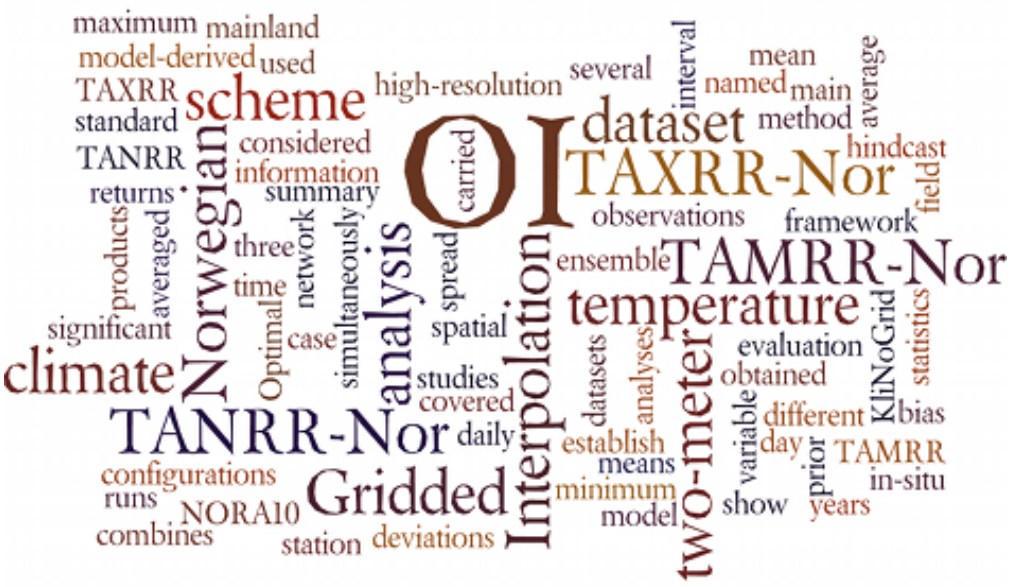
Norwegian Meteorological Institute

METreport

No. 02/2017
ISSN 2387-4201
Free

Spatial Interpolation of daily minimum, maximum and mean temperature

KliNoGrid: TANRR-Nor, TAXRR-Nor and TAMRR-Nor
Cristian Lussana



Norwegian Meteorological Institute

METreport

Title	Date
Spatial Interpolation of daily minimum, maximum and mean temperature	February 10, 2017
Section	**Report no.**
Division for Climate Services	02/2017
Author(s)	**Classification**
Cristian Lussana	● Free ○ Restricted
Client(s)	**Client's reference**
The Norwegian Water Resources and Energy Directorate	Tuomo Saloranta

Abstract

A spatial interpolation method based on Optimal Interpolation (OI) has been developed for the daily (i.e. 06-06 UTC): minimum (TANRR), maximum (TAXRR) and mean (TAMRR) two-meter temperature on Norwegian mainland. The OI combines a model-derived background field with in-situ observations from the climate station network. The model considered is the NORA10 high-resolution hindcast dataset. For each day and variable, the OI scheme runs simultaneously with several different configurations, such that an ensemble of analysis is obtained. The two main products of the interpolation are the analysis ensemble mean and spread. The OI scheme has been used to establish three gridded datasets within the Norwegian gridded climate dataset (KliNoGrid): TANRR-Nor, TAXRR-Nor and TAMRR-Nor. The time interval covers 33 years: from 1980 to 2012. The evaluation has been carried out by means of summary statistics and case studies. In general, TANRR-Nor, TAXRR-Nor and TAMRR-Nor are unbiased estimates of the actual temperature and their precision is on average between 1°C and 2°C.

Keywords

temperature, spatial interpolation, optimal interpolation, climate, hydrology

Disciplinary signature	Responsible signature
Hans Olav Hygen	Cecilie Stenersen

Contents

1 Introduction

The Norwegian gridded climate dataset (KliNoGrid) is an ongoing project based on the combination of remote sensing and/or numerical model output fields with in-situ observations. KliNoGrid aims at supporting hydrology, meteorology, climate research and operational applications. The combination of information relies on statistical methods to post-process model, or remote sensing, fields in order to to improve the accuracy and precision of the final predictions by including the small-scale processes measured by the point observations. Because of our interest in climate and hydrology, the gridded datasets should extend as far back in time as possible. In addition, the predicted fields are available on a regular grid with 1 Km of spacing in both Northing and Easting directions. The grid covers the relevant catchments for the Norwegian hydrology. The KliNoGrid dataset might be seen as complementary to the conventional climatological datasets, in the sense discussed in *Simmons et al.* (2016), such as the seNorge datasets (*Tveito et al.*, 2000; *Lussana et al.*, 2016b) which are based on in situ (i.e. point) observations only.

Currently, the KliNoGrid dataset includes daily and hourly precipitation fields based on the combination of radar-derived precipitation estimates with rain gauge observations (*Lussana et al.*, 2016a). Furthermore, we are working on the production of high-resolution gridded datasets of hourly wind covering a time interval of about 60 years.

In this report, the KliNoGrid products for two-meter daily minimum, maximum and mean temperature are introduced. For consistency with the aggregation time currently used for precipitation and temperature in hydrology, the daily aggregation is defined between 06 UTC of the previous day and 06 UTC of the day considered. As a consequence, given the standard for the parameter labels in the climate database of the Norwegian Meteorological Institute (MET Norway) the labels of the daily variables are: TANRR for the minimum temperature; TAXRR for the maximum temperature and TAMRR for the mean temperature. Note that TANRR and TAXRR have been defined specifically within this work, for they were not present in the climate database.

We have decided to combine the point observations for TANRR, TAXRR and TAMRR with the information derived from the NORA10 high-resolution hindcast dataset available at MET-Norway, which goes back in time to 1957. In particular, the two-meter temperature fields have been downscaled from their original grid spacing (around 10 Km) to the 1 Km grid and aggregated at the daily timescale as required by the three variables of interest. The final daily gridded datasets are named: TANRR-Nor for the minimum

temperature; TAXRR-Nor for the maximum temperature and TAMRR-Nor for the mean temperature, such that both the observation labels and the NORA10 model background are reported in the names.

The spatial interpolation is based on Optimal Interpolation (OI: *Gandin and Hardin*, 1965; *Daley*, 1991), which is a well-established statistical interpolation method in atmospheric sciences (*Lorenc*, 1986), where it is used in data assimilation to provide the initial condition for numerical simulations. In our case, the OI scheme has been adapted to adjust NORA10 model output fields given actual atmospheric in-situ observations, which are assumed to be a more accurate and precise representation of the true atmospheric state. The NORA10 two-meter temperature fields provide the background fields (prior), which are combined with the TANRR, TAXRR and TAMRR observations. Because of the mismatch between the full range of atmospheric scales influencing the point observations and the areal averages simulated by the numerical models, the issue of observation representativeness must be taken into account. A discussion of the OI observation representativeness error can be found in *Lussana et al.* (2010), here we simply mention that even in the fortuitous case that an observation and a grid point share the same location one should not expect an exact correspondence between the analyses and the observed values. Rather, the spatial interpolation scheme should be able to filter out the effects of small-scale processes influencing the point observations but which are not properly resolved by the numerical model.

The OI is based on the assumption of Gaussian probability density functions for the observation and background errors (i.e. deviations from the unknown truth), which is an approximation that we have considered valid in our work, nonetheless it might be worth mentioning that for daily extremes, such as TANRR-Nor and TAXRR-Nor, this is a less satisfactory approximation than for TAMRR-Nor. As a consequence, we might be justified in expecting a better quality (i.e. less uncertainty) for TAMRR-Nor fields than for TAXRR-Nor and TANRR-Nor.

The OI scheme includes a spatial consistency test, as described in *Lussana et al.* (2010, 2016b).

The original parts of our OI scheme are the parameter estimation procedure and the idea to consider an ensemble of analysis instead of a single best (i.e. minimum analysis error variance), linear, unbiased analysis. Our parameter estimation procedure selects those OI configurations which are expected to provide the best analysis fields, as measured by a likelihood function based on leave-one-out cross-validation. In general, the selection

includes around 25 possible OI configurations, which we have considered enough to sample the distribution of OI parameter values or at least to give us an idea of their optimal values. Consequently, several OI are performed for the same case, with the different configurations. As a result, an ensemble of analysis fields is obtained, which is intended to represent the sensitivity of the OI scheme to the selection of parameter values. For each day and variable, the main final products of our OI scheme are the analysis ensemble mean and spread, which is conveyed by the the standard deviation of the analysis ensemble.

The OI scheme has been used to establish the gridded datasets of TANRR-Nor, TAXRR-Nor and TAMRR-Nor, which covers the time period from 1980 to 2012.

The report is organized as follow. The observations and the NORA10 characteristics are reported in section 2. The spatial interpolation scheme is described in detail in section 3.2 and its evaluation over a recent period of more than 30 years is presented in section 4, together with two case studies.

2 Data

2.1 Observations

The Climate Database of the Norwegian Meteorological Institute (i.e. KDVH, *Klima Data Vare Huset*) has been used as our source of observations. The daily total precipitation for the generic day D is a key parameter for hydrology and it is defined as the accumulated precipitation between 06 UTC of day $D-1$ to 06 UTC of day D (i.e. parameter RR in KDVH). Our choice has been to define the minimum, maximum and average two-meter air temperature for the generic day D in a consistent way to RR. As a consequence, we have:

- TANRR: daily minimum temperature (06-06 UTC);
- TAXRR: daily maximum temperature (06-06 UTC);
- TAMRR: daily mean temperature (06-06 UTC), arithmetic mean of 24 hourly values or a formula based mean value computed from fewer observations;

TAMRR is available in KDVH. Unfortunately, TANRR and TAXRR are not directly available, they must be obtained by post-processing the available temperature observations. In particular, the two parameters we have used are:

- TAN_12: lowest observed temperature in the last 12 hours;
- TAX_12: highest observed temperature in the last 12 hours;

the TAN_12 and TAX_12 observations are available twice a day: at 06 UTC and 18 UTC. TANRR and TAXRR for day D are defined as:

$$\text{TANRR} = min\left(\text{TAN\_12}_{\text{day}=(D-1),\text{time}=18UTC}, \text{TAN\_12}_{\text{day}=D,\text{time}=06UTC}\right) \quad (1)$$

$$\text{TAXRR} = max\left(\text{TAX\_12}_{\text{day}=(D-1),\text{time}=18UTC}, \text{TAX\_12}_{\text{day}=D,\text{time}=06UTC}\right) \quad (2)$$

In Figure 1 the time series of available observations for TANRR, TAXRR and TAMRR in the time period from 1980 to 2012 are shown. At the beginning of the time period, the number of daily observations is around 200 for each variable, then they gradually decrease to: around 100 for both TAXRR and TANRR; around 150 for TAMRR. From 2007 onward, a sharp increase in the data availability is clearly evident and in 2012 the maximum values of around 270 daily observations for each variable are reached.

The Figure 2 gives a broad overview of the (spatially averaged) temperature trends in the period under study as the 365-day centered moving anomaly respect to the mean TAMRR value. The Figure might be useful in the interpretation of our results.

2.2 Model data

The NORA10 high-resolution hindcast dataset is described in *Reistad et al.* (2011). The dataset has been obtained as a dynamical downscaling based on ERA40 from 1957 to 2002 and on ECMWF operational analyses from 2002 onwards. The NORA10 two-meter temperature fields are available on a regular grid covering part of Northern Europe and having around 10 Km of grid spacing in both Easting and Northing directions. The fields are available at hourly time resolution, then the definitions for TANRR, TAXRR and TAMRR reported in section 2.1 have been used to aggregated the NORA10 hourly fields into the corresponding daily values.

3 Methods

3.1 NORA10 downscaling to the seNorge2 grid

The NORA10 two-meter temperature fields described in section 2.2 are available on a 10 Km- grid but our final grid has 1 Km of grid spacing, then we had to downscale the NORA10 fields to the finer 1 Km-grid.

The 1 Km-grid is the same grid used for the seNorge version 2 products (*Lussana et al.*, 2016b) and it is characterized by the parameters:

- Coordinate Reference system. proj4 string="+proj=utm +zone=33 +datum=WGS84 +units=m +no_defs +ellps=WGS84 +towgs84=0,0,0"
- dimensions (number of grid points): easting = 1195 ; northing = 1550 ;
- grid spacing: easting = 1000 meter ; northing = 1000 meter.

The downscaling of the hourly temperature has been realized by means of a two-step process:

1. a nearest-neighbor interpolation algorithm has been applied from the 10 Km- to the 1 Km-grid. The so-obtained temperature values have been adjusted based on the elevation difference to the 1 Km-grid and a fixed gradient of $-0.0065\,^{\circ}$C/m has been used:

$$T = T\,(\text{nearest-neighbor}) - 0.0065 \cdot [z\,(1\,\text{Km\_grid}) - z\,(10\,\text{Km\_grid})] \qquad (3)$$

   where $z(\ldots)$ indicates the point elevation on the specified grid and T stands for temperature.

2. for each grid point, compute the vertical gradient by considering the temperature values in a neighborhood box of ± 20 Km in both east-west and north-south directions, then apply the gradient correction taking into account the average temperature value and elevation within the ± 20 Km box:

$$T = T\,(\text{box\_average}) + \text{box\_gradient} \cdot [z\,(1\,\text{Km\_grid}) - z\,(\text{box\_average})] \qquad (4)$$

   If the gradient is not computable (too unstable), use $-0.0065\,^{\circ}$C/m as default gradient.

As a downscaling tool, we have used the Gridded Post-Processor (*gridpp*, available at https://github.com/metno/gridpp). Command lines:

1. gridpp input_file output_file -v T -d gradient constantGradient=-0.0065 minElevDiff=-999
2. gridpp input_file output_file -v T -d gradient minElevDiff=-999 searchRadius=20 defaultGradient=-0.0065 averageNeighbourhood=1

The NORA10 downscaled hourly temperature fields are then aggregated to the daily timescale as specified in section 2 for each of the three variables of interest: TANRR, TAXRR and TAMRR.

3.2 Statistical Interpolation

The Optimal Interpolation (OI) scheme combines the observed values stored in the KDVH with the background field derived from NORA10. The OI scheme generates an ensemble of possible analyses. For each of daily minimum, maximum and daily temperature, respectively: TANRR-Nor, TAXRR-Nor and TAMRR-Nor, the OI scheme generates an ensemble of possible analyses and the outputs are the ensemble mean and standard deviation fields derived from such an ensemble of two-meter temperature analyses.

3.2.1 Optimal Interpolation

The Optimal Interpolation (OI: *Gandin and Hardin*, 1965; *Kalnay*, 2003) scheme has been used to obtain the *analysis* vector. The same OI method has been applied to the three variables: TANRR, TAMRR and TAXRR, though with different parameters. The OI method implemented here is similar to the ones described in *Uboldi et al.* (2008); *Lussana et al.* (2016b).

The notation introduced in *Ide et al.* (1997) is adopted. The vectors denoted by $\mathbf{x}^{\cdots}$ indicate quantities at grid points, while the vectors denoted by $\mathbf{y}^{\cdots}$ indicate quantities at station locations.

The linear observation operator $\mathbf{H}$ transforms quantities from the grid-space onto the observation-space. For example, the background at station locations $\mathbf{y}^{\mathrm{b}}$ is obtained as:

$$\mathbf{y}^{\mathrm{b}} = \mathbf{H}\mathbf{x}^{\mathrm{b}} \tag{5}$$

The observation operator we have used is similar to the second step of section 3.1. A box of $\pm 10\,\mathrm{Km}$ around each station location is considered and the gradient is computed by using a linear regression on the temperature values (at grid points) within the box as a function of their elevations. Then, the m-th component of the background vector (i.e. at the m-th station location) is:

$$y^{\mathrm{b}}_m = \mathbf{x}^{\mathrm{b}}\,(\text{box\_average}) + \text{box\_gradient} \cdot [z_m - z\,(\text{box\_average})] \tag{6}$$

where: z_m is the elevation at the m-th station location; $\mathbf{x}^{\mathrm{b}}\,(\text{box\_average})$ is the average value of temperature within the box; $z\,(\text{box\_average})$ is the average elevation within the

box. Both $\mathbf{x}^{\mathrm{b}}\,(\text{box\_average})$ and $z\,(\text{box\_average})$ are computed by considering quantities at grid points only.

The OI analysis $\mathbf{x}^{\mathrm{a}}$ on grid points is:

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \tilde{\mathbf{G}}\left(\tilde{\mathbf{S}} + \varepsilon^2\mathbf{I}\right)^{-1}\left(\mathbf{y}^{\mathrm{o}} - \mathbf{y}^{\mathrm{b}}\right) \tag{7}$$

Where the subscripts a, b and o denote: analysis, background (i.e. NORA10 field) and observation, respectively. The scalar $\varepsilon^2 \equiv \sigma_{\mathrm{o}}^2/\sigma_{\mathrm{b}}^2$ is the ratio between the background and the observation error variances: $\varepsilon^2 = 0$ implies assuming perfect observations, hence exact interpolation; $\varepsilon^2 > 1$ implies a greater confidence in the background field rather than in the observations.

The OI gain matrix $\mathbf{K} = \mathbf{G}\left(\tilde{\mathbf{S}} + \varepsilon^2\mathbf{I}\right)^{-1}$ is expressed by means of: the background error correlation matrix at station locations $\tilde{\mathbf{S}}$ (the covariance matrix is $\mathbf{S} = \sigma_{\mathrm{b}}^2\tilde{\mathbf{S}}$); the $\tilde{\mathbf{G}}$ correlation matrix having as elements the correlations between the background error at grid points and the background error at station locations. Note that the observation error covariance matrix, usually indicated as $\mathbf{R}$, is set to $\sigma_{\mathrm{o}}^2\mathbf{I}$, with all the observations having the same error variance.

The correlation between two generic points $\mathbf{r}_i = (x_i, y_i, z_i)$ and $\mathbf{r}_j = \left(x_j, y_j, z_j\right)$ is specified by means of the correlation function γ:

$$\gamma\left(\mathbf{r}_i, \mathbf{r}_j\right) = \exp\left\{-\frac{1}{2}\left[\left(\frac{d\left(\mathbf{r}_i, \mathbf{r}_j\right)}{D^h}\right)^2 + \left(\frac{\Delta z\left(\mathbf{r}_i, \mathbf{r}_j\right)}{D^z}\right)^2\right]\right\} \tag{8}$$

Where $d\left(\mathbf{r}_i, \mathbf{r}_j\right)$ is the horizontal distance between the two points, and $\Delta z\left(\mathbf{r}_i, \mathbf{r}_j\right)$ is the difference between their elevations. D^h and D^z are the de-correlation distances in the horizontal and vertical directions, respectively. The sum between square brackets in Eq. (8) defines a new three-dimensional distance, where distances in the vertical and in the horizontal directions have different weights. The function γ has been used for the specification of $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{G}}$.

In addition, a Spatial Consistency Test (SCT) has been implemented within the OI scheme as described in *Lussana et al.* (2010, 2016b). In the current work, the i-th observation is flagged as suspect by the SCT and consequently not used in the OI if:

$$\frac{\left(y_i^o - \check{y}_i^a\right)\left(y_i^o - y_i^a\right)}{\sigma_o^2} > 40 \tag{9}$$

Where y_i^a is the leave-one-out cross-validated analysis at the i-th location.

The setting of OI parameter values is discussed in section (3.2.2). The gain matrix in Eq. (7) is completely defined by the values of the three parameters: D^h, D^z, and ε^2, moreover the value of σ_o^2 is needed for the SCT.

3.2.2 OI Parameters

The OI parameter values are estimated for each day. Instead of a single best combination of values, we have decided to take into account a set of OI parameter values for each day. The analyses obtained by the OI scheme with the different configurations constitute our ensemble.

In order to limit the computational time, the selection of the best OI parameter values for $\left(D^h, D^z, \varepsilon^2\right)$ is constrained among all the possible combinations of the following sets:

- $D^h = (10, 25, 50, 75, 100, 125, 150, 175, 200, 300)$ [Km]
- $D^z = (100, 250, 500, 750, 1000, 1250, 1500, 2000)$ [m]
- $\varepsilon^2 = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)$ [-]

Consider the generic day d. In the OI parameter estimation, a time interval of 31 days centered on d has been taken into account. Then, for each combination of parameters $\left(D^h, D^z, \varepsilon^2\right)$ the value of σ_o^2 has been computed as (*Lussana et al.*, 2010):

$$\sigma_o^2 = \frac{\sum_{t=1}^{31} \left[\left(\mathbf{y}^{\mathrm{o}} - \mathbf{y}^{\mathrm{a}}\right)_t^{\mathrm{T}} \left(\mathbf{y}^{\mathrm{o}} - \mathbf{y}^{\mathrm{b}}\right)_t \right]}{\sum_{t=1}^{31} M_t} \tag{10}$$

where: the time index t assumes the values of the 31 days in the interval centered on d; M_t indicates the number of observations available at time t.

The selection of the best set of OI parameter values are based on:

- the Cross-Validation score (*Uboldi et al.*, 2008; *Lussana et al.*, 2016b):

$$\mathrm{CVscore} = \frac{1}{31} \sum_{t=1}^{31} \left\{ \sqrt{\frac{1}{M_t} \sum_{m=1}^{M_t} \left(y_m^o - \check{y}_m^a\right)^2} \right\} \tag{11}$$

- a likelihood function relating the empirical estimate of $\sigma_o^2 + \sigma_b^2$ to the theoretical one.

  The empirical estimate is (*Desroziers et al.*, 2005):

$$\left(\sigma_o^2 + \sigma_b^2\right)_{emp} = \frac{1}{31} \sum_{t=1}^{31} \left\{ \frac{1}{M_t} \sum_{m=1}^{M_t} \left(y_m^o - y_m^b\right)^2 \right\} \tag{12}$$

The theoretical one is based on Eq. (10) and on the definition of ε^2, then it can be written as:

$$\left(\sigma_o^2+\sigma_b^2\right)_{theo}=\sigma_o^2+\frac{\sigma_o^2}{\varepsilon^2} \tag{13}$$

Note that Eq. (12) is based on observations and model data, while Eq. (13) depends on the analysis, thus on the values of $\left(D^h,D^z,\varepsilon^2\right)$. As a consequence, these two Equations provides a further relation between the observed and model data and the OI parameters.

In practice, to implement the relation we have defined the two vectors $\mathbf{L}_{emp}$ and $\mathbf{L}_{theo}$:

$$\mathbf{L}^{emp} = \text{pnorm}\left(innovation,\text{mean}=0,\text{sd}=\left(\sigma_o^2+\sigma_b^2\right)_{emp}\right) \tag{14}$$

$$\mathbf{L}^{theo} = \text{pnorm}\left(innovation,\text{mean}=0,\text{sd}=\left(\sigma_o^2+\sigma_b^2\right)_{theo}\right) \tag{15}$$

The set *innovation* includes all the I innovation values $y_i^{\text{o}}-y_i^{\text{b}}\, i=1,\ldots,I$ for all the station locations and for the 31 days considered. The "pnorm" symbol indicates the Gaussian distribution function values returned for each element of the *innovation* set, given the parameters: mean value (mean) and standard deviation (sd). The two vectors $\mathbf{L}_{emp}$ and $\mathbf{L}_{theo}$ have the same number of elements I as the *innovation* set. The value of a generic element of $\mathbf{L}_{emp}$ quantifies the likelihood of having the corresponding innovation value given our estimate for the standard deviation as $\left(\sigma_o^2+\sigma_b^2\right)_{emp}$. A similar statement holds true for $\mathbf{L}_{theo}$ and $\left(\sigma_o^2+\sigma_b^2\right)_{theo}$.

The likelihood function J which quantifies the goodness-of-fit of the theoretical estimate $\left(\sigma_o^2+\sigma_b^2\right)_{theo}$ to the empirical one $\left(\sigma_o^2+\sigma_b^2\right)_{emp}$ is then defined as:

$$J=\frac{1}{I}\sum_{i=1}^{I}\left(L_i^{emp}-L_i^{theo}\right)^2 \tag{16}$$

where L_i^{emp} and L_i^{theo} are the i-th elements of the vectors defined in Eq. (15). In our definition of J, we have considered the entire distribution of the innovation values instead of considering only the agreement between two values, such as the mean values defined in Eqs. (12)-(13).

Let us go back to the selection of values for the OI parameters $\left(D^h,D^z,\varepsilon^2\right)$. Once set a value for the triplet $\left(D^h,D^z,\varepsilon^2\right)$, the σ_o^2 value is automatically obtained by Eq. (10).

Consider the possible combinations of $\left(D^h,D^z,\varepsilon^2\right)$ listed above, then for each of them the CVscore and J values are computed as in Eqs. (11)-(16), respectively. Our selection

of best values for $\left(D^h, D^z, \varepsilon^2\right)$ includes those triplets for which: CVscore is within the smallest 10% of values and at the same time J is within the smallest 25% of values.

The number of elements within a selection of best values of $\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)$ is usually around 25.

In the Figures 3-5 the time series for the median of D^h, D^z and ε^2 distribution of daily "best" values are shown. To facilitate the intercomparison between TAMRR-Nor, TANRR-Nor and TAXRR-Nor, each Figure is composed by three panels (one for each variable).

In Figure 3, the median of D^h is roughly around 100 Km for the three variables.

The Figure 4 shows D^z. For TAMRR-Nor and TANRR-Nor, the median of D^z varies mostly between 500 m and 1000 m, while for TAXRR-Nor it varies more around 500 m.

With respect to ε^2, in Figure 5 both TAMRR-Nor and TANRR-Nor show a gradual increase of the median of their distributions, which means that the OI parameter optimization procedure tends to give more weight to the NORA10-derived background over time. In particular, the improvements in the model background due to the introduction of the ECMWF operational analyses from 2002 onward are evident. The situation seems to be different for TAXRR-Nor, which presents a ε^2 rather stable in time.

In the Figures 6-8 the typical year for the empirical distribution of D^h, D^z and ε^2 best values are shown. In general, the seasonal cycle is more important for D_z, which has higher values during summer. In the case of TAXRR-Nor, ε^2 in Figure 8 shows a pronounced seasonal cycle too, which results in the attribution of more weight to the observations in the summer.

The combined effects of the station distribution and of the OI parameter choices are summarized in Figure 9 by means of the CVIDI-score (*Lussana et al.*, 2016b) for the three variables. The (dimensionless) CVIDI-score can assume values between 0 and 1 and as reference values one may consider: < 0.45 for isolated grid points; > 0.85 for grid points located in dense station areas. In general, the CVIDI-score is greater than 0.8: most of the grid point are located in dense station areas and, as a consequence, the adjustment to the background due to the observations can influence a large portion of the domain.

It might be worth noticing that for TAXRR-Nor in 2007 there is a period presenting a sharp discontinuity in the CVIDI-score, which deserves further investigations.

3.2.3 Implementation choices

Consider the selection of best values for $\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)$, then the best analyses of TANRR-Nor, TAXRR-Nor and TAMRR-Nor are the mean of the corresponding ensemble of analyses, $\overline{\mathbf{x}}^{\mathrm{a}}$:

$$\overline{\mathbf{x}}^{\mathrm{a}} = \mathrm{mean}\left(\mathbf{x}^{\mathrm{a}}\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)\right) \tag{17}$$

As an estimate of the ensemble spread, the standard deviations of the TANRR-Nor, TAXRR-Nor and TAMRR-Nor analysis fields have been computed:

$$\mathrm{sd}^{\mathrm{a}} = \mathrm{standard\_deviation}\left(\mathbf{x}^{\mathrm{a}}\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)\right) \tag{18}$$

sd^{a} is related to the uncertainty due both to the station distribution and to our selection of the OI parameters.

As diagnostic tools, the average Integral Data Influence (IDI: *Uboldi et al.*, 2008; *Lussana et al.*, 2016b) and its standard deviation have been computed:

$$\overline{\mathbf{x}}^{\mathrm{IDI}} = \mathrm{mean}\left(\mathbf{x}^{\mathrm{IDI}}\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)\right) \tag{19}$$

$$\mathrm{sd}^{\mathrm{IDI}} = \mathrm{standard\_deviation}\left(\mathbf{x}^{\mathrm{a}}\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)\right) \tag{20}$$

In our OI scheme, a SCT is performed for each element of the selection of best values $\left(D^h, D^z, \varepsilon^2, \sigma_o^2\right)$. We have decided to flag an observation as suspect and discard it from the OI if the observation fails just one of those SCT.

4 Results

In Figures 10-12 the time series for the daily biases for TAMRR-Nor, TANRR-Nor and TAXRR-Nor, respectively, are shown. Each Figure is composed by three panels: the daily averaged (over all the station locations) value of the innovation (i.e. observation minus background) is shown on the top panel; the middle panel refers to the daily averaged observation minus CV-analysis and it can be interpreted as a result valid for the analysis at grid points; the observation minus analysis daily average is shown in the bottom panel. Ideally, the average of all these quantities should be close to 0°C. However, the time series for the innovation shows the presence of a bias for all the variables: for TAMRR-Nor, a negative bias is present and its magnitude is reduced over time; for TANRR-Nor a negative bias is present too and it remains stable in time; for TAXRR-Nor a positive, stable in time bias is present. The innovation biases show a seasonal cycle, with the higher values in winter. Our spatial interpolation scheme is able to effectively reduce the bias for all

variables and throughout the whole year, as can be easily seen by comparing the ordinate scales of the top panels with the other panels in the Figures.

The time series for the daily root-mean-square (RMS) for the innovation, CV-analysis and analysis residuals are shown in Figures 13-15 for TAMRR-Nor, TANRR-Nor and TAXRR-Nor, respectively. The daily root-mean-square of the CV-analysis residuals is also called CVscore (*Lussana et al.*, 2016b). As for the bias, also in the case of the RMS of the different diagnostics our OI scheme is able to improve the precision of the predicted temperature. Most noticeably, in the case of TANRR-Nor the RMS of the innovation is on average around 3°C, and up to 8°C in the worst cases, while the CVscore is a bit less than 2°C on average and 4°C for the worst cases.

The improved precision of the analyses compared to the model background is clearly evident also in Figure 16, where the RMS are shown for a typical year, and in Figure-17 by considering the Northern and the Southern parts of the domain separately. The differences between the analysis at station locations and at grid points are on average less than 0.5°C. Despite the difference in station densities between the sparser distribution of stations in the North of Norway and the denser station network of the South, the precision of the analysis is comparable. Note that both Figures 16- 17 show the seasonal cycle of the RMS deviations. For TAMRR-Nor and TANRR-Nor, there is a clear signal of higher uncertainties in the winter and a better precision in the summer. On the other hand, for TAXRR-Nor the situation is different: the analyses are on average more precise in spring and autumn, while during winter and summer they are characterized by similar uncertainties.

The evaluation based on the (empirical) joint probability distributions for: background and observation; CV-analysis and observation; analysis and observation are shown in Figures 18- 23. In particular, the Figures 18, 20 and 22 show the density plots with the empirical joint probability distributions; the Figures 19, 21 and 23 show the empirical conditional probability density functions given a selection of the observed values and the corresponding parameters ($N(\mu, \sigma)$, where μ is the mean and σ the standard deviation) for the best-fitting Gaussian probability density function. The Figures 18, 20 and 22 are quite similar to each other and they show the impact of the OI scheme on the bias adjustment and in the reduction of the uncertainty, especially for extreme values. The Fig-

ures 19, 21 and 23 quantify the accuracy and precision of the background, CV-analysis and analysis across the range of observed values. For negative temperatures, the analysis is a less precise estimate of the true temperature than for positive values. For the most part of the observed range of temperatures, our analysis does not show a significant bias and its standard deviation is around 1°C. In the case of TANRR-Nor, the most challenging cases correspond to the occurrence of extremely low temperatures: for observed values around −25°C, our CV-analysis still has on average a bias of 4°C and a standard deviation of 3°C.

In the final part of our evaluation, two case studies are presented. We will focus on the coldest and the warmest days within the time interval 1980-2012, which are: 11 January 1987 (coldest) and 2 August 1994 (warmest). In Figures 24-29 the maps for the two case studies are displayed for TAMRR-Nor, TANRR-Nor and TAXRR-Nor. The analysis fields are shown on the left panel, while the model background fields based on NORA10 are shown on the right. The dots both mark the station locations and show the observed values. The overall adjustment of the background field towards the observations is evident and its effect is quantified by the significant improvements in the bias correction and in the CVscore reduction (see the main title of each Figure). For example, in the case of TANRR-Nor for 11 January 1987, the 192 observations available allow us to reduce the bias of the predicted two-meter minimum field and to improve the precision of almost 2°C on average. Besides, it is clear that the OI scheme would strongly benefit from a denser station network in Sweden and Finland.

As an example of the other available analysis products (see section 3.2.3), Figures 30 and 31 show the IDI fields and the ensemble spread (i.e. standard deviation). The IDI fields in Figure 30 are the mean (on the left) and the standard deviation (on the right) of the IDI ensemble. The standard deviations of the analysis ensemble, which represents the ensemble spread, in the two case studies for TANRR-Nor (left panel, 11 January 1987) and TAMRR-Nor (right panel, 2 August 1994) are shown in Figure 31. Both these products are expected to provide the users an estimate of the uncertainties related to our OI scheme. In particular, the IDI-related products focuses on the impact of the station distribution on the analysis, while the ensemble spread is related to the uncertainties introduced by our selection of the OI parameters.

5 Conclusions

A spatial interpolation scheme for daily minimum, average and maximum two-meter temperature (TANRR-Nor, TAXRR-Nor and TAMRR-Nor, respectively) based on statistical interpolation has been developed and implemented at MET Norway. The statistical method is an implementation of Optimal Interpolation (OI), which combines a model derived background with in-situ observations from the climate database.

The model considered is the NORA10 high-resolution hindcast dataset, which has been downscaled on the 1-Km grid often used for hydrology and climatology in Norway. Two original procedures to obtain the daily (06-06 UTC) minimum and maximum two-meter temperature observations from the standard parameters available in the climate database have been implemented.

The standard OI scheme relies on an optimization procedure to set the OI parameters, which are: the background and observation error variances and the error correlation functions. The OI configuration plays a crucial part in the determination of the final analysis quality. In this work, a selection of around 25 different OI configurations has been used for each day and variable to obtain an ensemble of analysis fields. The OI configurations have been selected among the ones which guarantees the best-possible results. The OI products available to the users are: the analysis ensemble mean, the ensemble spread and a diagnostic product based on the Integral Data Influence (IDI). The ensemble spread reflects the sensitivity of our OI scheme to the choice of the OI parameters.

The evaluation shows that the NORA10 background fields present systematic differences if compared with the observed values. The OI scheme is able to adjust the analysis for the presence of bias in NORA10 temperature fields and to improve the precision of the predicted values. It has been verified that such improvements occur for the daily minimum, maximum and average temperatures and throughout the whole year. Only for extremely low temperature (below -25°C), a significant bias is still present in the analysis. For the most part of the observed range of temperatures, our analysis does not show a significant bias and its standard deviation is around 1°C.

Future developments might include: a refined downscaling of NORA10 onto the high-resolution grid based on geographical elements; the inclusion of more data in the climate database, especially from the neighboring countries of Sweden and Finland; the evaluation of the ensemble spread as a descriptor of the actual analysis uncertainty.

Acknowledgements

The presented activity was partly funded within the framework of the *Felles aktiviteter NVE-MET tilknyttet nasjonal flom- og skredvarslingstjeneste*.

6 Figures

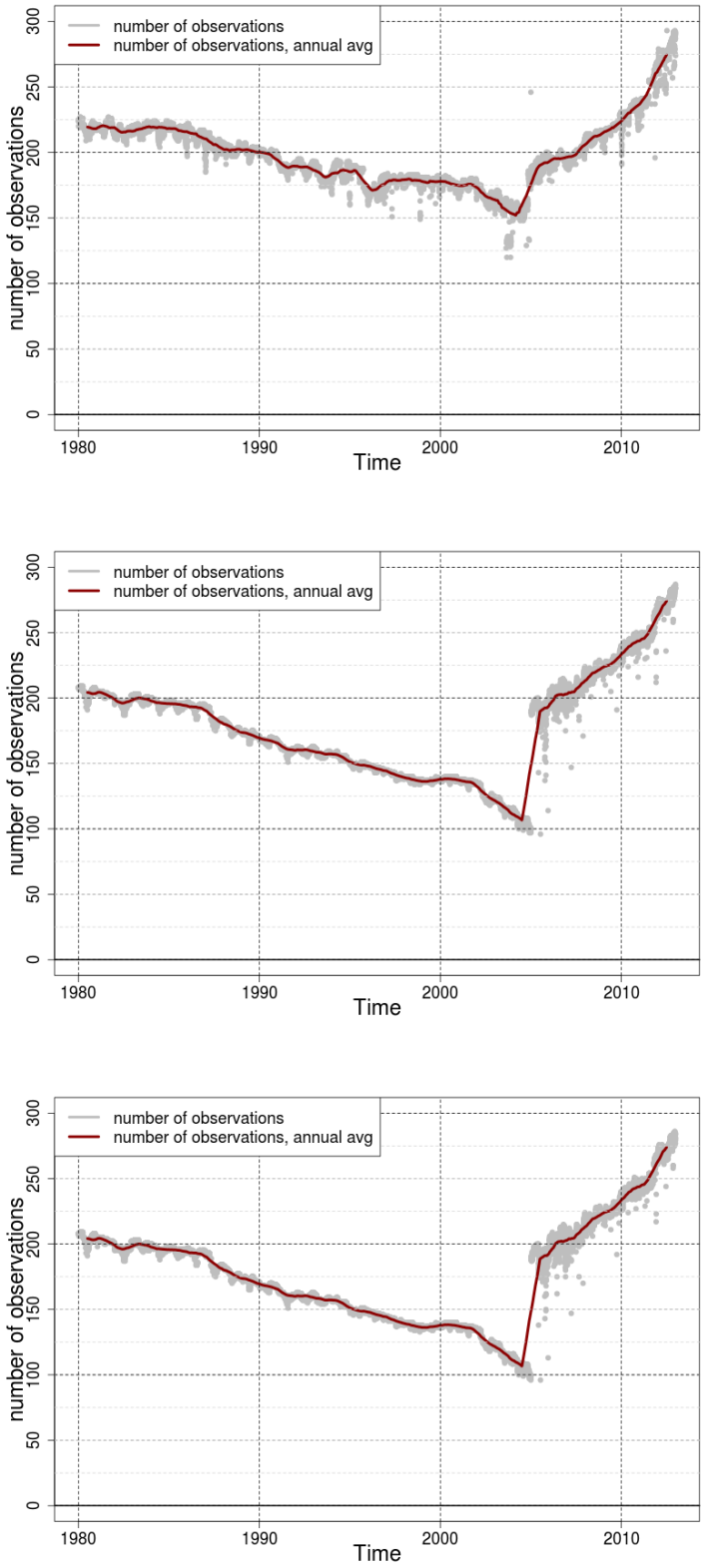


Figure 1: Time series of the number of available observations (1 point=1 day) for: TAMRR-Nor (top panel); TANRR-Nor (middle); TAXRR-Nor (bottom). Time interval: 1980-2012.

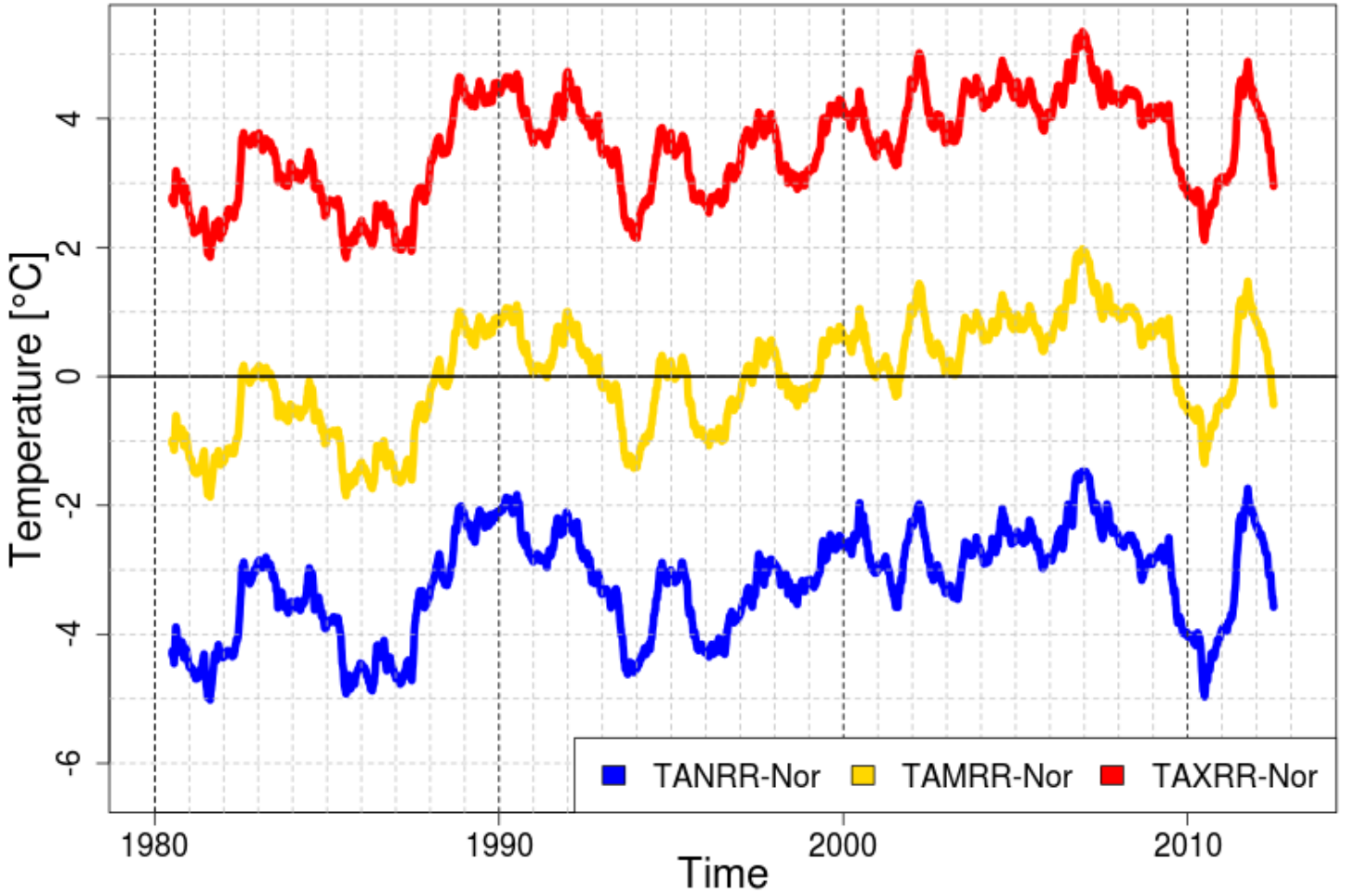


Figure 2: Time series of the TANRR-Nor, TAMRR-Nor, TAXRR-Nor 1-year centered moving average of anomalies respect to the mean TAMRR-Nor. Time interval: 1980-2012.

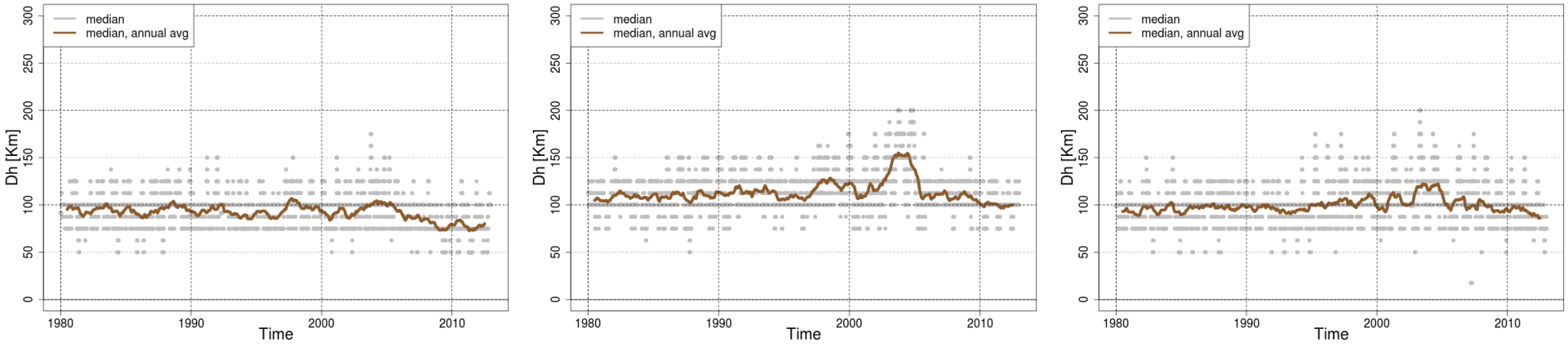


Figure 3: Time series of the median of the D^h distribution for: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Time interval: 1980-2012.

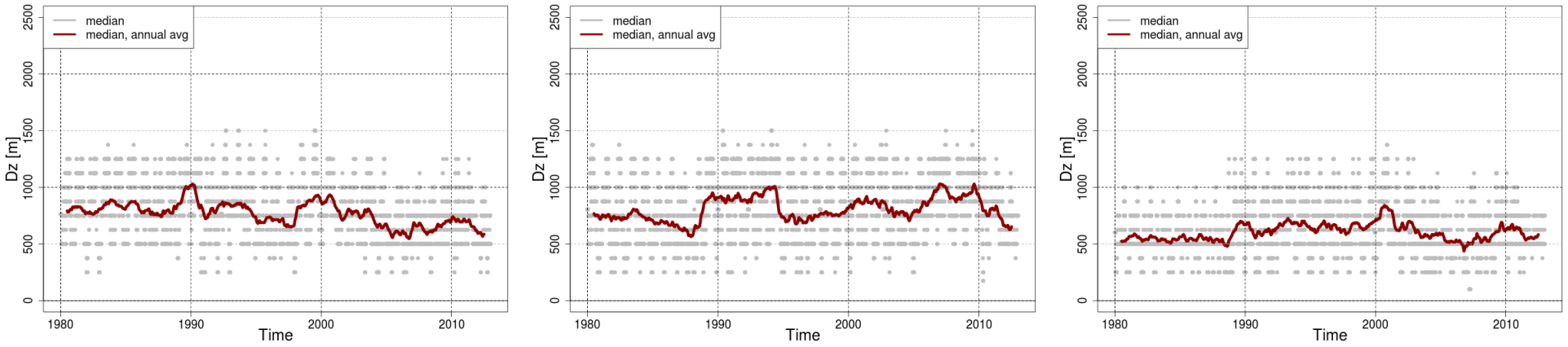


Figure 4: Time series of the median of the D^z distribution for: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Time interval: 1980-2012.

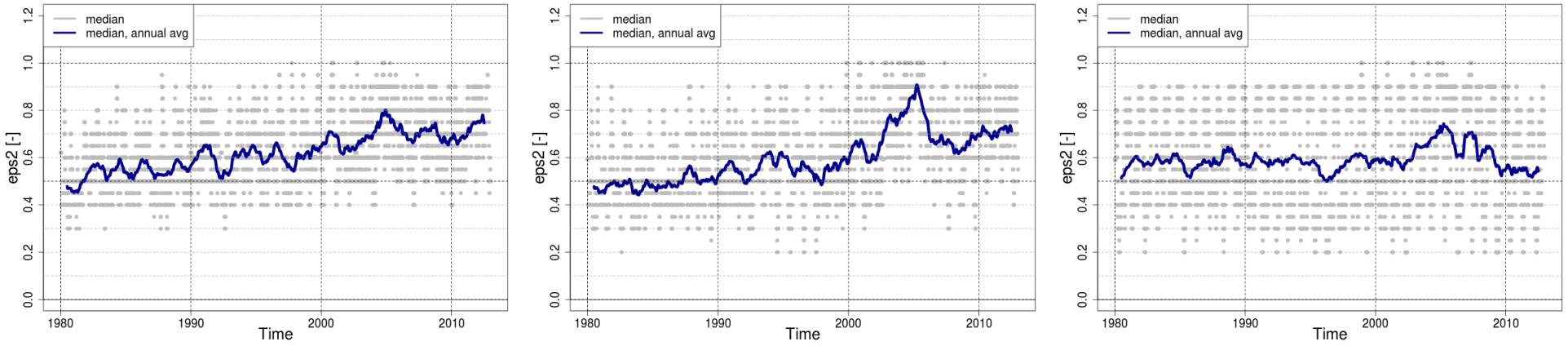


Figure 5: Time series of the median of the ε^2 distribution for: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Time interval: 1980-2012.

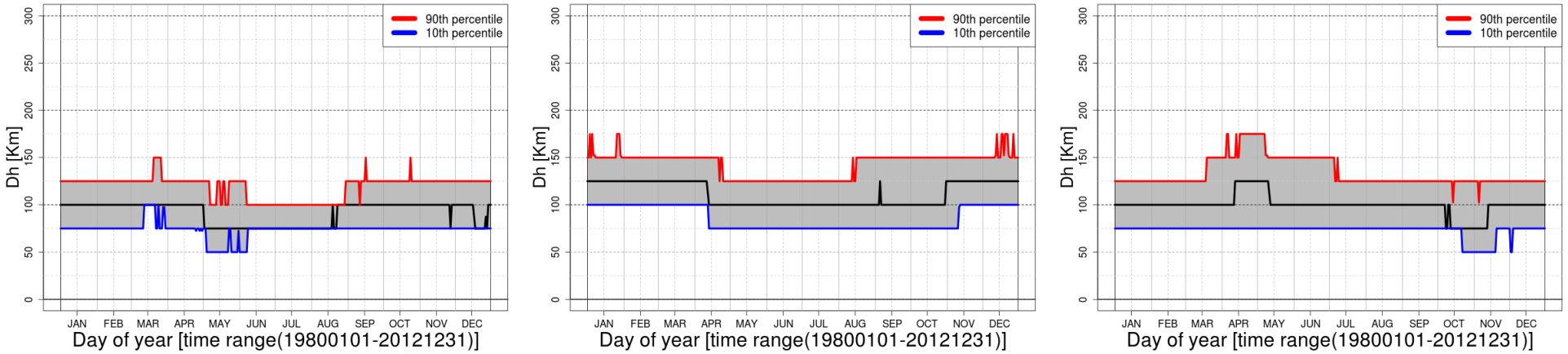


Figure 6: Empirical distribution of D^h for each day of the year: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Based on 1980-2012 data.

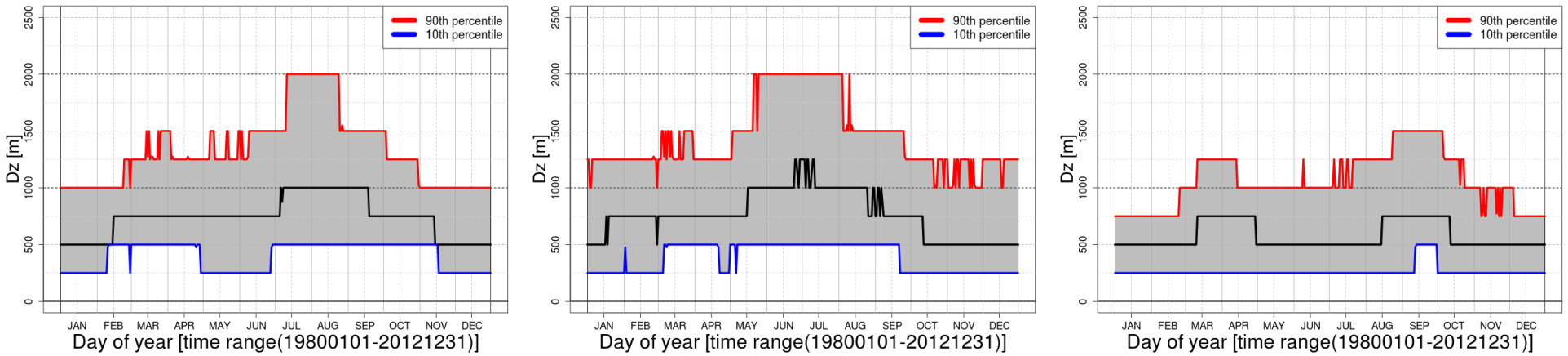


Figure 7: Empirical distribution of D^z for each day of the year: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Based on 1980-2012 data.

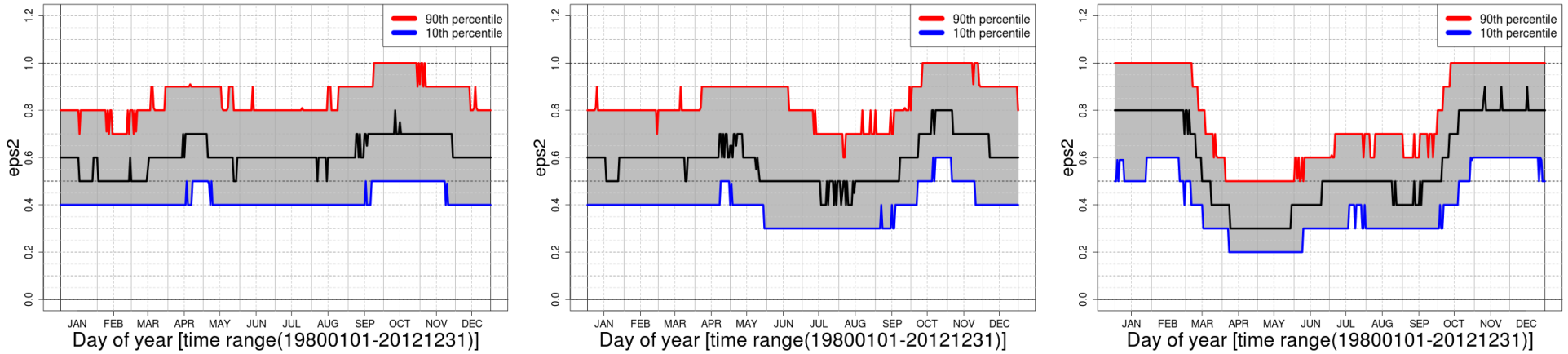


Figure 8: Empirical distribution of ε^2 for each day of the year: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). Thick lines show the 1-year centered moving averages. Based on 1980-2012 data.

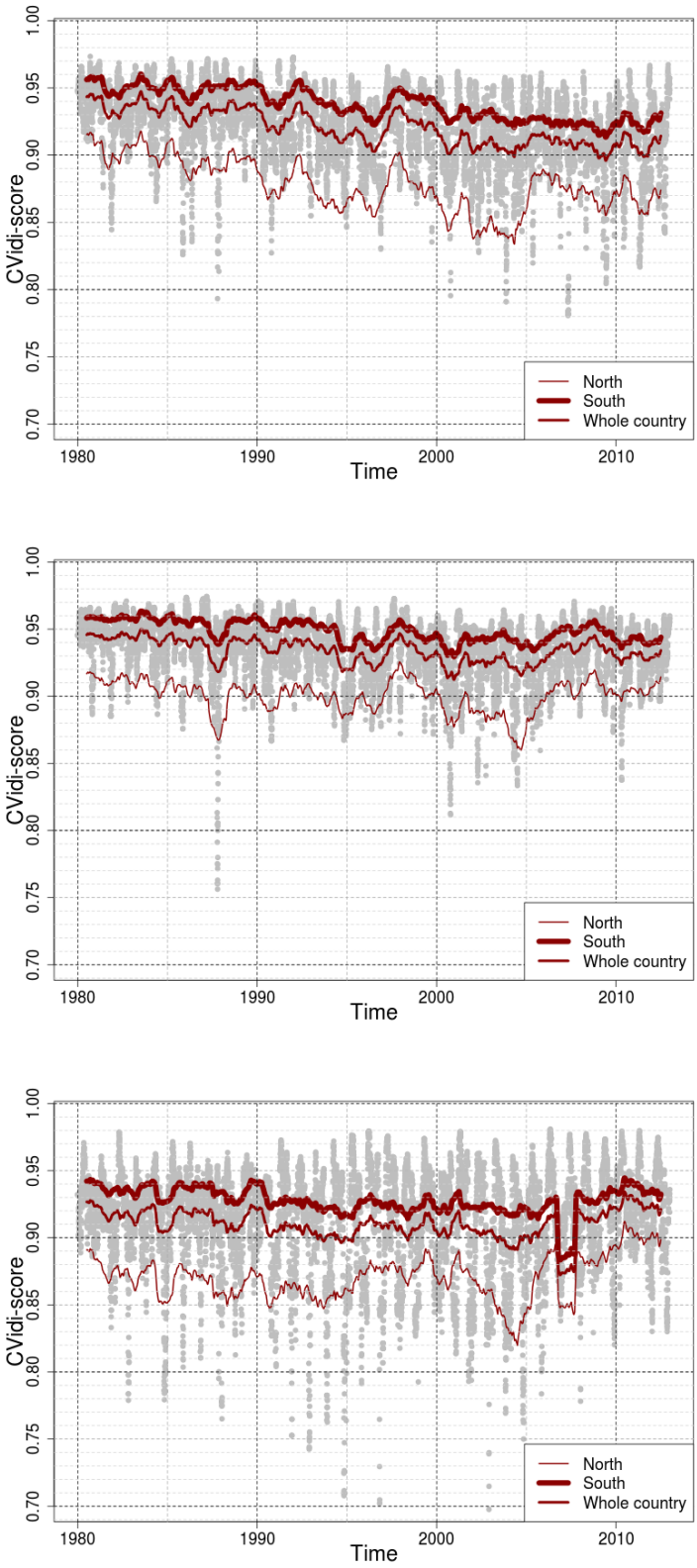


Figure 9: Time series of the CVIDI-score (1 point=1 day) for: TAMRR-Nor (top panel); TANRR-Nor (middle); TAXRR-Nor (bottom). Time interval: 1980-2012.

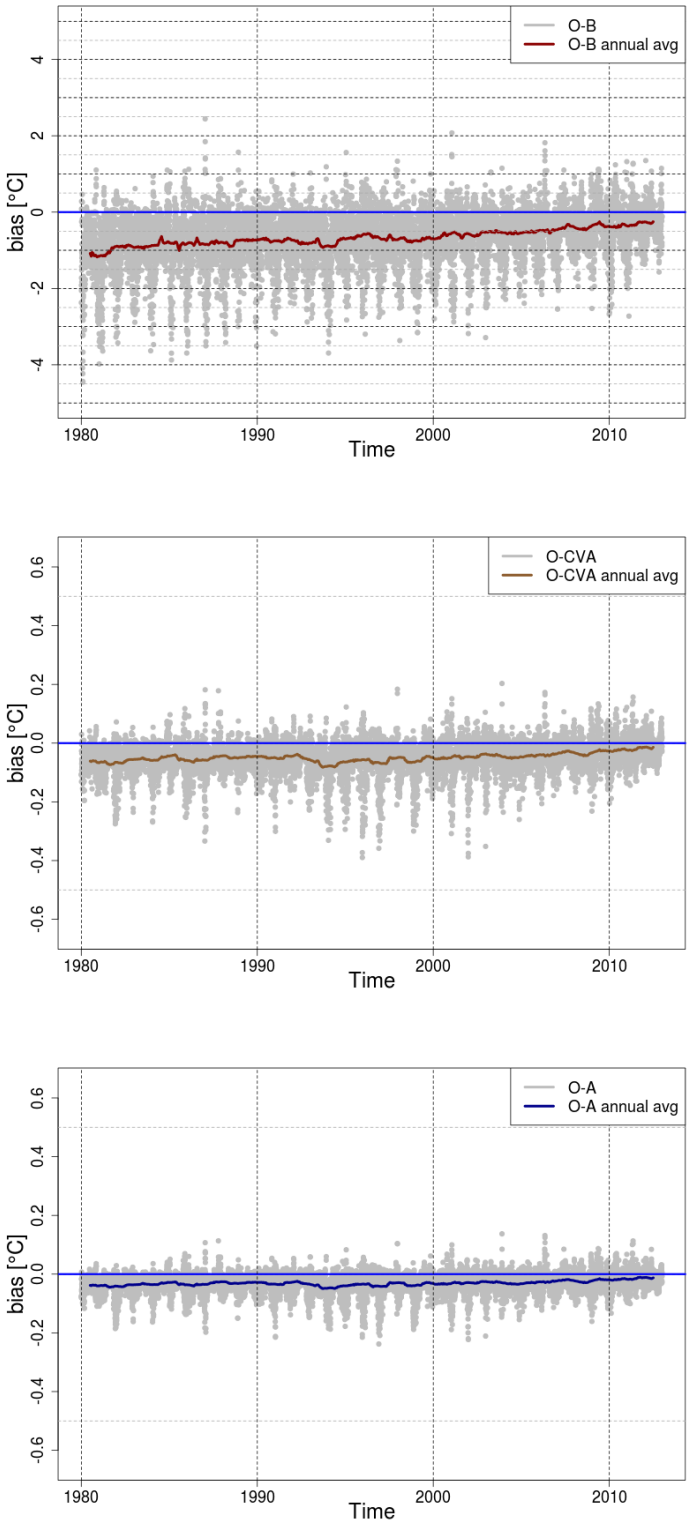


Figure 10: TAMRR-Nor. Times series of the bias for: innovation O-B (top panel); CV-analysis O-CVA (middle); residual O-A (bottom)

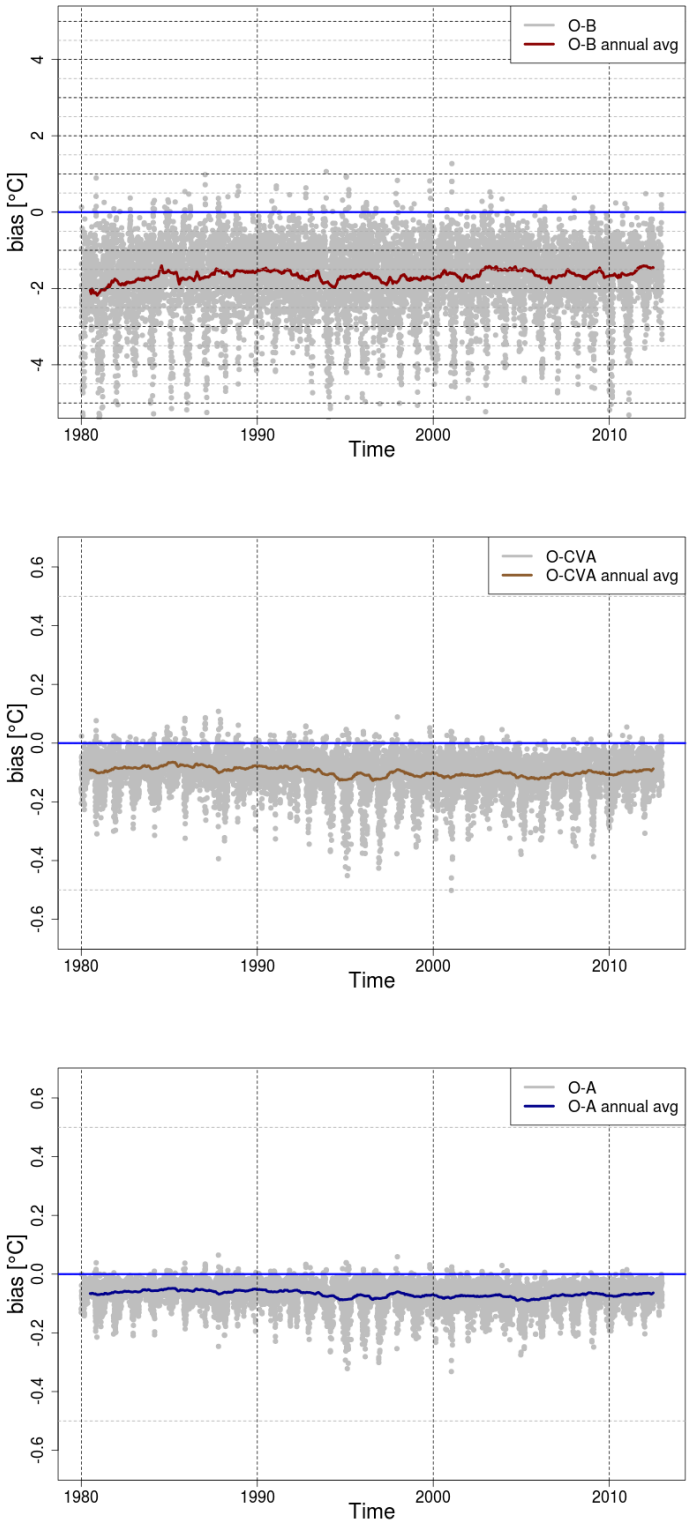


Figure 11: TANRR-Nor. Times series of the bias for: innovation O-B (top panel); CV-analysis O-CVA (middle); residual O-A (bottom)

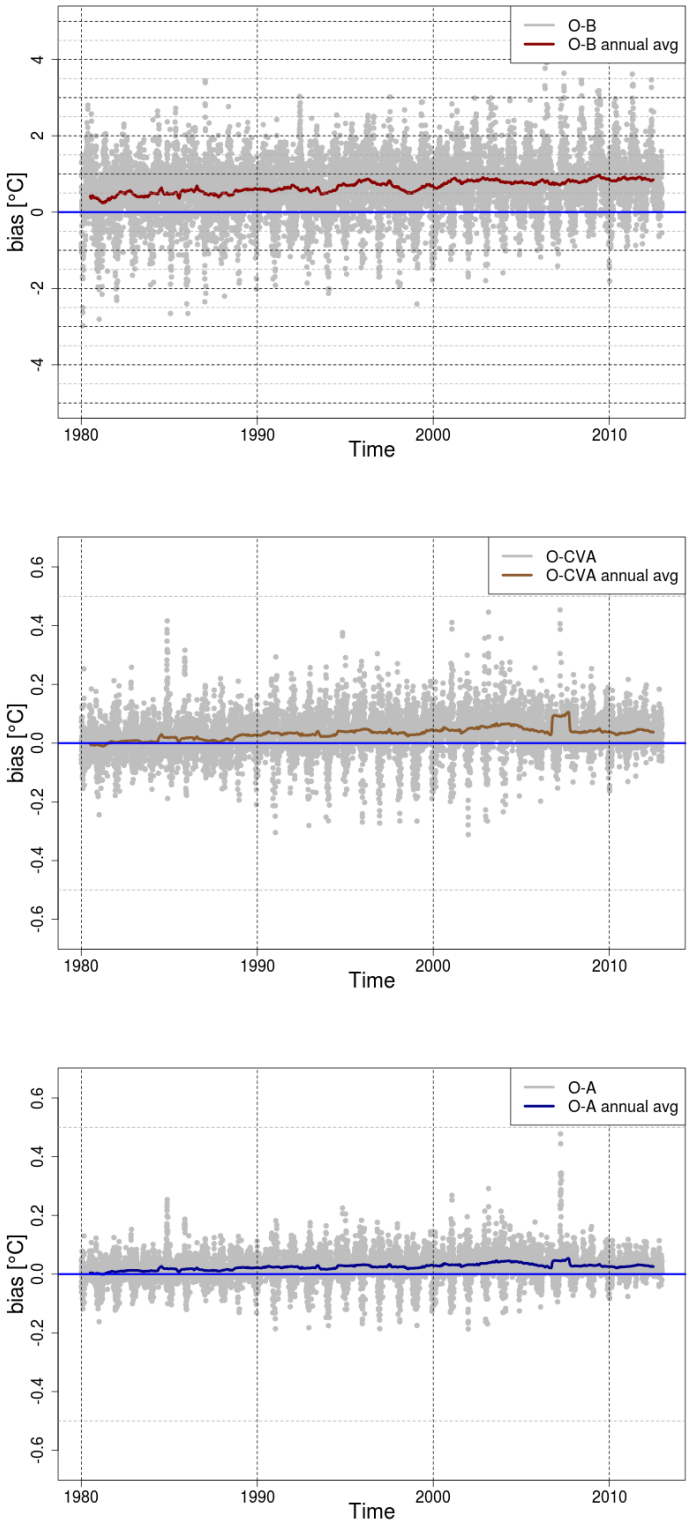


Figure 12: TAXRR-Nor. Times series of the bias for: innovation O-B (top panel); CV-analysis O-CVA (middle); residual O-A (bottom)

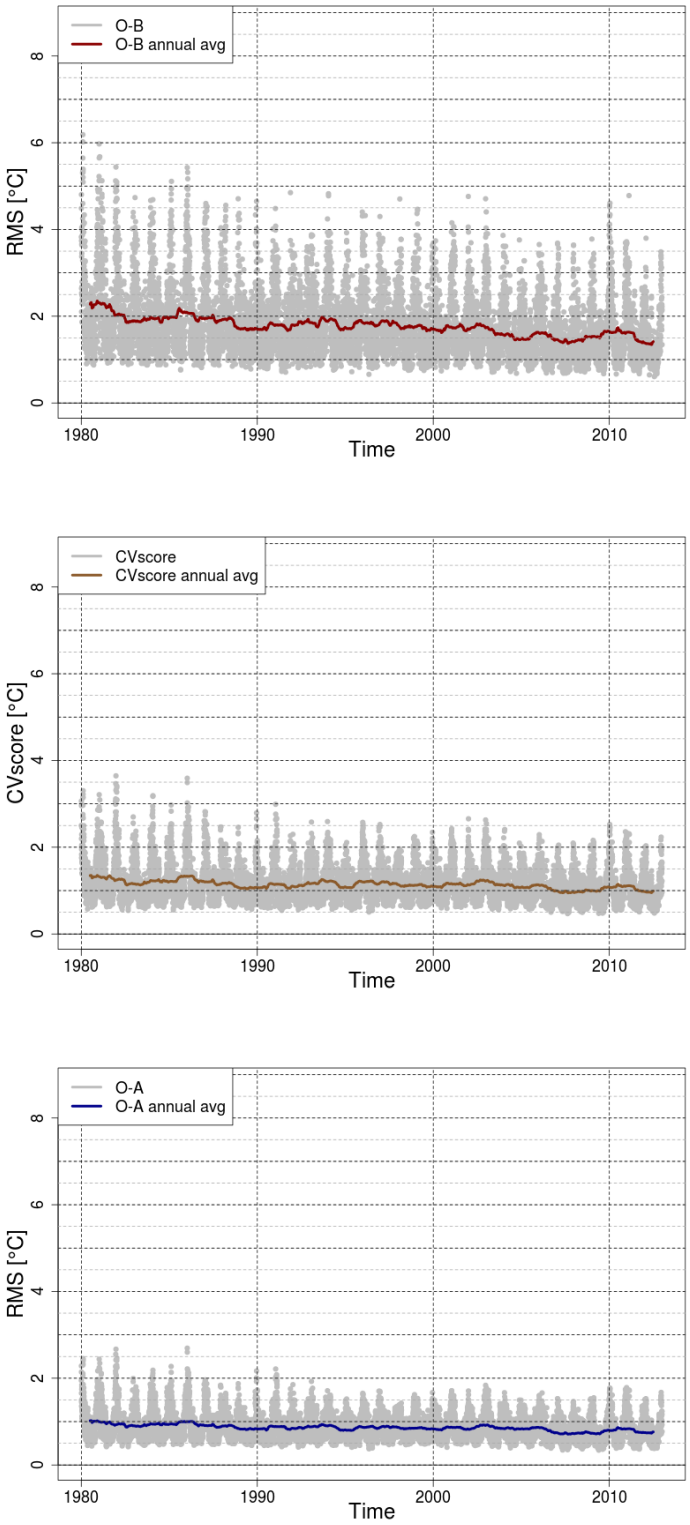


Figure 13: TAMRR-Nor. Times series of the root mean square for: innovation O-B (top panel); CV-analysis O-CVA or CVscore (middle); residual O-A (bottom)

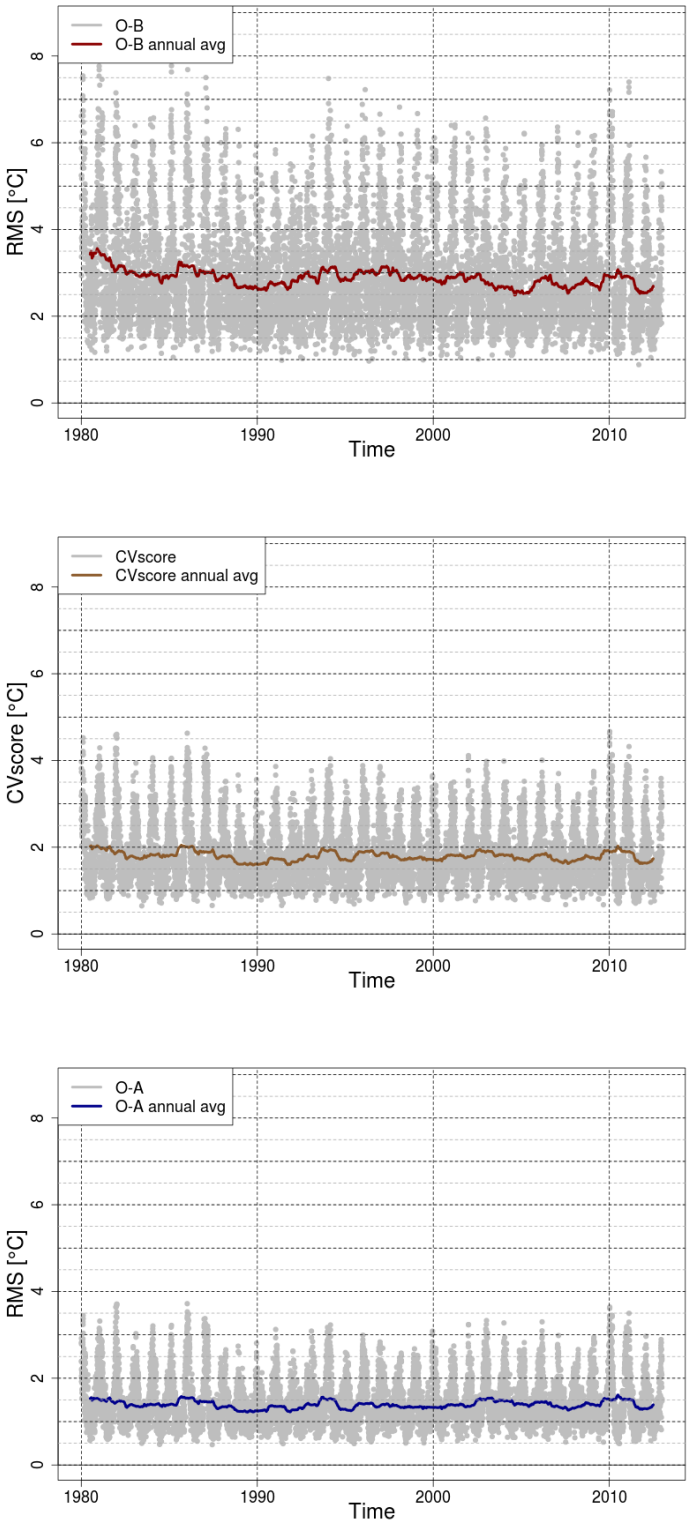


Figure 14: TANRR-Nor. Times series of the root mean square for: innovation O-B (top panel); CV-analysis O-CVA or CVscore (middle); residual O-A (bottom)

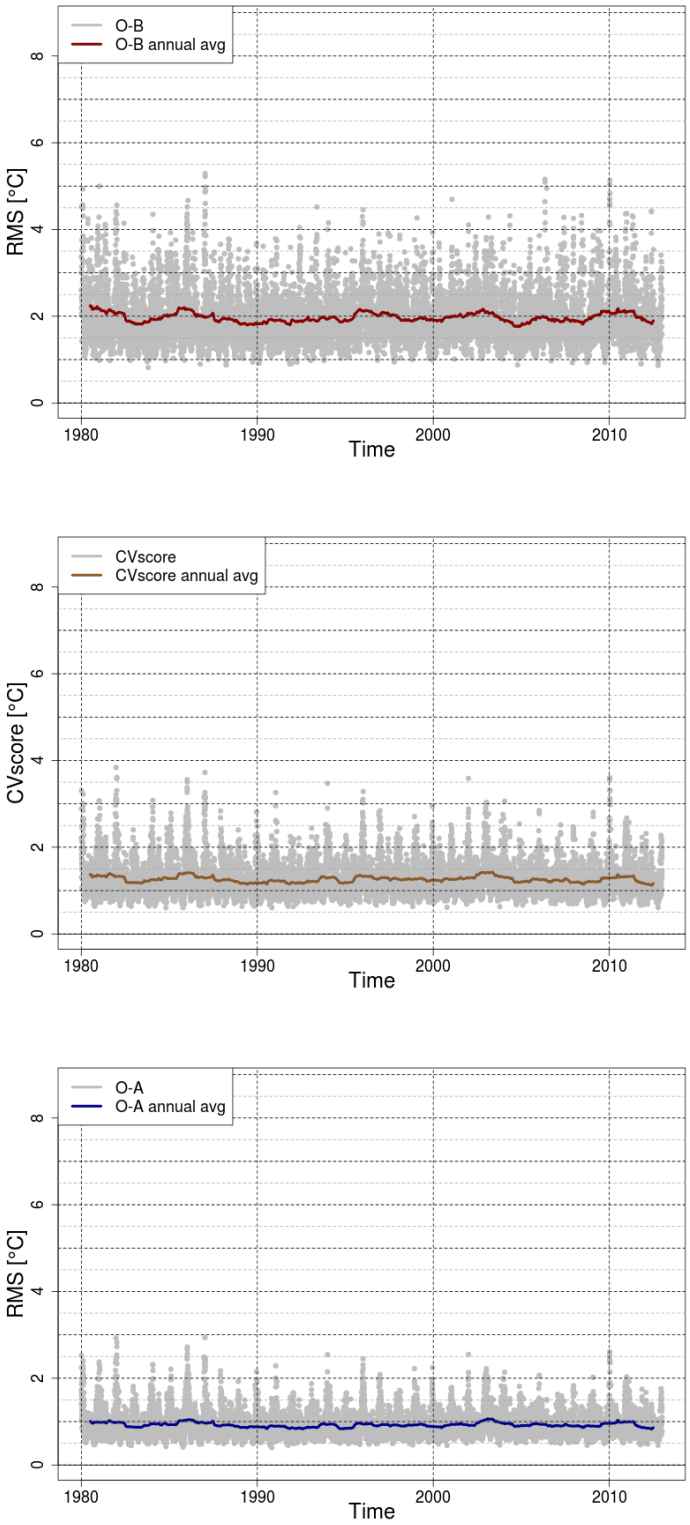


Figure 15: TAXRR-Nor. Times series of the root mean square for: innovation O-B (top panel); CV-analysis O-CVA or CVscore (middle); residual O-A (bottom)

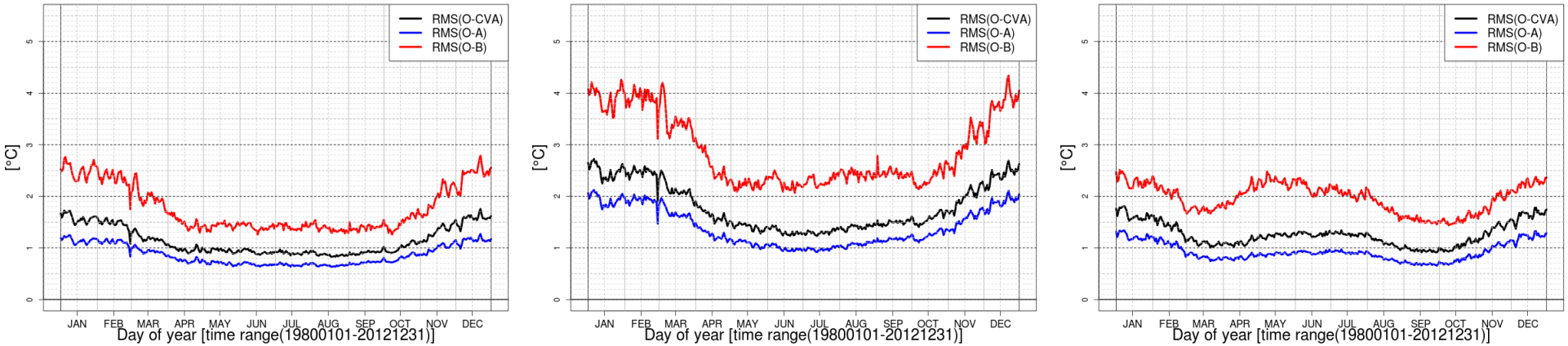


Figure 16: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). average root mean square for each day of the year of: CV residual (black); analysis residual (blue); innovation (red).

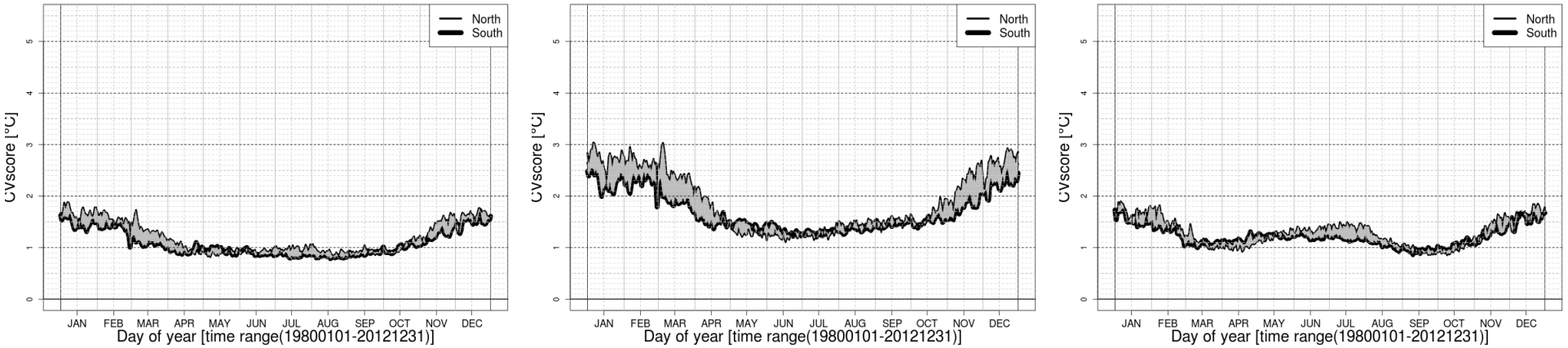


Figure 17: *TAMRR-Nor* (left), *TANRR-Nor* (central), *TAXRR-Nor* (right). average CVscore for each day of the year of. Northern and Southern Norway have been considered separately.

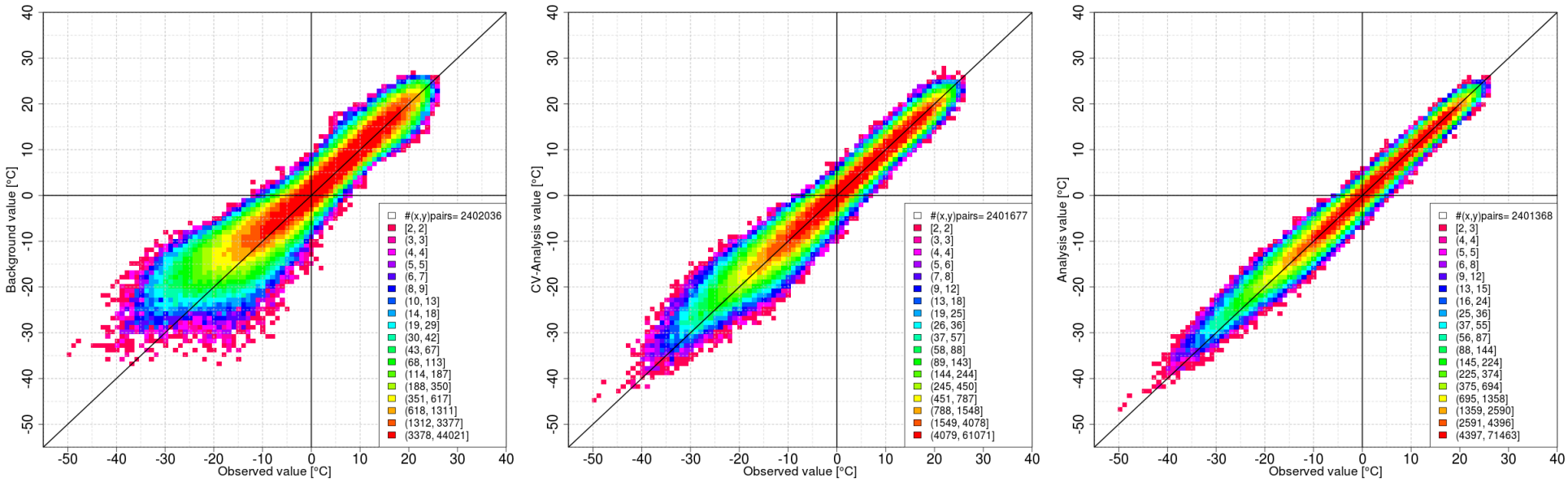


Figure 18: TAMRR-Nor. Density plots of: background *vs* observations (left); CV-analysis *vs* observations (central); analysis *vs* observations (right). Based on 1980-2012 data. The shading indicates the number of pairs (or collocations) within each 1°C by 1°C cell and the total number of collocations is reported in the legend.

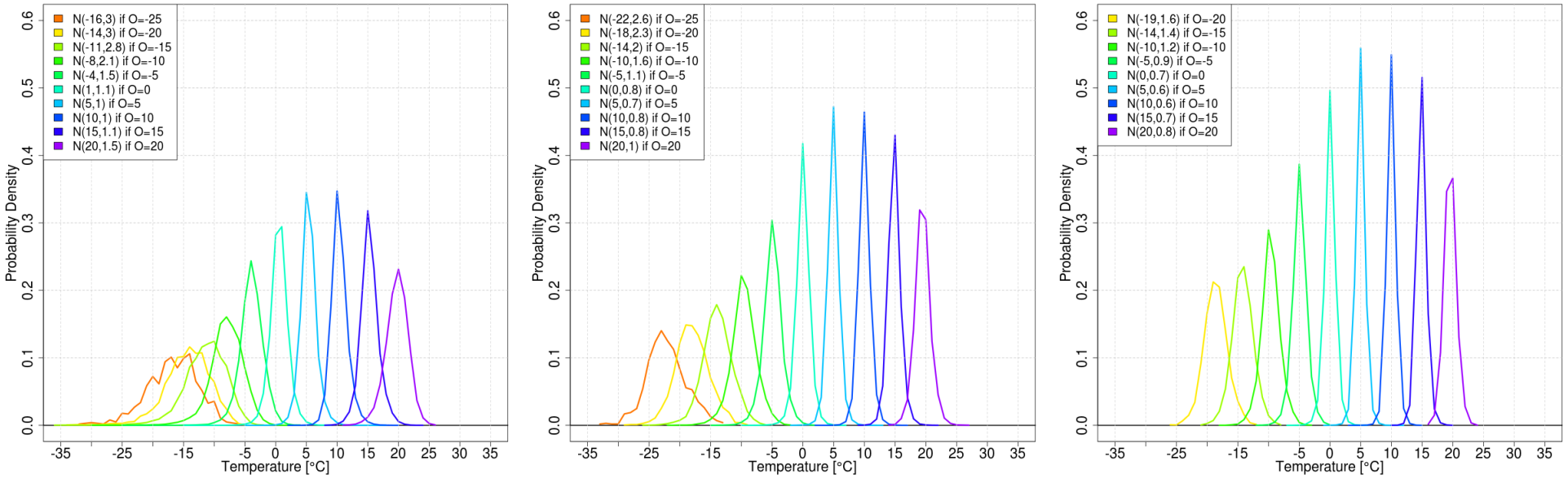


Figure 19: TAMRR-Nor. Empirical conditional probability density function given the observed value O for: background (left); CV-analysis (central); analysis (right). Based on 1980-2012 data.

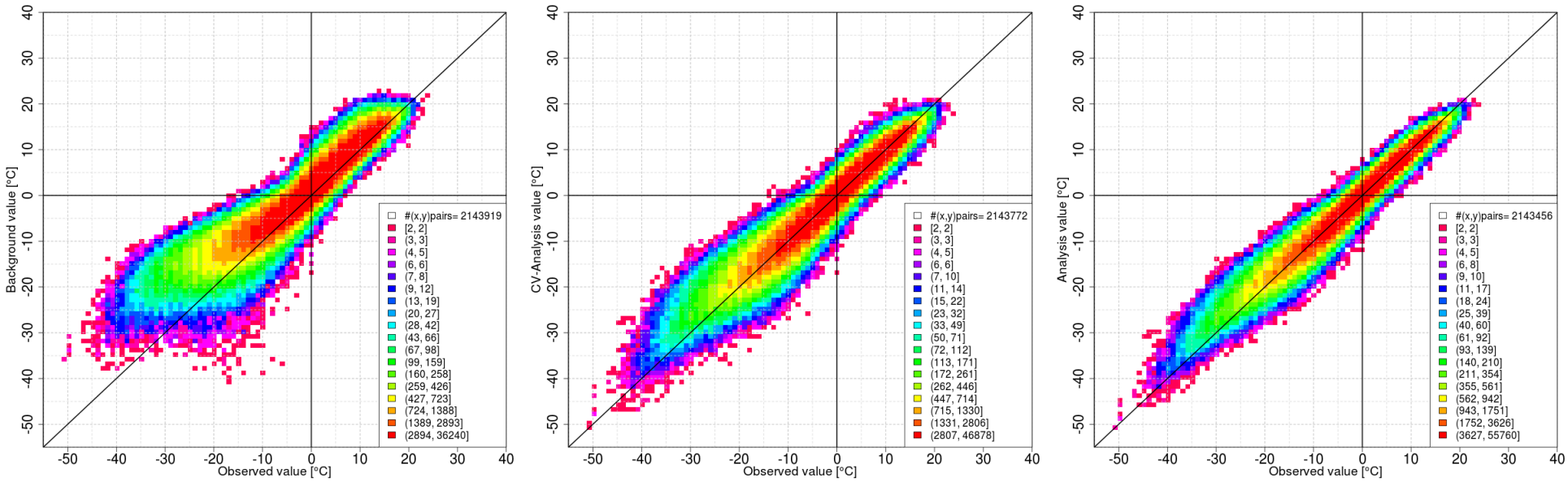


Figure 20: TANRR-Nor. Density plots of: background *vs* observations (left); CV-analysis *vs* observations (central); analysis *vs* observations (right). Based on 1980-2012 data. The shading indicates the number of pairs (or collocations) within each 1°C by 1°C cell and the total number of collocations is reported in the legend.

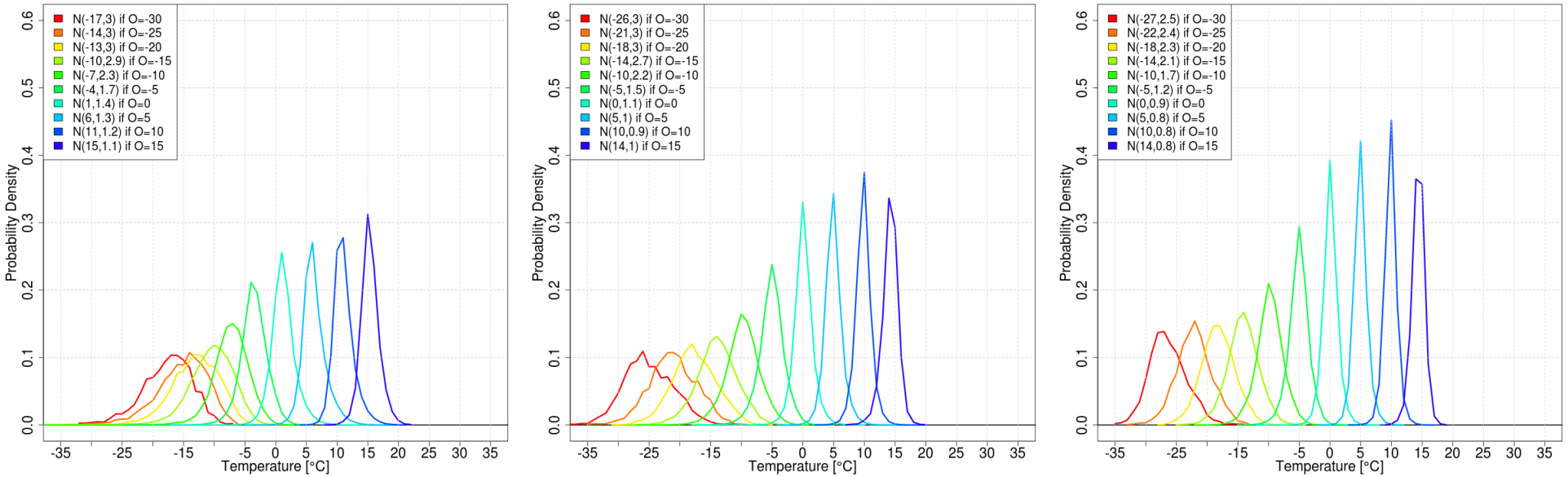


Figure 21: TANRR-Nor. Empirical conditional probability density function given the observed value O for: background (left); CV-analysis (central); analysis (right). Based on 1980-2012 data.

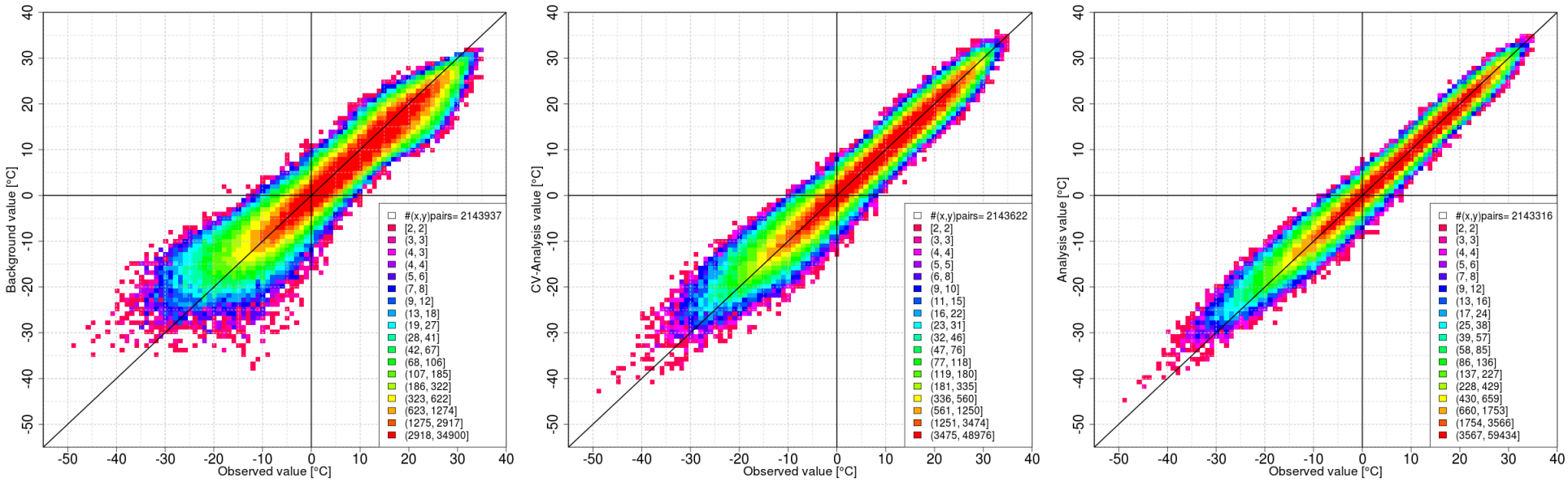


Figure 22: TAXRR-Nor. Density plots of: background *vs* observations (left); CV-analysis *vs* observations (central); analysis *vs* observations (right). Based on 1980-2012 data. The shading indicates the number of pairs (or collocations) within each 1°C by 1°C cell and the total number of collocations is reported in the legend.

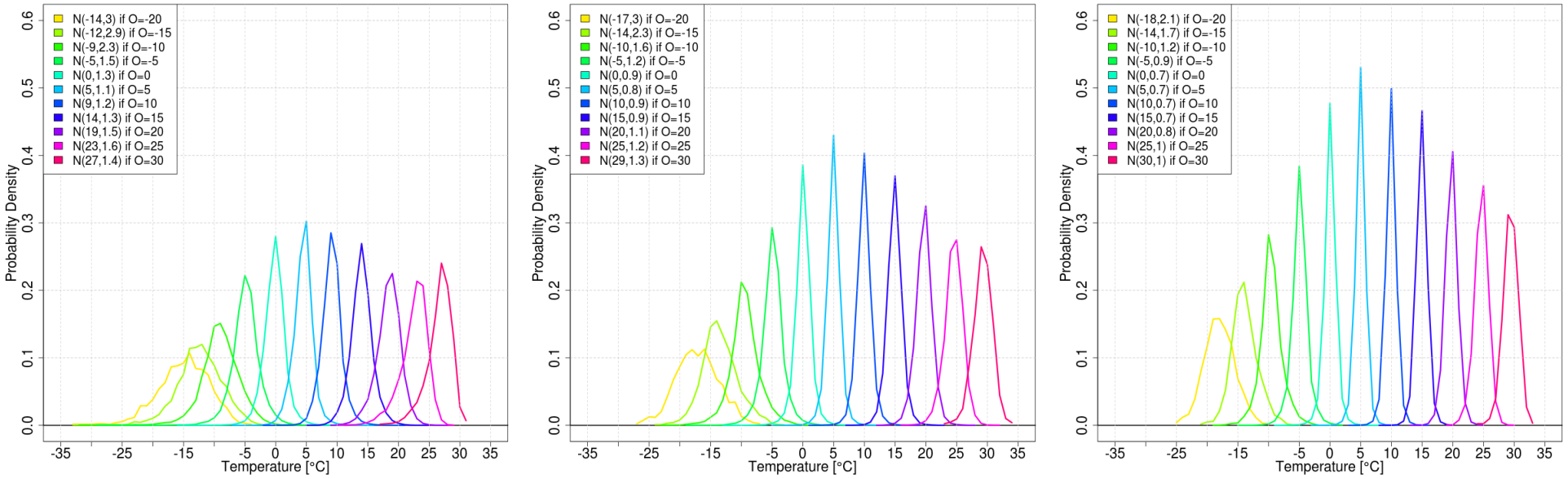


Figure 23: TAXRR-Nor. Empirical conditional probability density function given the observed value O for: background (left); CV-analysis (central); analysis (right). Based on 1980-2012 data.

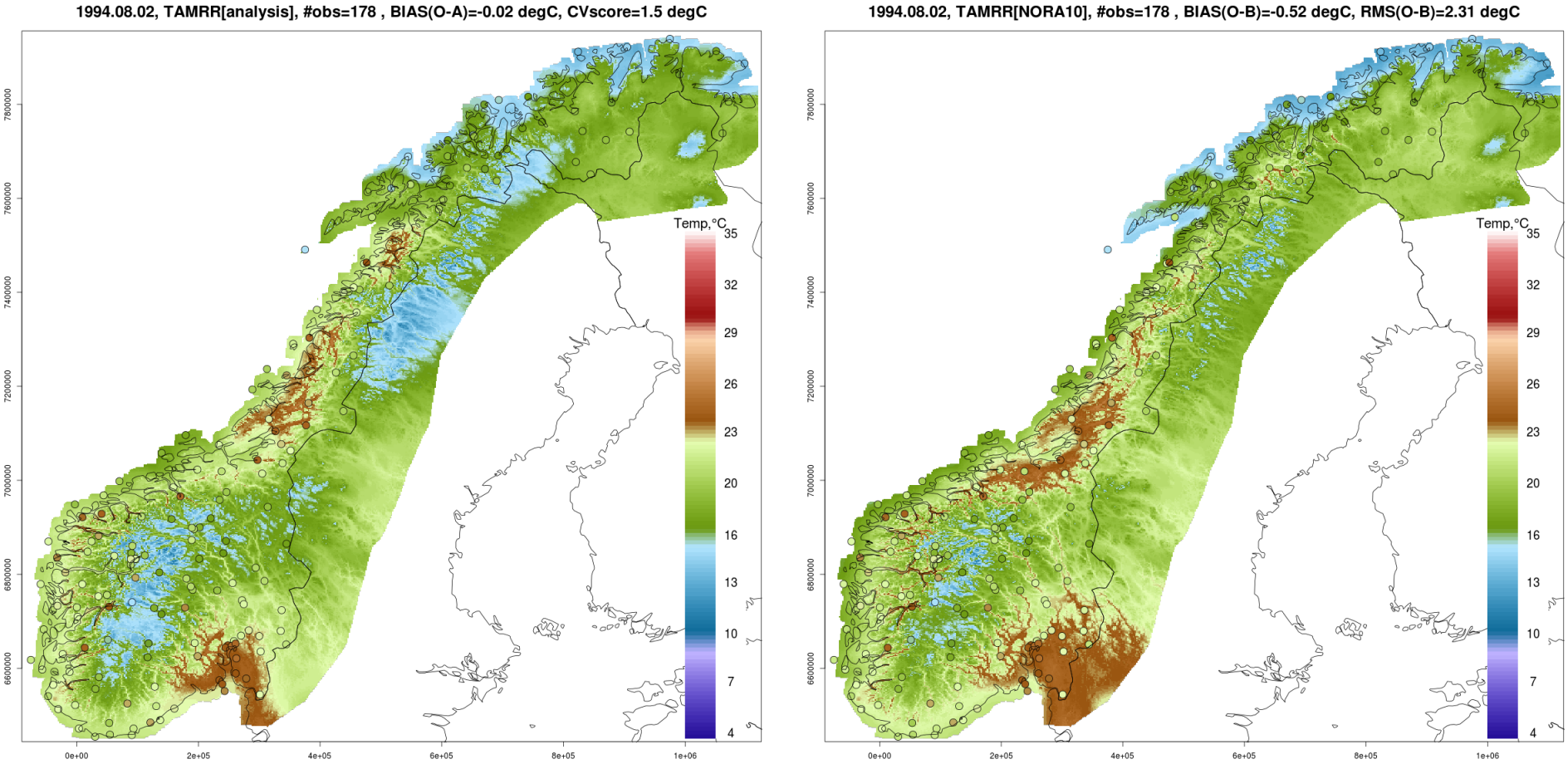


Figure 24: TAMRR-Nor. 2 August 1994. Daily (06-06 UTC) average of the two-meter temperature field: analysis (left); model background based on NORA10 (right).

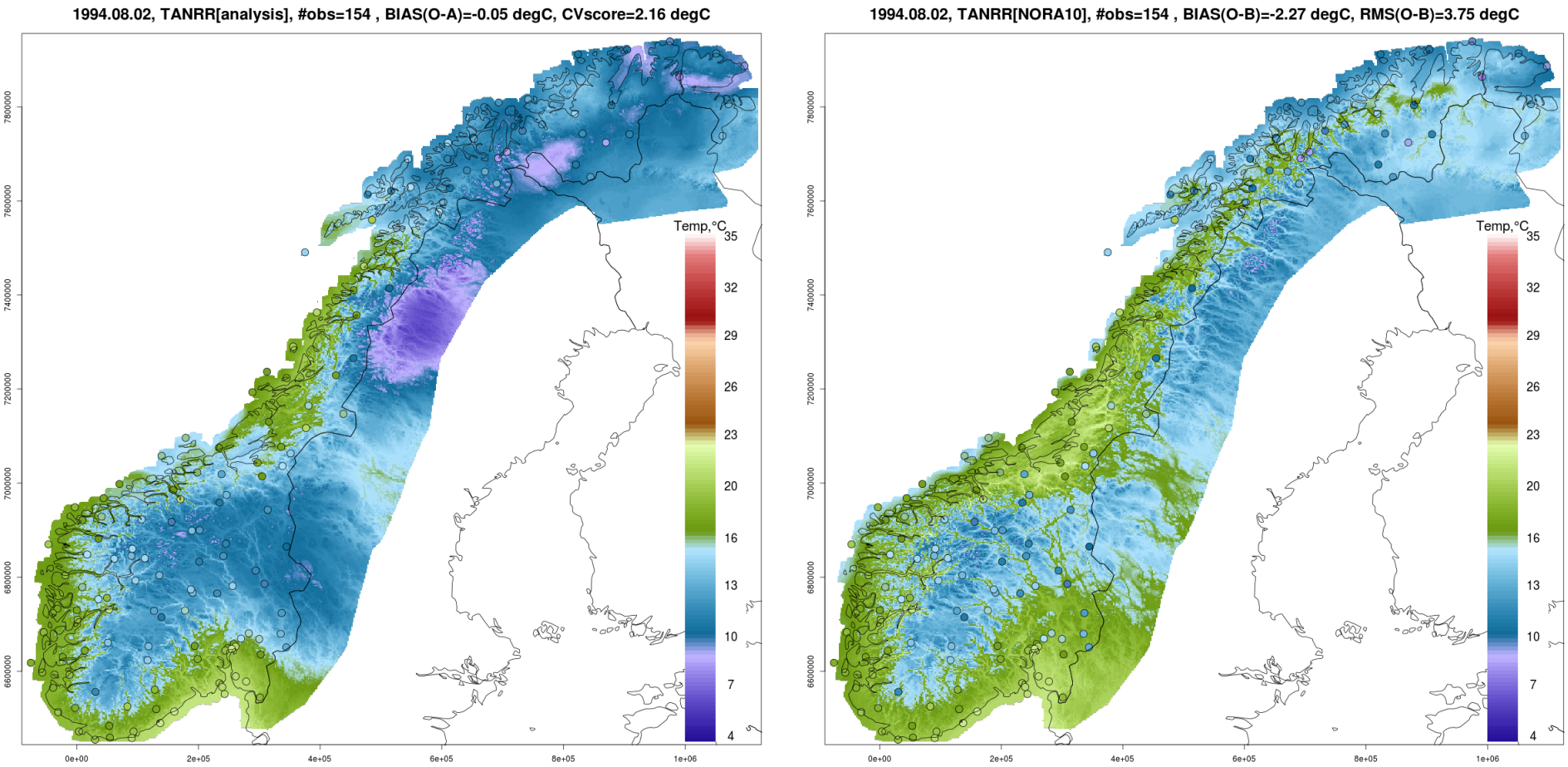


Figure 25: TANRR-Nor. 2 August 1994. Daily (06-06 UTC) minimum two-meter temperature field: analysis (left); model background based on NORA10 (right).

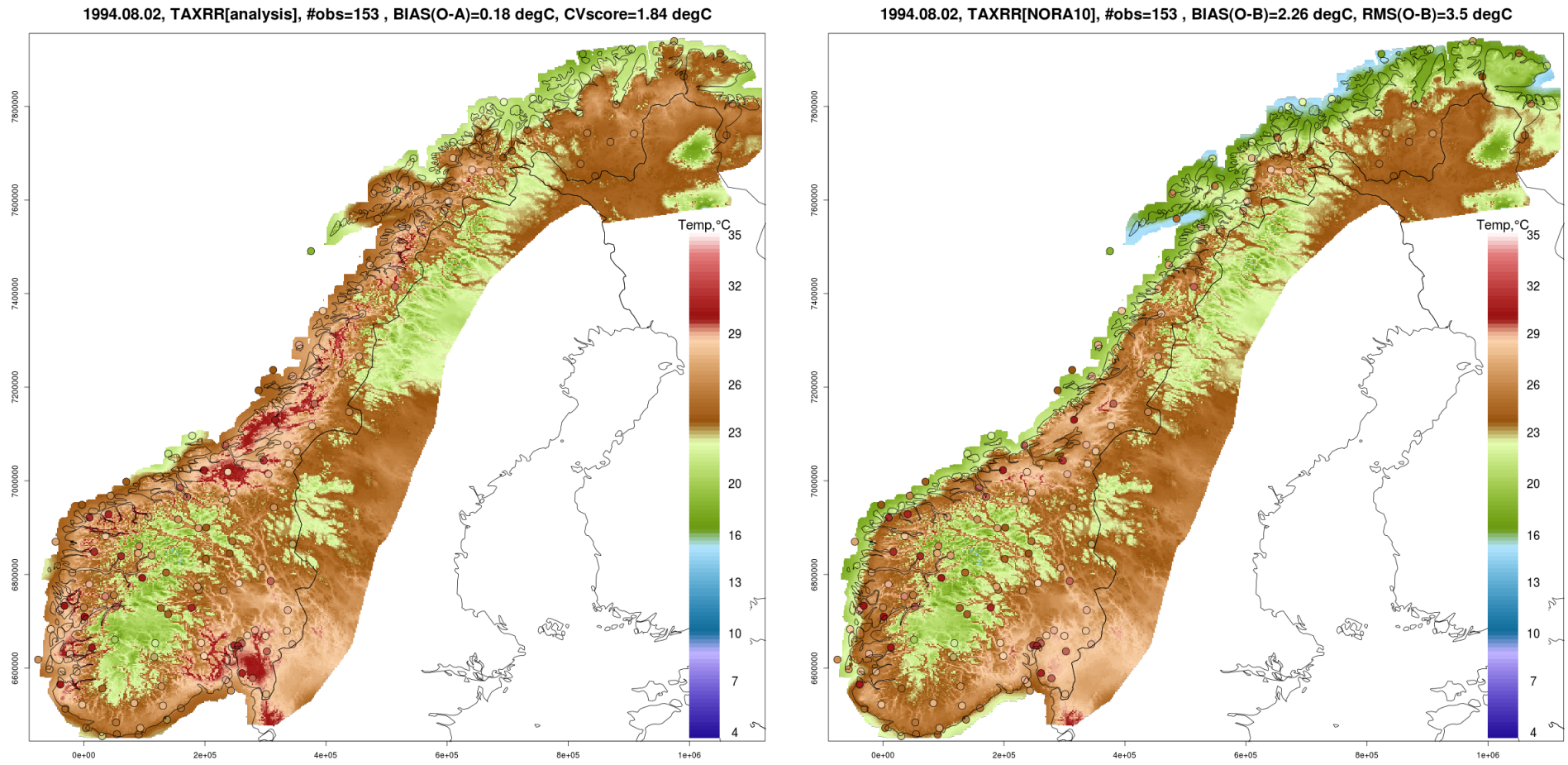


Figure 26: TAXRR-Nor. 2 August 1994. Daily (06-06 UTC) maximum two-meter temperature field: analysis (left); model background based on NORA10 (right).

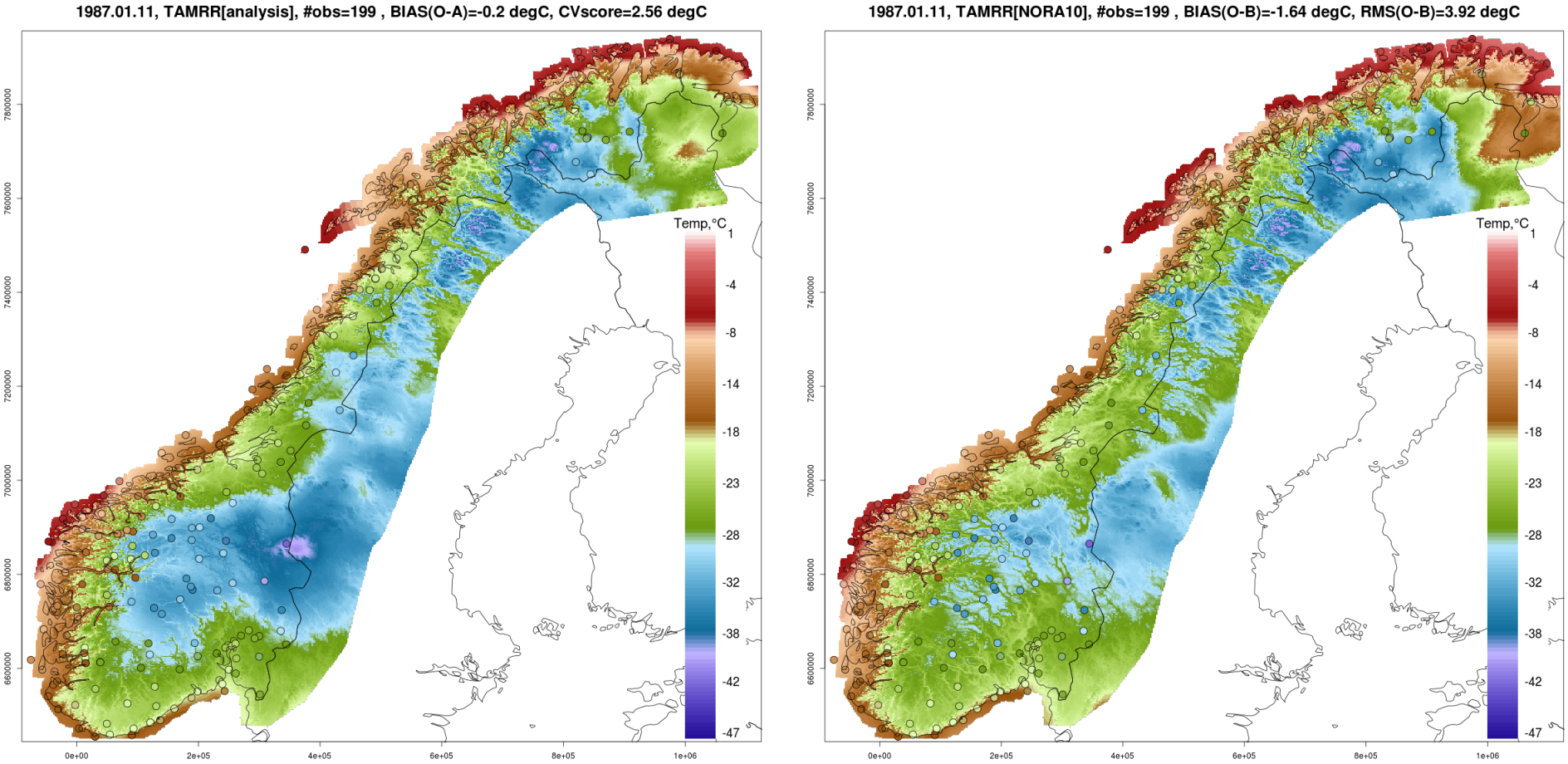


Figure 27: TAMRR-Nor. 11 January 1987. Daily (06-06 UTC) average of the two-meter temperature field: analysis (left); model background based on NORA10 (right).

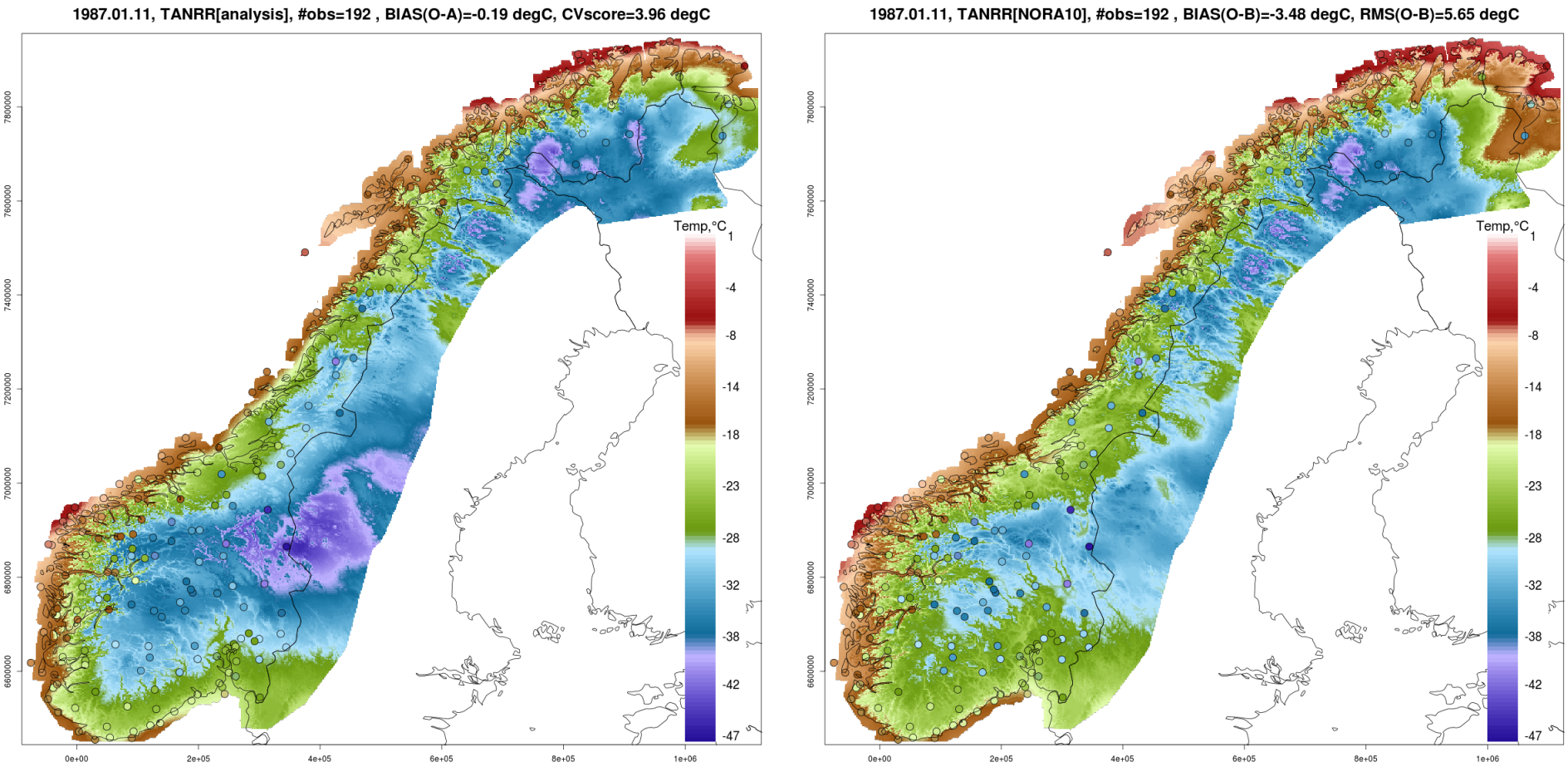


Figure 28: TANRR-Nor. 11 January 1987. Daily (06-06 UTC) minimum two-meter temperature field: analysis (left); model background based on NORA10 (right).

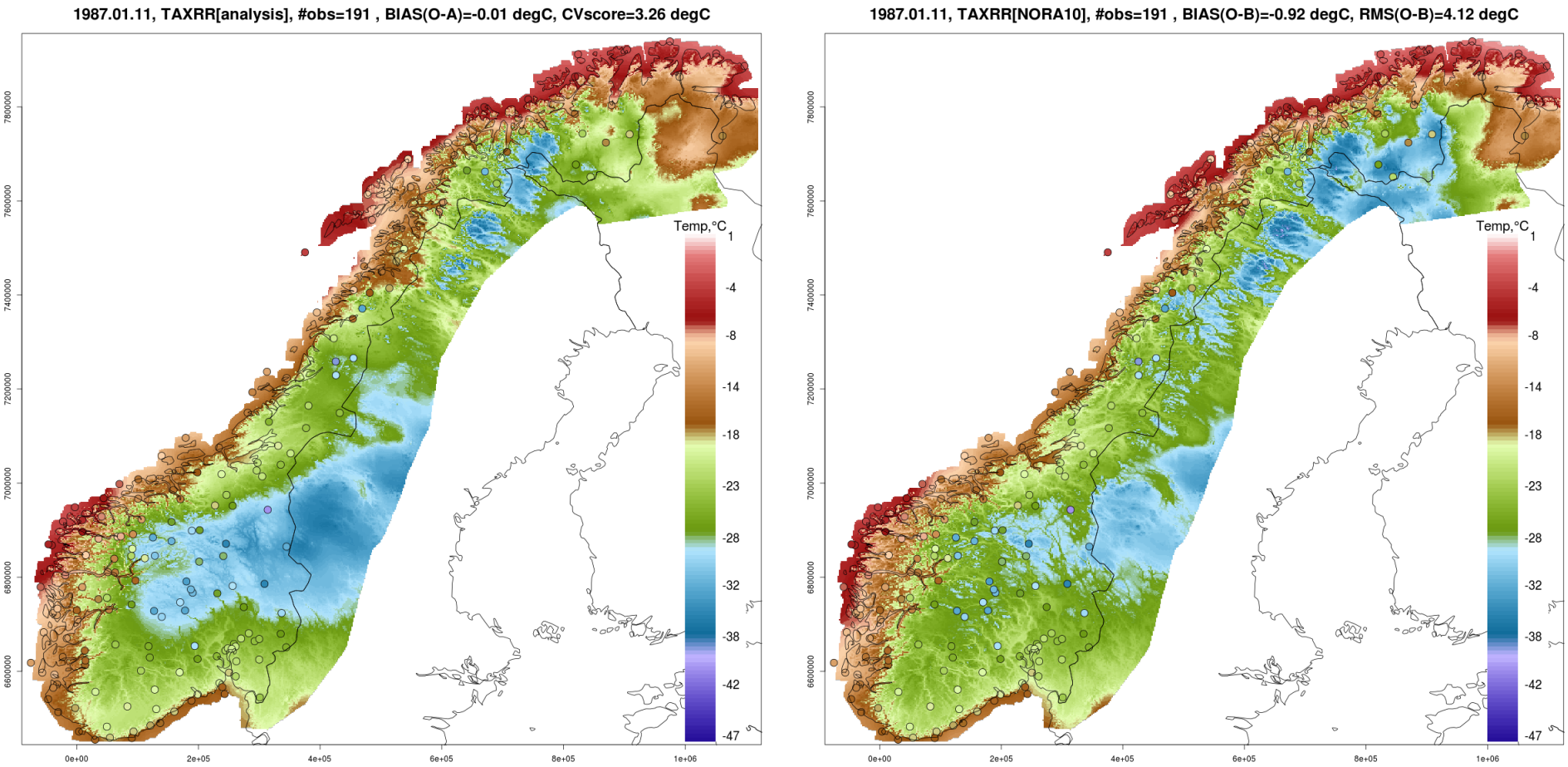


Figure 29: TAXRR-Nor. 11 January 1987. Daily (06-06 UTC) maximum two-meter temperature field: analysis (left); model background based on NORA10 (right).

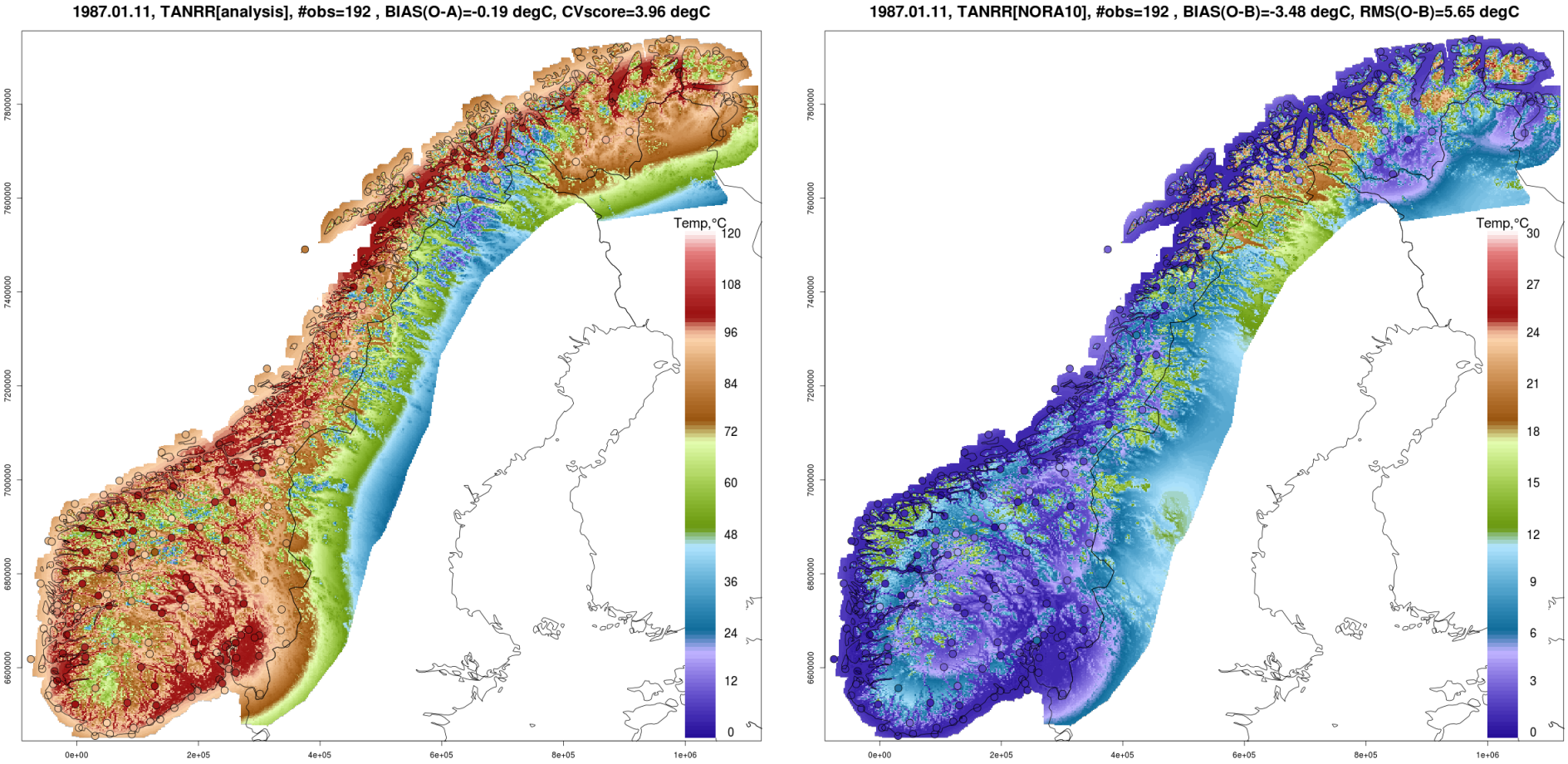


Figure 30: TANRR-Nor. 11 January 1987. IDI ensemble mean (left); IDI ensemble spread (i.e. standard deviation, right).

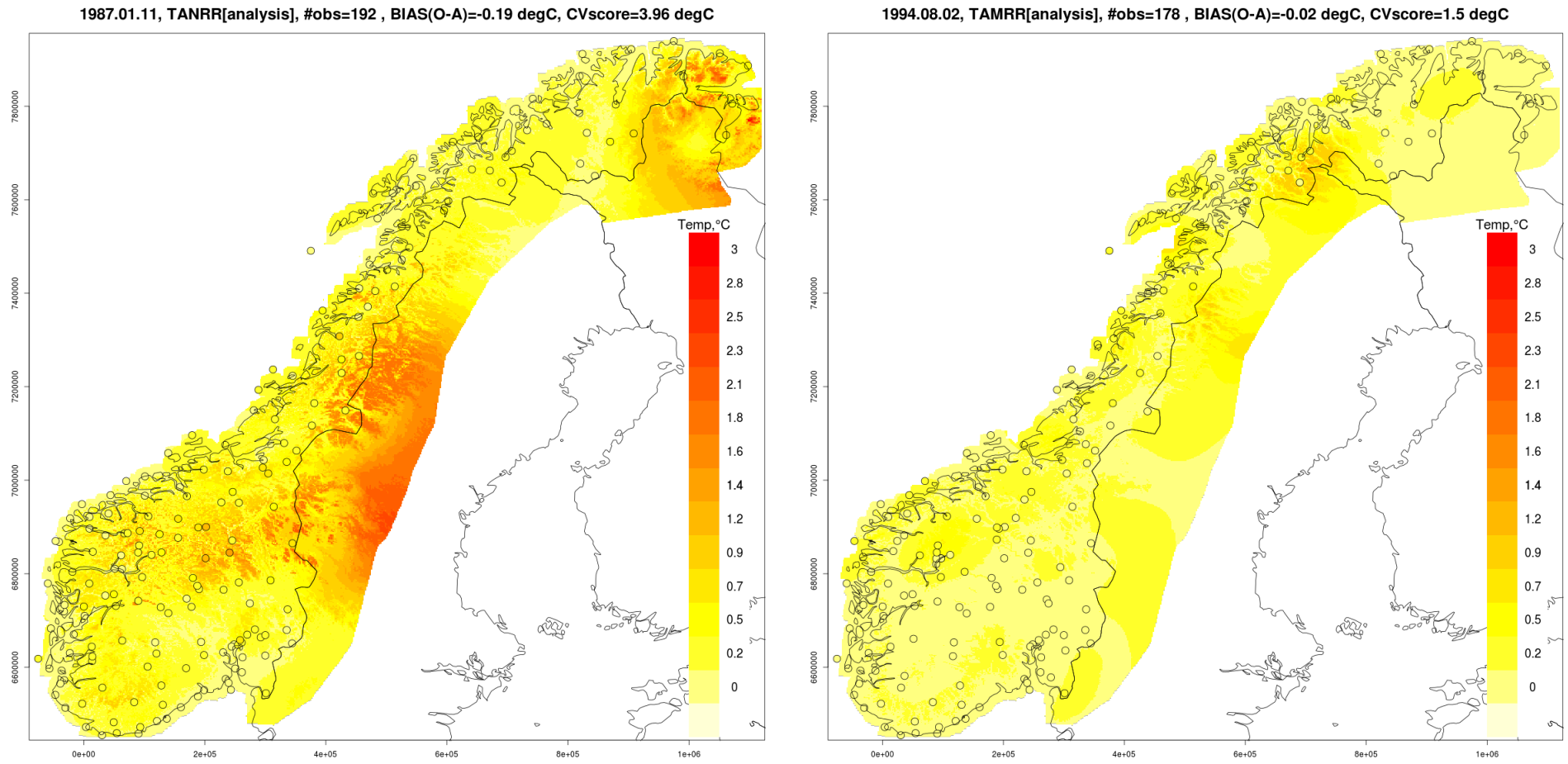


Figure 31: analysis ensemble spread (i.e. standard deviation) for: TANRR-Nor 11 January 1987 (left); TAMRR-Nor 2 August 1994 (right).

References

Daley, R. (1991), *Atmospheric Data Analysis*, 457 pp., Cambridge University Press.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli (2005), Diagnosis of observation, background and analysis-error statistics in observation space, *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3385–3396.

Gandin, L. S., and R. Hardin (1965), *Objective analysis of meteorological fields*, vol. 242, Israel program for scientific translations Jerusalem.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc (1997), Unified notation for data assimilation: operational, sequential and variational, *Practice*, *75*(1B), 181–189.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, 341 pp., Cambridge University Press.

Lorenc, A. (1986), Analysis methods for numerical weather prediction, *Quart. J. Roy. Meteorol. Soc.*, *112*, 1177–1194.

Lussana, C., F. Uboldi, and M. R. Salvati (2010), A spatial consistency test for surface observations from mesoscale meteorological networks, *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 1075–1088.

Lussana, C., C. A. Elo, and S. S. Rønning (2016a), Klinogrid rr-rad: combination of radar-derived precipitation fields and raingauge observations.

Lussana, C., O. E. Tveito, and F. Uboldi (2016b), senorge v2.0: an observational gridded dataset of temperature for norway.

Reistad, M., Ø. Breivik, H. Haakenstad, O. J. Aarnes, B. R. Furevik, and J.-R. Bidlot (2011), A high-resolution hindcast of wind and waves for the north sea, the norwegian sea, and the barents sea, *Journal of Geophysical Research: Oceans*, *116*(C5).

Simmons, A., P. Berrisford, D. Dee, H. Hersbach, S. Hirahara, and J.-N. Thépaut (2016), A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets, *Quarterly Journal of the Royal Meteorological Society*.

Tveito, O., E. Førland, R. Heino, I. Hanssen-Bauer, H. Alexandersson, B. Dahlström, A. Drebs, C. Kern-Hansen, T. Jónsson, E. Vaarby Laursen, et al. (2000), Nordic temperature maps, *DNMI report*, *9*(00).

Uboldi, F., C. Lussana, and M. Salvati (2008), Three-dimensional spatial interpolation of surface meteorological observations from high-resolution local networks, *Meteorological Applications*, *15*(3), 331–345.