 8. Inversions are observed because the KED method utilizes external drift with an optimal offset and scaling factor. When used for extrapolation outside the area covered by observations, the scaling factor is unconstrained by observations, and might provide unrealistic values. The Nordnes AWS illustrates a more general, but frequent case of observed variance at some locations being much lower than actual variance in the study area, and the correlation between the sampled values and predictor being weak.

We propose to address this decorrelation problem through artificial observations. The artificial observations are taken from T^M at four arbitrary, but remote points, $\mathbf{x}_{ap}$. Those points are shown by black triangles at the domain corners in Fig. 8. Thus, the artificial

observations are taken from the LES data in under-sampled areas at large ρ_{ix} . Obviously, adding $T^M(\mathbf{x}_{ap})$ decreases weights of the real observations, and therefore this engineering should be applied with care. KEDAP could be written as

$$T_{1KEDAP}(\mathbf{x}) = \text{KED}\{T(\mathbf{x}_i) \cup T^M(\mathbf{x}_{ap}), \mathbf{X}, \gamma[T^M(\mathbf{x})], T^M\}. \quad (21)$$

It corresponds to the KED method 5. The KED method 6 becomes

$$\begin{aligned} T_{2KEDAP}(\mathbf{x}) &= T_0(\mathbf{x}) + \Delta T_{KEDAP}(\mathbf{x}), \\ \Delta T_{KEDAP}(\mathbf{x}) &= \text{KED}\{\Delta T(\mathbf{x}_i) \cup \Delta T^M(\mathbf{x}_{ap}), \\ &\quad \mathbf{X}, \gamma[\Delta T^M(\mathbf{x})], \Delta T^M\}. \end{aligned} \quad (22)$$

Similarly, the weighted method 7, which returns $T_{12KED}(\mathbf{x})$, could be modified using $T_{1KEDAP}(\mathbf{x})$ and $T_{2KEDAP}(\mathbf{x})$. Figure 8 illustrates the temperature deviations ΔT^M , ΔT_{KED} , and ΔT_{KEDAP} for the NE case. The NE case reveals the strongest influence of the AWS lumping in the central area on the mapping quality. This case comes out particularly poor in a cross validation without the Nordnes AWS.

Summing up, we remind the reader that the 10 kriging methods are implemented in the MATLAB programming environment and applied to the 8 weather cases as

FIG. 9. The reference scheme for the 10 kriging methods that presents the data flow. Here, GRD is the MATLAB grid data routine; γ is a routine that returns the fitted variogram model; OK is an ordinary kriging routine; KED is a routine for the universal kriging with external drift. A number of technical routines such as interpolation between different grids are not shown. The KEDAP methods are not presented in this scheme as they differ from the KED methods only by usage of the artificial “virtual” observations.

listed in Table 2. For convenience, we sketched the data flow in and between these methods in Fig. 9. The maps of ΔT for all cases are given in the supplemental material (section S4). The temperature maps themselves are not included in this presentation because, by eye, the choice of the best method is not obvious. The next section, however, presents an objective evaluation of the methods by a cross-validation method.

5. Cross validation of the temperature mapping methods

Although the choice of interpolation methods may be subjective in each specific case, it is generally advised to reconstruct the detailed temperature map with minimum temperature difference between the observed and interpolated values at each location. We evaluate the 10 kriging methods in this section with respect to this minimization criterion. We use quantitative cross validation (jackknifing) and evaluation as it is commonly applied in geostatistics (e.g., Benavides et al. 2007). The

method is suitable here because no reserved data are required for the validation. It is a highly desired feature when the observational network is small. The cross validation requires to remove one (or more) locations with observations from the network. The reduced network is used to obtain a new temperature map, which is compared with the original map based on the complete observational network. The point is to evaluate the temperature difference at a withheld location $\mathbf{x}_k$. Let $T_{jk}(\mathbf{x}_{i \neq k})$ denote the reconstructed temperature at $\mathbf{x}_k$ in the case j . Then, a residual error is

$$r_{jk}(\mathbf{x}_k) = T_{jk}(\mathbf{x}_{i \neq k}) - T(\mathbf{x}_k). \quad (23)$$

The residual errors for $T_0(\mathbf{x}_k)$ and $T^M(\mathbf{x}_k)$ are calculated as $r_0(\mathbf{x}_k) = T_0(\mathbf{x}_k) - T(\mathbf{x}_k)$ and $r^M(\mathbf{x}_k) = T^M(\mathbf{x}_k) - T(\mathbf{x}_k)$. Thus, there are $k = 1, \dots, n$; $n = 7$ values of $r_{jk}(\mathbf{x}_k)$ for each of $j = 1, \dots, N$; $N = 8$ cases. The average value for each case is obtained as $r_j = n^{-1} \sum_{k=1}^n r_{jk}(\mathbf{x}_k)$. The averaging over all cases is applied to get the unique value, which characterizes the overall performance of the

FIG. 10. The cross validation of the interpolation methods using (a) RMSE and (b) MXE. The gray vertical bars present RMSE and MXE; the colored bars—RMSE_N and MXE_N. The gray solid lines give the average RMSE and MXE over all 8 cases; the gray dotted lines—the average over 5 “good” cases. The black solid and dotted lines are RMSE_N and MXE_N.

method. The residual error is used to calculate a root-mean-square error as

$$\text{RMSE} = \sqrt{N^{-1} \sum_{j=1}^N r_j^2}, \quad (24)$$

and the absolute maximum error as

$$\text{MXE} = \max_{j=1, \dots, N} |r_j|. \quad (25)$$

Additionally, we calculate RMSE_N and MXE_N over 6 AWSs where we do not remove the Nordnes AWS at $\mathbf{x}_7$. We also calculate the average value over 5 “good” cases where the EE, SS, and SW cases (see Table 3) have been excluded. It has been recognized that in the EE, SS, and SW cases, the LES strongly underestimates magnitude of spatial variance of ΔT (see Fig. S4.1 in the online supplemental material). The coastal locations such as that of Nordnes were affected. The ratio of the observed to modeled standard deviation of ΔT was larger than 2.0 in the three “bad” cases, while this ratio in the five “good” cases varied from 0.97 to 1.77.

The quality of the kriging methods is improving with their sophistication and inclusion of additional information. Figure 10 and Table 5 show the cross-validation errors, averaged over considered cases. The supplemental material (section S5) includes values of the errors in all cases. Simple temperature interpolations are given by the baseline T_0 and T^M maps. The baseline maps utilize the vertical temperature gradient to account for the temperature change at different surface elevations, $z(\mathbf{x})$. In fact, T^M is more complicated as it is a simulated temperature, which includes local dynamics, turbulent mixing and physical properties of the surface. The RMSE averaged over all cases is 1.87 K for T_0 (1.77 K for the “good” cases). The LES simulations reduce this value to RMSE = 1.21 K (1.06 K) for T^M . This reduction confirms that the LES model has an added value for the high-resolution temperature mapping despite of its highly idealized setup.

As it reported in literature, the OK methods 1 and 2 are able to provide rather satisfactory temperature maps. We found, however, that T_{10K} is highly inaccurate almost everywhere, and that T_{20K} is still unrealistic

TABLE 5. The average quality of the temperature mapping in terms of RMSE and MXE in K. The values for each certain case are given in the supplemental material (section S5).

Selection of cases		RMSE		RMSE _N		MXE		MXE _N	
		All	Good	All	Good	All	Good	All	Good
Basic	T_o	1.87	1.77	1.28	1.77	3.75	3.73	2.00	1.75
	T^M	1.21	1.06	1.12	1.06	2.23	1.96	1.92	1.96
OK	T_{1OK}	1.34	1.36	1.23	1.36	2.30	2.34	2.24	2.28
	T_{2OK}	2.00	2.03	1.13	2.03	4.46	4.46	2.10	2.20
OK, variogram from LES	T_{1OKLES}	1.78	1.92	1.79	1.92	2.81	2.91	2.81	2.91
	T_{2OKLES}	1.83	1.87	0.92	1.87	4.25	4.27	1.60	1.74
KED	T_{1KED}	1.10	0.91	0.89	0.91	2.09	1.63	1.58	1.51
	T_{2KED}	1.45	1.68	0.90	1.68	3.18	3.72	1.67	1.68
	T_{12KED}	0.95	0.79	0.85	0.79	1.86	1.51	1.58	1.51
KED with additional sampled points	T_{1KEDAP}	1.10	0.91	0.89	0.91	2.09	1.63	1.58	1.51
	T_{2KEDAP}	0.99	0.90	0.83	0.90	1.96	1.76	1.51	1.59
	$T_{12KEDAP}$	0.94	0.80	0.80	0.80	1.78	1.44	1.44	1.44

in the remote areas and over the physically distinct surfaces. Surprisingly, T_{1OK} exhibits rather small averaged RMSE = 1.35 K and MXE = 2.3 K in average—those values are significantly smaller than RMSE and MXE of the baseline T_o . The errors are even smaller than for method 2, which returns T_{2OK} with RMSE = 2.0 K and MXE = 4.5 K. A closer look reveals that this deterioration of T_{2OK} is due to the Nordnes AWS. Thus, the Nordnes AWS—and, more generally, the coastal locations—is a challenge for the kriging methods.

We looked at the best theoretical variogram in the mapping with and without the Nordnes AWS. The variogram for T_{1OK} is rapidly increasing with distance ρ . It reaches the sill already at $\rho > 250$ m. When the distant observational locations are excluded, the method extrapolates the mean temperature of the remaining observational network as shown in Fig. 4a. It is just coincidence that the temperature at Nordnes is closer to the mean temperature of the remaining network than to one obtained through the kriging. If the Nordnes AWS is removed, T_{2OK} is obtained with positive temperature deviations ΔT_{OK} , whereas strongly negative values are expected at Nordnes. This error is further exacerbated by the positive error of T_o , resulting in the temperature bias of +4 K. If the Nordnes AWS is retained in the cross validation, then T_{2OK} errors are significantly reduced and remain below those for T_{1OK} . The average RMSE_N is 1.26 K for T_{1OK} and 1.13 K for T_{2OK} .

The LES further improves the temperature maps and makes them more realistic. Method 4 uses the LES data to fit the theoretical variogram. It has lower average RMSE = 1.83 K and RMSE_N = 0.92 K for T_{2OKLES} than method 2 for T_{2OK} . The Nordnes site, however, deteriorates the cross-validation results by 0.91 K, that is, by 100%. It increases the RMSE of T_{2OKLES} to be larger than that of the baseline T_o . We achieve noticeable

improvement only with the KED methods (RMSE = 1.1 K for T_{1KED} , and =1.45 for T_{2KED}). The larger RMSE for T_{2KED} is due to the sign reversal of ΔT in the NE case (see Fig. 8). Such a reversal results in a temperature overestimation at Nordnes by as much as 11 K. Without Nordnes, no significant differences in RMSE_N between T_{1KED} and T_{2KED} are found. The weighted combination of two KED methods—method 7 for T_{12KED} —further reduces the error at Nordnes in the NE case. The difference between RMSE and RMSE_N scores is practically eliminated in T_{12KED} . The average RMSE and MXE values for T_{12KED} are 0.95 and 1.86 K, respectively, which is smaller than they are T_o and even for T^M .

The artificial points in the remote parts of the domain eliminate the sign reversal and improve the temperature maps. Method 8 does not demonstrate advantages, but methods 9 and 10 for ΔT kriging do; the average RMSE and MXE values for T_{2KEDAP} are 0.99 and 1.96 K, respectively. The averaging only over “good” cases gives RMSE = 0.90 K and MXE = 1.76 K. The best results are achieved by the weighted combination of these two KEDAP methods in method 10. The average RMSE and MXE values of $T_{12KEDAP}$ are 0.99 and 1.78 K.

The two most advanced methods, 7 and 10, which return T_{12KED} and $T_{12KEDAP}$, respectively, are the best temperature mapping methods in terms of the cross validation averaged over all 8 cases. The cross-validation errors for these methods are about 50% of that for T_o and for the kriging methods (T_{2OK} , T_{1OKLES} , T_{2OKLES}). Specifically, the average RMSE for $T_{12KEDAP}$ is 2.24 times smaller than the RMSE for T_o and 2.41 times smaller than the RMSE for T_{2OK} . Considering only the “good” cases, the errors are 2.83 and 2.50 times smaller, respectively. The average MXE for $T_{12KEDAP}$ is 2.63 times smaller than MXE for T_o and 3.19 times

smaller than MXE for T_{2OK} . For the “good” cases, MXE is 3.17 and 3.84 times smaller. T_{12KED} and $T_{12KEDAP}$ remain physically realistic even in the case with the retained Nordnes AWS.

6. Conclusions and recommendations

The high-resolution temperature mapping over physically and topographically complex domain is required in many practical applications. The literature review reveals a broad diversity of different interpolation methods for temperature mapping. The air temperature strongly depends on elevation. Interpolation methods return unrealistic temperature maps if they do not include information on the relief. Other methods include such information through the vertical temperature gradient; either the climatic mean or actual gradients can be used. They return more realistic temperature maps, as we demonstrate with the temperature mapping in T_0 . Still, it is highly desirable to account not only for the vertical temperature gradient but also for the local temperature differences induced by variations in land use–land cover, physical surface properties, and local atmospheric dynamics. Horizontal temperature differences are typically comparable with the vertical differences even within a mountainous area.

This study explores the 10 kriging methods addressing the needs of detailed local temperature mapping. The key novel element of this study is in incorporating of the LES results. We run the LES model PALM. Eight different weather cases are simulated. The cases are mainly distinguished by the wind direction, and therefore by the advection of cold and warm air in the coastal zone and within the valleys. The initial and boundary conditions for the high-resolution model simulations are taken from observations, but the model do not assimilate the observations inside the domain. Hence, the LES does not simulate the observed weather conditions, and it cannot be directly used for temperature mapping. At the same time, the LES resolves local circulations, turbulent mixing and the surface heterogeneity. Thus, the model output can provide additional information to obtain variogram and external drift in the kriging methods.

This study creates the temperature maps of the Bergen area using a dense network of 7 AWSs and the vertical temperature profile from the MTP instrument. The network is lumped in the central Bergen area. Moreover, the AWSs are located at different elevations. Thus, it is challenging to extract homogeneous statistical information for interpolation routines from such a network. Offshore and coastal areas, elevated plateaus, and other, smaller, valleys are not covered with observations.

We demonstrate that simple ordinary kriging provides a highly unrealistic temperature map T_{1OK} . We propose to use temperature deviations, $\Delta T = T - T_0$, in order to improve the map. The resulting T_{2OK} is significantly more realistic. The temperature maps are further improved in more sophisticated kriging methods, which use the LES results in the calculations of variogram and external drift. The best results are returned by the proposed weighted KED method 7. This method intensively utilizes the LES results for the variogram and the external drift terms. We observe that the quality of the temperature maps is geographically uneven. The temperature field in some area is better reconstructed by one method, whereas another method is better off in some other area. We propose to weight the different temperature maps to further improve the reconstruction in the whole domain. The maps are weighted by the inverse variance of temperature (or temperature difference) at each location. Method 7 in such a procedure returns T_{12KED} that is of significantly better quality, in terms of cross validation, than the temperature maps from the other methods.

Considering the NE case, we observed that the largest contribution in the errors is coming from the remote AWS in Nordnes at the interface between warmer land and colder sea areas. As it is poorly correlated to the other stations, ΔT may reverse its sign at this location, thus adding instead of cancelling, to the differences. T_{12KED} is less affected by this problem than that from purely ΔT -based methods 2, 4, and 6. We propose to engineer a correction involving artificial points with observations from the LES. Those points are located in remote parts of the domain, so that they decrease the weight of the Nordnes AWS in calculations. This approach improves $T_{12KEDAP}$ quality making it the best method of this study.

We conclude that even the idealized LES runs can improve the high-resolution temperature mapping in the complex areas, areas with a sparse and lumped observational network as well as in the case when the observations include inaccurate and unrepresentative data. We recommend the weighted methods 7 and 10, which return T_{12KED} and $T_{12KEDAP}$, for practical applications.

Acknowledgments. This study was supported by the ReSiS project of the Norwegian Ministry of Environment, by the Belmont Forum project Anthropogenic Heat Islands in the Arctic: Windows to the Future of the Regional Climates, Ecosystems, and Societies (Project 247468) and by the Russian Foundation for Basic Research (RFBR) and the Moscow Government Project 19-35-70009. Kriging methods efficiency evaluation for urban landscapes was performed with financial

support of the Russian Science Foundation Project 19-77-30012. We are grateful to Dr. Laurent Bertino at the Nansen Environmental and Remote Sensing Center and Dr. Timofey Samsonov at the Lomonosov Moscow State University for their advice and discussions of the methods and results.

REFERENCES

- Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, and E. F. Wood, 2018: Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data*, **5**, 180214, <https://doi.org/10.1038/sdata.2018.214>.
- Benavides, R., F. Montes, A. Rubio, and K. Osoro, 2007: Geostatistical modelling of air temperature in a mountainous region of northern Spain. *Agric. For. Meteorol.*, **146**, 173–188, <https://doi.org/10.1016/j.agrformet.2007.05.014>.
- Bivand, R. S., E. Pebesma, and V. Gómez-Rubio, 2013: *Applied Spatial Data Analysis with R*. Springer, 414 pp., <https://doi.org/10.1007/978-1-4614-7618-4>.
- Bois, B., D. Joly, H. Quénot, P. Pieri, J.-P. Gaudillère, D. Guyon, E. Saur, and C. van Leeuwen, 2018: Temperature-based zoning of the Bordeaux wine region. *OENO One*, **52**, 291–306, <https://doi.org/10.20870/oeno-one.2018.52.4.1580>.
- Bosveld, F. C., and Coauthors, 2014: The third GABLS inter-comparison case for evaluation studies of boundary-layer models. Part B: Results and process understanding. *Bound.-Layer Meteorol.*, **152**, 157–187, <https://doi.org/10.1007/s10546-014-9919-1>.
- Buscaill, C., E. Upegui, and J.-F. Viel, 2012: Mapping heatwave health risk at the community level for public health action. *Int. J. Health Geogr.*, **11**, 38, <https://doi.org/10.1186/1476-072X-11-38>.
- Chilès, J.-P., and P. Delfiner, 2012: *Geostatistics: Modeling Spatial Uncertainty*. 2nd ed. John Wiley & Sons, 699 pp., <https://doi.org/10.1002/9781118136188>.
- Courault, D., and P. Monestiez, 1999: Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France. *Int. J. Climatol.*, **19**, 365–378, [https://doi.org/10.1002/\(SICI\)1097-0088\(19990330\)19:4<365::AID-JOC369>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0088(19990330)19:4<365::AID-JOC369>3.0.CO;2-E).
- Cressie, N., 1991: *Statistics for Spatial Data*. John Wiley & Sons, 900 pp.
- Davy, R., and I. Esau, 2016: Differences in the efficacy of climate forcings explained by variations in atmospheric boundary layer depth. *Nat. Commun.*, **7**, 11690, <https://doi.org/10.1038/ncomms11690>.
- , —, S. Outten, A. Chernokulsky, and S. Zilitinkevich, 2017: Diurnal asymmetry to the observed global warming. *Int. J. Climatol.*, **37**, 79–93, <https://doi.org/10.1002/JOC.4688>.
- Deardorff, J. W., G. E. Willis, and B. H. Stockton, 1980: Laboratory studies of the entrainment zone of a convectively mixed layer. *J. Fluid Mech.*, **100**, 41–64, <https://doi.org/10.1017/S0022112080001000>.
- DeGaetano, A. T., and B. N. Belcher, 2007: Spatial interpolation of daily maximum and minimum air temperature based on meteorological model analyses and independent observations. *J. Appl. Meteor. Climatol.*, **46**, 1981–1992, <https://doi.org/10.1175/2007JAMC1536.1>.
- Dobrowski, S. Z., 2011: A climatic basis for microrefugia: The influence of terrain on climate. *Global Change Biol.*, **17**, 1022–1035, <https://doi.org/10.1111/j.1365-2486.2010.02263.x>.
- Esau, I., and V. Miles, 2018: Exogenous drivers of surface urban heat islands in northern west Siberia. *Geogr. Environ. Sustainability*, **11**, 83–99, <https://doi.org/10.24057/2071-9388-2018-11-3-83-99>.
- , T. Wolf, E. Miller, I. Repina, Y. Troitskaya, and S. Zilitinkevich, 2013: The analysis of results of remote sensing monitoring of the temperature profile in lower atmosphere in Bergen (Norway). *Russ. Meteor. Hydrol.*, **38**, 715–722, <https://doi.org/10.3103/S1068373913100099>.
- Fernández, F., L. Alvarez-Vazquez, N. Garcia-Chan, A. Martinez, and M. Vazquez-Mendez, 2015: Optimal location of green zones in metropolitan areas to control the urban heat island. *J. Comput. Appl. Math.*, **289**, 412–425, <https://doi.org/10.1016/j.cam.2014.10.023>.
- Gräler, B., E. Pebesma, and G. Heuvelink, 2016: Spatio-temporal interpolation using gstat. *R J.*, **8**, 204–218, <https://doi.org/10.32614/RJ-2016-014>.
- Gregg, J. W., C. G. Jones, and T. E. Dawson, 2003: Urbanization effects on tree growth in the vicinity of New York City. *Nature*, **424**, 183–187, <https://doi.org/10.1038/nature01728>.
- Gronemeier, T., S. Raasch, and E. Ng, 2017: Effects of unstable stratification on ventilation in Hong Kong. *Atmosphere*, **8**, 168, <https://doi.org/10.3390/atmos8090168>.
- Hengl, T., B. M. Gerard, M. P. Tadić, and E. J. Pebesma, 2011: Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theor. Appl. Climatol.*, **107**, 265–277, <https://doi.org/10.1007/s00704-011-0464-2>.
- Hjort, J., J. Suomi, and J. Käyhkö, 2016: Extreme urban-rural temperatures in the coastal city of Turku, Finland: Quantification and visualization based on a generalized additive model. *Sci. Total Environ.*, **569–570**, 507–517, <https://doi.org/10.1016/j.scitotenv.2016.06.136>.
- Ho, H. C., A. Knudby, Y. Xu, M. Hodul, and M. Aminipouri, 2016: A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (Humidex), for the greater Vancouver area. *Sci. Total Environ.*, **544**, 929–938, <https://doi.org/10.1016/j.scitotenv.2015.12.021>.
- Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei, 2008: Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.*, **113**, D21110, <https://doi.org/10.1029/2008JD010100>.
- Hudson, G., and H. Wackernagel, 1994: Mapping temperature using kriging with external drift: theory and an example from Scotland. *Int. J. Climatol.*, **14**, 77–91, <https://doi.org/10.1002/joc.3370140107>.
- Ishida, T., and S. Kawashima, 1993: Use of cokriging to estimate surface air temperature from elevation. *Theor. Appl. Climatol.*, **47**, 147–157, <https://doi.org/10.1007/BF00867447>.
- Jonassen, M. O., H. Ólafsson, J. Reuder, and J. A. Olseth, 2012: Multi-scale variability of winds in the complex topography of southwestern Norway. *Tellus*, **64A**, 11962, <https://doi.org/10.3402/tellusa.v64i0.11962>.
- Kadyrov, E. N., and D. R. Pick, 1998: The potential for temperature retrieval from an angular-scanning single-channel microwave radiometer and some comparisons with in situ observations. *Meteor. Appl.*, **5**, 393–404, <https://doi.org/10.1017/S1350482798001054>.
- Lapen, D. R., and H. N. Hayhoe, 2003: Spatial analysis of seasonal and annual temperature and precipitation normals in southern Ontario, Canada. *J. Great Lakes Res.*, **29**, 529–544, [https://doi.org/10.1016/S0380-1330\(03\)70457-2](https://doi.org/10.1016/S0380-1330(03)70457-2).
- Letzel, M. O., M. Krane, and S. Raasch, 2008: High resolution urban large-eddy simulation studies from street canyon to

- neighbourhood scale. *Atmos. Environ.*, **42**, 8770–8784, <https://doi.org/10.1016/j.atmosenv.2008.08.001>.
- Li, J., and A. D. Heap, 2011: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecol. Inform.*, **6**, 228–241, <https://doi.org/10.1016/j.ecoinf.2010.12.003>.
- , and —, 2014: Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Modell. Software*, **53**, 173–189, <https://doi.org/10.1016/j.envsoft.2013.12.008>.
- Maronga, B., and Coauthors, 2015: The Parallelized Large-Eddy Simulation Model (PALM) version 4.0 for atmospheric and oceanic flows: Model formulation, recent developments, and future perspectives. *Geosci. Model Dev.*, **8**, 2515–2551, <https://doi.org/10.5194/gmd-8-2515-2015>.
- Meier, F., D. Fenner, T. Grassmann, M. Otto, and D. Scherer, 2017: Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Climate*, **19**, 170–191, <https://doi.org/10.1016/j.uclim.2017.01.006>.
- Miles, V., and I. Esau, 2016: Spatial heterogeneity of greening and browning between and within bioclimatic zones in northern west Siberia. *Environ. Res. Lett.*, **11**, 115002, <https://doi.org/10.1088/1748-9326/11/11/115002>.
- Mironova, V., N. Shartova, A. Beljaev, M. Varentsov, and M. Grishchenko, 2019: Effects of climate change and heterogeneity of local climates on the development of malaria parasite (*Plasmodium vivax*) in Moscow megacity region. *Int. J. Environ. Res. Public Health*, **16**, 694, <https://doi.org/10.3390/ijerph16050694>.
- Muñoz-Esparza, D., J. K. Lundquist, J. A. Sauer, B. Kosović, and R. R. Linn, 2017: Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign: From weather to boundary-layer eddies. *J. Adv. Model. Earth Syst.*, **9**, 1572–1594, <https://doi.org/10.1002/2017MS000960>.
- Pebesma, E. J., and C. G. Wesseling, 1998: A program for geostatistical modelling, prediction and simulation. *Comput. Geosci.*, **24**, 17–31, [https://doi.org/10.1016/S0098-3004\(97\)00082-4](https://doi.org/10.1016/S0098-3004(97)00082-4).
- Resler, J., and Coauthors, 2017: PALM-USM v1.0: A new urban surface model integrated into the PALM large-eddy simulation model. *Geosci. Model Dev.*, **10**, 3635–3659, <https://doi.org/10.5194/gmd-10-3635-2017>.
- Salvetti, M. V., B. Geurts, J. Meyers, and P. Sagaut, Eds., 2011: *Quality and Reliability of Large-Eddy Simulations II*. Springer, 430 pp., <https://doi.org/10.1007/978-94-007-0231-8>.
- Scambos, T. A., G. G. Campbell, A. Pope, T. Haran, A. Muto, M. Lazzara, C. H. Reijmer, and M. R. van den Broeke, 2018: Ultralow surface temperatures in East Antarctica from satellite thermal infrared mapping: The coldest places on Earth. *Geophys. Res. Lett.*, **45**, 6124–6133, <https://doi.org/10.1029/2018GL078133>.
- Schalkwijk, J., H. J. J. Jonker, A. P. Siebesma, and F. C. Bosveld, 2015: A year-long large-eddy simulation of the weather over Cabauw: An overview. *Mon. Wea. Rev.*, **143**, 828–844, <https://doi.org/10.1175/MWR-D-14-00293.1>.
- Shao, J., J. Swanson, R. Patterson, P. Lister, and A. McDonald, 1997: Variation of winter road surface temperature due to topography and application of thermal mapping. *Meteor. Appl.*, **4**, 131–137, <https://doi.org/10.1017/S135048279700042X>.
- Smoliak, B. V., P. K. Snyder, T. E. Twine, P. M. Mykleby, and W. F. Hertel, 2015: Dense network observations of the Twin Cities canopy-layer urban heat island. *J. Appl. Meteor. Climatol.*, **54**, 1899–1917, <https://doi.org/10.1175/JAMC-D-14-0239.1>.
- Stahl, K., R. D. Moore, J. A. Floyer, M. G. Asplin, and I. G. McKendry, 2006: Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agric. For. Meteorol.*, **139**, 224–236, <https://doi.org/10.1016/j.agrformet.2006.07.004>.
- Szymanowski, M., and M. Kryza, 2009: GIS-based techniques for urban heat island spatialization. *Climate Res.*, **38**, 171–187, <https://doi.org/10.3354/cr00780>.
- Tachikawa, T., and Coauthors, 2011b: ASTER Global Digital Elevation Model version 2—Summary of validation results. Japan Space Systems, 27 pp., https://ssl.jspacesystems.or.jp/ersdac/GDEM/ver2Validation/Summary_GDEM2_validation_report_final.pdf.
- , M. Hato, M. Kaku, and A. Iwasaki, 2011a: The characteristics of ASTER GDEM version 2. *2011 IEEE Int. Geoscience and Remote Sensing Symp.*, Vancouver, BC, Canada, IEEE, 3657–3660, <https://doi.org/10.1109/IGARSS.2011.6050017>.
- Taheri-Shahraiyi, H., and S. Sodoudi, 2017: High-resolution air temperature mapping in urban areas: A review on different modelling techniques. *Therm. Sci.*, **21**, 2267–2286, <https://doi.org/10.2298/TSCI150922094T>.
- Valved, A. S., 2012: Local flow conditions in the Bergen valley based on observations and numerical simulations. MS thesis. Geophysical Institute, University of Bergen, 30 pp, <http://bora.uib.no/handle/1956/7016>.
- Varentsov, M., P. Konstantinov, A. Baklanov, I. Esau, V. Miles, and R. Davy, 2018: Anthropogenic and natural drivers of a strong winter urban heat island in a typical Arctic city. *Atmos. Chem. Phys.*, **18**, 17 573–17 587, <https://doi.org/10.5194/acp-18-17573-2018>.
- Wackernagel, H., 2003: *Multivariate Geostatistics*. Springer, 387 pp., <https://doi.org/10.1007/978-3-662-05294-5>.
- , C. Lajaunie, N. Blond, C. Roth, and R. Vautard, 2004: Geostatistical risk mapping with chemical transport model output and ozone station data. *Ecol. Modell.*, **179**, 177–185, <https://doi.org/10.1016/j.ecolmodel.2004.06.018>.
- Wicker, L. J., and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097, [https://doi.org/10.1175/1520-0493\(2002\)130<2088:TSMFEM>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2).
- Williamson, J. H., 1980: Low-storage Runge-Kutta schemes. *J. Comput. Phys.*, **35**, 48–56, [https://doi.org/10.1016/0021-9991\(80\)90033-9](https://doi.org/10.1016/0021-9991(80)90033-9).
- Wolf, T., I. Esau, and J. Reuder, 2014: Analysis of the vertical temperature structure in the Bergen valley, Norway, and its connection to pollution episodes. *J. Geophys. Res. Atmos.*, **119**, 10 645–10 662, <https://doi.org/10.1002/2014JD022085>.
- , L. Pettersson, and I. Esau, 2020: A very high-resolution assessment and modelling of urban air quality. *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/ACP-2019-294>, in press.
- Wolf-Grosse, T., I. Esau, and J. Reuder, 2017: Sensitivity of local air quality to the interplay between small- and large-scale circulations: A large-eddy simulation study. *Atmos. Chem. Phys.*, **17**, 7261–7276, <https://doi.org/10.5194/acp-17-7261-2017>.
- Yao, X., B. Fu, Y. Lü, F. Sun, S. Wang, and M. Liu, 2013: Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. *PLOS ONE*, **8**, e54660, <https://doi.org/10.1371/journal.pone.0054660>.
- Zhang, K., E. M. Oswald, D. G. Brown, S. J. Brines, C. J. Gronlund, J. L. White-Newsome, R. B. Rood, and M. S. O'Neill, 2011: Geostatistical exploration of spatial variation of summertime temperatures in the Detroit metropolitan region. *Environ. Res.*, **111**, 1046–1053, <https://doi.org/10.1016/j.envres.2011.08.012>.
- Zhang, Y., A. T. Murray, and B. Turner, 2017: Optimizing green space locations to reduce daytime and nighttime urban heat island effects in Phoenix, Arizona. *Landscape Urban Plann.*, **165**, 162–171, <https://doi.org/10.1016/j.landurbplan.2017.04.009>.