Towards High Resolution Climate Reconstruction Using an Off-line Data Assimilation and COSMO-CLM 5.00 Model

Bijan Fallah[1], Emmanuele Russo[1], Walter Acevedo[2], Achille Mauri[3], Nico Becker[1], and Ulrich Cubasch[1]

[1]Institute of Meteorology, Freie Universität Berlin,Carl-Heinrich-Becker Weg 6-10, 12165 Berlin, Germany
[2]Deutscher Wetterdienst, Offenbach, Frankfurt am Main, Germany
[3]European Commission Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra (VA), Italy.

Correspondence to: Bijan Fallah (bijan.fallah@met.fu-berlin.de)

Abstract. Data assimilation (DA) methods have been used recently to constrain the climate model forecasts by paleo-proxy records. Both DA and climate models are computationally very expensive. Moreover, in paleo-DA, the time step of consequence observations is usually too long for a dynamical model to follow the previous analysis state and the chaotic behavior of the model becomes dominant. The majority of the recent paleoclimate studies using DA have performed low or intermediate resolution global simulations along with an "off-line" DA approach. In an "off-line" DA, the re-initialisation cycle is completely removed after the assimilation step. In this paper, we design a computationally affordable DA to assimilate yearly pseudo and real observations into an ensemble of COSMO-CLM high resolution regional climate model (RCM) simulations over Europe, where the ensemble members slightly differ in boundary and initial conditions. Within a perfect model experiment, the performance of the applied DA scheme is evaluated with respect to its sensitivity to the noise levels of pseudo-observations. It was observed that the injected bias in the pseudo-observations does linearly impact the DA skill. Such experiments can serve as a tool for selection of proxy records, which can potentially reduce the state estimation error when they are assimilated. Additionally, the sensitivity of the COSMO-CLM to the boundary conditions is addressed. The geographical regions, where the model exhibits high internal variability are identified. Two sets of experiments are conducted by averaging the observations over summer and winter. Furthermore, the effect of the spurious correlations within the observation space is studied and a optimal correlation radius, within which the observations are assumed to be correlated, is detected. Finally, the pollen-based reconstructed quantities at the mid-Holocene are assimilated into the RCM and the performance is evaluated against a test data-set. We conclude that the DA approach is a promising tool for creating high resolution yearly analysis quantities. The affordable DA method can be applied to efficiently improve the climate field reconstruction efforts by combining high resolution paleo-climate simulations and the available proxy records.

1 Introduction

It is now well known that many of the long-term processes (millennial scale) in the climate system have a large impact on predicting the climate even for the shorter time scales (decadal and yearly scales) (Evans et al., 2013; Acevedo et al., 2015; Latif et al., 2016; Steiger and Smerdon, 2017). The improvement of future predictions of the climate models, depends largely on the understanding of such processes. However, cutting-edge science appears to be immature in describing processes, which take place on time scales beyond the yearly cycle of the climate system (Latif et al., 2016). One of the main challenges is the lack of information for longer time-periods, with observational data-sets usually covering less than the recent century. Alternative information of past climate behavior whose time scales expand beyond the available observational data-sets are indeed required. Two methods are commonly employed for reconstructing the climate prior to the instrumental records: paleo-proxy reconstructions and climate models (Hakim et al., 2013).

Climate proxy archives (tree ring, coral, sediment and glacial) are examples of indirect climate observations, which suffer from several ill structural conditions (Acevedo et al., 2015; Jones and Mann, 2004). The data recording processes involved in such archives are very complex, encompassing physical, biological and chemical processes (Evans et al., 2013). The temporal resolution of climate proxies does not exceed seasonal time scales. Differentiating the climate and human impact on proxies is also a challenging work. Furthermore, inverting proxy records into climate information is traditionally done in the framework of statistical modeling and multivariate linear regression techniques dominate this area (Acevedo et al., 2015). Using more sophisticated statistical models is problematic, because the overlapping time span between the instrumental (weather station observations) and the proxy records becomes too short to train the statistical models.

Climate models may serve as an alternative method for the investigation of the long-term paleoclimate variability. They create dynamically consistent climate states by using numerical methods (Goosse, 2016). However, their reconstructed states are very sensitive to the initial conditions and to the imposed forcings, as well as to the parametrisation schemes used for the representation of sub-grid scale processes (small-scale processes which are not explicitly resolved by the model). An improvement of the spatial resolution of climate models is thought to be of crucial importance for the study of past climate changes, in particular when comparing their results against proxy data which are highly affected by local-scale processes (Renssen et al., 2001; Bonfils et al., 2004; Masson et al., 1999; Fallah and Cubasch, 2015; Russo and Cubasch, 2016; Russo, 2016). As a consequence, along with GCMs, high resolution regional climate models (RCMs) are recently being applied in paleoclimate studies. Several recent studies have applied a time-slice climate simulation method (Kaspar and Cubasch, 2008) to downscale dynamically the global paleo-climate simulations with a higher resolution (Prömmel et al., 2013; Fallah et al., 2016; Russo and Cubasch, 2016). Simulation of the climate of the past using RCMs is a challenging approach due to their high computational costs and dependency of such models on the driving general circulation model (GCM). Given the computational costs of RCMs, previous studies, could conduct only a single time-slice climate simulation. An individual model simulation may not provide a sophisticated measure of the uncertainty in the climate state.

In addition to the two above-mentioned methodologies, a novel and appealing technique for the reconstruction of the climate of the past is Data Assimilation (DA). According to Talagrand (1997), assimilation of the observations is the process through

which the state of the atmospheric or oceanic flow is estimated by using the available observations and the physical laws which govern its evolution, presented in the form of a numerical model. DA has recently emerged as a powerful tool for paleoclimate studies, mathematically blending together the information from proxy records and climate models (Evensen, 2003; Hughes et al., 2010; Brönnimann, 2011; Bhend et al., 2012; Dee et al., 2016; Hakim et al., 2013; Steiger et al., 2014; Matsikaris et al., 2015, 2016; Hakim et al., 2016; Acevedo et al., 2017; Okazaki and Yoshimura, 2017; Perkins and Hakim, 2017). For a review of the DA techniques which are applied in paleoclimate studies, we refer to the works of Acevedo et al. (2015, 2017) and Steiger and Hakim (2015). One limiting aspect of classical DA methods is the fact that the realistic climate models may not have predictive skills longer than several months (Acevedo et al., 2017; Perkins and Hakim, 2017). The models lose their forecasting skills shortly after initialization and evolve freely until the next step where the observations are available. Considering the complexity of implementation of DA methods into climate models, their high computational expenses and the short forecasting horizon of the models, an alternative DA approach ("off-line" or "no-cycling") was applied by several scholars (Franke et al., 2017; Okazaki and Yoshimura, 2017; Dee et al., 2016; Acevedo et al., 2017; Steiger and Hakim, 2015; Chen et al., 2015; Steiger et al., 2014). In an offline DA, the model initialization from the analysis has been abandoned and the background ensembles are derived from precomputed simulations (Okazaki and Yoshimura, 2017). In this paper we conduct a test by assimilating yearly averaged pesudo and real observations into an ensemble of RCM simulations over Europe via an offline DA approach. The main purpose of our study is to contribute to the simulation of the climate in general, and in particular, to the paleo-DA efforts, by addressing following points:

(i) How is the geographic distribution of model bias in an ensemble of regional climate simulations (members differ slightly in boundary and initial conditions)?

(ii) What is the optimum radius within which the observations are correlated?

(iii) Can this particular offline DA approach be used to constrain the regional climate simulations in time and space?

(iv) What is the range of observation errors within which the observations reduce the model's bias via DA?

This paper is structured as follows. Section 2 starts with introducing the offline DA basics and the experiment design. Furthermore, the concept of the perfect model experiment and the metrics, which measure the performance of this method are described. In Sec.3 we present our results of constrained model simulations. We discuss and summarize our work in Sec.4.

2 Data and Methods

2.1 Optimal Interpolation Basics

Prior to describing the experimental design, we give a brief review of the optimal interpolation (OI) method (for the full review see Barth et al. (2008a, b)). OI or "objective analysis" or "kriging" is one of the most commonly used and simple DA methods applied since 1970s (Barth et al., 2008b). The unknown state of the climate (X) is a vector, which has to be estimated

conditioned on the available observations (Y). Given the state vector X, the state on observations' location is obtained by an interpolation method (here nearest neighbor). Here, this operation is noted as matrix H (observation operator) and consequently the state at the observations' location is defined as HX. If the observations are in the form of proxy data, the operator H will bring the model state to the proxy state by means of a proxy forward model. The ultimate goal of the OI is to estimate the real state of the climate, which is labeled as "Nature" run (X^N). After any DA cycle, the so-called analysis (X^A) is calculated given the observation (Y) and the background X^B (first guess or forecast). The background and observation can be written as:

$$X^B = X^N + \eta^B \tag{1}$$

$$Y = HX^N + \varepsilon \tag{2}$$

where ε and η^B denote the observation and background errors, respectively. In the applied OI scheme here, it is assumed that the background and observations are unbiased:

$$E[\eta^B] = 0 \tag{3}$$

$$E[\varepsilon] = 0 \tag{4}$$

Other hypotheses are that the information about the observation and background errors are known (*prior* knowledge) and they are independent:

$$E[\eta^B \eta^{B^T}] = P^B \tag{5}$$

$$E[\varepsilon \varepsilon^T] = R \tag{6}$$

$$E[\eta^B \varepsilon^T] = 0 \tag{7}$$

The OI scheme is considered as the best linear unbiased estimator (BLUE) of the nature state X^N. A BLUE has the following characteristics:

1. It is linear for Y and X^B

2. It is not biased:

$$E[X^A] = X^N \tag{8}$$

3. It has the lowest error variance (optimal error variance).

The unbiased linear equation between X^B and Y can be written as :

5   $$X^A = X^B + K(Y - HX^B) \tag{9}$$

where K is the "Kalman gain" matrix. Equation 9 can be written as :

$$\eta^A = \eta^B + K(\varepsilon - H\eta^B) = (I - KH)\eta^B + K\varepsilon \tag{10}$$

Thus the error covariance of the analysis will be:

$$P^A(K) = E[\eta^A \eta^{A^T}] = (I - KH)P^B(I - KH)^T + KRK^T \tag{11}$$

10   The trace of matrix P^A indicates the total error covariance of the analysis:

$$\begin{aligned} Trace(P^A(K)) = Trace(P^B) + Trace(KHP^BH^TK^T) \\ -2Trace(P^BH^TK^T) + Trace(KRK^T) \end{aligned} \tag{12}$$

Given that the total error variance of analysis has its minimum value, a small δK will not modify the total variance:

$$\begin{aligned} Trace(P^{Analsis}(K + \delta K)) - Trace(P^A(K)) = 0 \\ = 2Trace(KHP^BH^T\delta K^T) - 2Trace(P^BH^T\delta K^T) + 2Trace(KR\delta K^T) \\ = 2Trace([K(HP^BH^T + R)]\delta K^T) \end{aligned} \tag{13}$$

Assuming that the δK is arbitrary, the Kalman gain is :

15   $$K = P^BH^T(HP^BH^T + R)^{-1} \tag{14}$$

Finally, the error covariance of the BLUE is given by:

$$\begin{aligned} P^A = P^B - KHP^B \\ = P^B - P^BH^T(HP^BH^T + R)^{-1}HP^B \end{aligned} \tag{15}$$

The calculation of the covariance matrices for RCM are very expensive. Therefore, an ensemble of the model states are applied to approximate the mean and covariance of the forecast (Evensen, 1994). Following a stochastic approach (Hamill, 20   2006), an observation ensemble is created by adding random noise (in the observational range) to the Y. In our experiment,

we tested the impact of different observation errors on the analysis skill in more detail. The error covariance of the BLUE, P^B can be localized through the following parametrization:

$$P^B(x_1, ..., x_n, y_1, ..., y_n) = \sigma(x_1, ..., x_n)^2 exp(-\frac{(x_1 - y_1)^2}{L^2_1})... - \frac{(x_n - y_n)^2}{L^2_n})) \tag{16}$$

where the $\sigma(x_1, ..., x_n)^2$ is the error variance and L_n the correlation length (or "localization radius"). Background state on each
5   grid point may be highly correlated with observations over regions far apart. The correlation length is defined to overcome these spurious relations.

2.2   Observation System Simulation Experiment

Models contain systematic errors which may have diverse origins (dynamical core, parametrization and initialization). DA schemes are also based on simplified hypotheses and are imperfect (e.g., here the Gaussian parametrization for P^B). The
10 interaction of error sources with one another obscures tracing of the origins of such biases. These caveats are neglected by using a simplified numerical experiment called observation system simulation experiment (OSSE). The usage of OSSEs is increasing in the field of climate DA as a validation tool (Annan and Hargreaves, 2012; Bhend et al., 2012; Steiger et al., 2014; Acevedo et al., 2015; Dee et al., 2016; Acevedo et al., 2017; Okazaki and Yoshimura, 2017). Firstly, one simulation is selected as the natural state of the climate X^N or the prediction target. Then by using the output from the nature run and
15 by adding random draws from a white noise distribution, the pseudo-observations are created which are interpolated over the observations' location. The location of 500 random meteorological stations of the "ENSEMBLES daily gridded observational dataset for precipitation, temperature and sea level pressure in Europe called E-OBS" (Haylock et al., 2008) are used to create the pseudo-observation data (Fig.1). Finally, the OI scheme is applied to assimilate the pseudo-observation into the free ensemble run (an unconstrained ensemble of simulations) and the observationally constrained run X^{DA} is obtained. Here we
20 neglect the reinitialization step of the DA and draw the forecast state from precomputed free ensemble simulation. Recently such assimilation methodologies are labeled as "off-line" DA (Huntley and Hakim, 2010).

2.3   Ensemble Generation Technique

RCM simulations tend to follow the trajectory of the driving GCM, however it is known that RCM deviate from the driving GCM, both on smaller scales, which are not resolved by the GCM and on larger scales which are resolved by the GCM (Becker
25 et al., 2015). RCMs are very sensitive to the choice of the domain and usually are tuned for the selected area. Modifying the RCM's domain in size or geographic area will set new boundary conditions for the RCM. In the context of the domain manipulation, several studies analysed the impact of the domain size on regional model simulations, finding it might have a significant impact on predictive skill and climatological characteristics of the models (Larsen et al., 2013; Goswami et al., 2012; Colin et al., 2010). In addition, a shift of the model domain generates different large-scale patterns in the simulation outputs
30 (Miguez-Macho et al., 2004), which are the consequence of the "large-scale secondary circulations" in the RCM and are relative to the driving data (Becker et al., 2015). There are different popular methods for generating RCM ensembles, e.g. the use of different parameterizations (Wang et al., 2011), different driving data (Verbunt et al., 2007) or different initialization dates

(Hollweg et al., 2008). However, the uncertainty introduced by the selection of a specific model domain is rarely considered, but the usefulness of this approach is shown in Pardowitz et al. (2016) and Mazza et al. (2017). The advantage of this method is that it is easy to implement and allows a consistent model setup among the ensemble members. The multiple members are created by shifting the model domains relative to each other.

2.4 Experimental Design

The CCLM model version cosmo_131108_5.00_clm8 (Asharaf et al., 2012) is used as the dynamical model and the nearest neighbor interpolation is applied as the observation forward model. We have to note that the model-data mapping is still a limiting factor in paleo DA (Dee et al., 2016; Goosse, 2016). For a review of proxy system models we refer to the recent work of Dee et al. (2016). Here, our focus is only on the DA scheme and its usability in paleo-climate. Two sets of observations are used by time-averaging of daily values over the winter (DJF) and summer (JJA) seasons. The time average climate analysis is conducted by using the OI scheme as the DA tool. The horizontal resolution of the simulations is set to $0.44° \times 0.44°$. Our OSSE consists of two sets of simulations:

- **Nature** simulation: a 10 year long simulation over Europe driven by global atmospheric reanalysis data (6-hourly) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), the so-called ERAInterim (Dee et al., 2011) (initial and boundary conditions are taken from ERAInterim). This run will be used as the "target" state of the climate in the investigations and is labeled as the "nature" run.

- **Shifted** domain simulations: all the model parameters are set as in the nature run but simulation domains are shifted in 4 distinct directions (northeast, northwest, southeast and southwest) and with 5 different shifting values (1 to 5 grid points) from the nature domain. These shifting directions allow for having optimal different boundaries compared to the nature run (Fig. 1). Due to high computational expenses of such simulations we conducted 10-year simulations (total number of $21 \times 10 = 210$ model years). The state vector is taken from the evaluation domain (thick black box in Fig. 1).

2.5 Model's Skill Metric

The skill of the analysis (X^A) can be quantified by the root mean square error (RMSE):

$$RMSE(\langle X^A \rangle) = \left(\overline{\left(X^N - \langle X^A \rangle \right)^2} \right)^{\frac{1}{2}}, \tag{17}$$

where $\overline{}$ and $\langle \, \rangle$ denote the time and ensemble mean operator. Mean and spread of the ensemble are presented as :

$$\langle X^A \rangle = \frac{1}{N} \sum_{n=1}^{N} x_n^A, \tag{18}$$

$$\sigma(i,j) = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(\epsilon_n(i,j))^2} \tag{19}$$

where $\epsilon_n(i,j) = X_m(i,j) - \overline{\langle X \rangle}(i,j)$, N is the ensemble index and (i,j) the indices of horizontal positions. Same approach is used to calculate the skill for the free ensemble quantities.

To maintain clarity we focus only on the temperature at 2 meter (T2M) variable. For evaluation, the relaxation zone where the boundary data is relaxed in the RCM (here 20 grid points next to the lateral boundaries) is removed. The sea surface temperature in our CCLM simulation is interpolated from the driving model (here ERAInterim) and not calculated by the model dynamics and the spread of the ensemble for T2M is zero over oceans. Therefore the RMSEs over oceans are masked out from the analysis and values over land are only shown.

3   Results

3.1   Unconstrained Ensemble Runs

Figure 2 shows the seasonal mean of RMSE and spread of the ensemble for T2M over the evaluation domain. A slight change in the boundary conditions (shifting of the domain) of the model leads to large RMSEs (up to $1K$ in seasonal means over 10 years) in the forecasted quantities (Fig.2). Maximum RMSE values are located on the regions where the spread of the ensemble is also large. An interesting feature of the RMSE pattern is the accumulation of the errors in the center and Northeast of the domain for both the winter and summer seasons. West and North of the domain is largely dominated by the ocean where the temperature quantities are forced by the reanalysis data. There is small difference between the ensemble members on these areas and the maximum error values are mostly located on land. On the other hand, the maximum lateral inflow occurs on the West and the Northwest boundaries. These features have to be considered cautiously when conducting long-term climate simulations using CCLM. The long seasonal time averaging filter relatively dampens the RMSE and the spread of the ensemble. The errors will increase drastically for daily quantities. One of our main motivations in this paper is to test the usage of OI in calculation of analysis quantities by assimilating the pseudo and real observations in the free ensemble simulations.

3.2   Constrained Ensemble Runs

Prior to conducting the analysis calculation, we search for an optimal correlation length (L) for the covariance localization, which minimizes the RMSE of the analysis. The mean RMSE over the evaluation domain is calculated for different values of L ($0.1° \leq L \leq 6°$). The RMSE values show a minimum at correlation length of $1.7°$ ($\sim 190\ Km$) for summer (JJA) and $2.1°$ ($\sim 230\ Km$) for winter (DJF) (Fig.3). For a long-term time averaging e.g., yearly or decadal periods, the correlation length will increase accordingly (Chen et al., 2015). The increase rate of the RMSE with respect to the correlation length is larger during the summer than in winter. The error reduction of the DA is influenced by the shape of the white noise added to create the pseudo-observation data. Here, we assess the noise levels by signal-to-noise ratio (SNR), which is the ratio of the variance of

clean observation to the variance of added white noise. Figure 4, shows the averaged seasonal near-surface temperature RMSE for the analysis of the ensemble run. The RMSE values are reduced linearly with increasing SNR. In the study of Acevedo et al. (2017), the RMSE values did not change significantly between 2 and 11 SNR values. One possible reason for such a behavior might be due to the regional domain of RCM over Europe, which is dominated by local effects. On the other hand, $\sim 70\%$ of the GCM domain in Acevedo et al. (2017) was covered by the oceans where no proxy was assimilated. In our experiment, the field-mean filter does not flatten the RMSE reduction rate for larger SNR values over Europe. For the rest of the paper the results are shown with $SNR = 3$.

Figures 5 and 6 show the error reduction of analysis for summer and winter. Assimilation of 500 pseudo-observations has significantly reduced the mean of errors for both seasons. However, the spread of the errors is more highly reduced in summer than in winter. The regions with the largest error reduction are located at highly populated areas by pseudo-observations (e.g., Germany, Sweden and Denmark). One interesting feature is the error reduction over the South of the domain where only 3 observations are available. Usually, in an offline DA the observational information would not be accumulated over time and could not be conveyed to unobserved regions (Acevedo, 2015). There exist two outliers in the ensemble spread during the winter, who contribute to the larger error spread during the winter (Fig. 6.b). The DA, has successfully detected these two outliers and reduced the errors significantly (Fig. 6.b).

3.3 Long Ensemble Runs

As defined by world meteorological organization (WMO) the climate is explained by averaging the weather state for a period of at least 30 years. Therefore, we conducted a new set of 5 36-year long simulations (one nature and 4 shifted runs). The computational costs of RCM was the only limiting factor to choose this number of members ($5 \times 36 = 180$ years of model run). The members were perturbed by shifting the domain 4 grid points to the Northeast, Northwest, Southwest and Southeast. The 36-year ensemble spread and RMSE show similar patterns as in the previous experiment using 20 members (Fig. 7). The analysis quantities indicate a significant error reduction both in median and spread of the ensemble (Fig. 8 and Fig. 9). Figure 10 illustrates time evolution of field-mean RMSEs for the free ensemble and analysis quantities. There exists a linear trend in free runs' RMSE. However, the linear trend is removed in the RMSE of the analysis for the summer. This feature was previously observed in an offline DA experiment using an intermediate complexity model and the EnKF approach (Acevedo et al., 2017). Furthermore, the results show that the analysis RMSEs are significantly reduced compared to the free ensemble run.

3.4 Random background selection

Following the methodology of Steiger et al. (2014) and Dee et al. (2016) for background selection, using a random climatology from a large climate simulation pool instead of an ensemble simulation at the observation instant, might remove the trend previously observed in the RMSE trends for winter. Under this set-up, the RMSE reduction shows a coherent pattern (Fig. 11.a) and the uprising trend in RMSE is removed (Fig. 11.b). However, the analysis RMSE mean is in the range of the background state. In our experiment, background itself had acceptable skills over large portions of the domain. Using random states affects

the model skills, for example over western part of the domain(i.e., Spain, Portugal, France and Morocco). We conclude that using random states would be beneficial if the model had no significant skills at any region of its domain. Such significant skills of the background might be a characteristics of this particular RCM, which is tuned for Europe.

3.5 Application to a paleo study: the case of summer temperatures over Europe at the mid-Holocene

A test with the real data will shed light on the applied DA efficiency. Therefore, we design additional experiments using real proxies and precomputed COSMO-CLM simulations at mid-Holocene. Here, we briefly explain the method and present the results for a single summer-time (JJA) time-slice of 6,000 years before present (6KBP). We use the pollen-based temperature reconstructions of Mauri et al. (2015) as observation and the model simulations of Russo and Cubasch (2016) as the background in our DA method. 20% of the proxy data is kept as the test data (randomly selected), and the rest 80% as the training data-set. The proxy data does not have a regular timing. Two approaches are applied for averaging the proxy data:

a) averaging with respect to their distance to the target year (6KBP): a time window centered on the target year, (e.g., reference time $\pm$ 500 years) is chosen and the values as well as their standard errors are weighted by their time distance to the target year. The weights are chosen from a normal distribution with standard deviation of 100. Totally 3 weighting time-spans are defined (5 bins). Figure 12 shows the weights assigned to each time interval with respect to the reference time. The data falling in each particular time-bin are weighted equally.

b) averaging with respect to their uncertainties provided as standard error by Mauri et al. (2015) at the recorded time: all reconstructions within the time window of reference year $\pm$500 years are chosen. The weighted arithmetic mean of the temperatures and its standard errors are used to calculate the time-slice values (6KBP). Each proxy is weighted first by its standard error.

Then the weighted mean is calculated by:

$$\bar{x} = \frac{\sum_{i=1}^{n}(x_i \sigma_i^{-2})}{\sum_{i=1}^{n} \sigma_i^{-2}} \tag{20}$$

The uncertainty of the weighted mean is given then by:

$$\sigma_{\bar{x}}^2 = \frac{1}{\sum_{i=1}^{n} \sigma_i^{-2}} \tag{21}$$

Where the σ is the standard error. Figure 13 shows the schematic of weighting of observations with respect to their standard errors.

Finally, for the model, the 25-year time-average is assigned as the expected value and the standard deviation from the mean as uncertainty measure. Figure 14 shows the schematic of the approach. Please note that here, a single model run of 25 years is used as the background. We assume that each year of the model simulation could serve as an ensemble member for the target time-slice (6KBP). Therefore, the analysis is done for a single step using the background information of the 25 years. Figure 15.a shows the analysis results for summer T2m temperature anomalies from the pre-industrial time (6KBP – 0.2KBP) over Europe, obtained by the application of approach (a) for averaging the proxy data. The testing proxy data are represented by circles, while their standard error is represented as superimposed squares. The analysis and the proxy show a good agreement

especially for proxy records with low standard error. In contrast, the model forecast (without assimilation) and the proxies (Fig. 15.b) show little agreements. The assimilated reconstructions are shown in Figure 16. The positive anomalous region over Romanian in the analysis is due to the cluster of proxy data with low standard error over this region. By applying method (b) for averaging, the analysis pattern is very similar to the previous method (Fig. 17 and Fig. 18). Overall, we conclude that the proposed DA approach is contributing to the error reduction in the analysis values where the pure model outputs might not capture the local patterns.

We have to mention that the analysis presented here is based on the combination of the proxy reconstructions and the climate model and any interpretation of the patterns should take into account the uncertainties provided by these two sources of information. Several reasons have driven our decision to apply the proposed DA method to this particular paleoclimate case study. Among these, the most important one was the attempt to contribute to the reconstruction of summer temperatures over the Mediterranean region at the mid-Holocene. This has been the subject of a long-standing debate within the paleoclimate community: while, on the one hand, climate model simulations (Braconnot et al., 2007; Fischer and Jungclaus, 2011; Bonfils et al., 2004; Russo and Cubasch, 2016) and some reconstructions based on specific proxy types (such as chyronomids Samartin et al. (2017)) have shown that the area was characterized by warmer summer conditions with respect to present days, on the other hand, climate reconstructions based on other proxies, in particular the ones based on pollen data (Cheddadi et al., 1996; Davis et al., 2003; Mauri et al., 2015), have given support to the idea of opposite, colder conditions over the area.

This dualism seemed to be finally solved in a recent study by Samartin et al. (2017), in which they showed that summer temperatures reconstructed from chyronomids, over northern Italy, were higher at the mid-Holocene when compared to present times. This pattern were also captured by other proxy records from different parts of the Mediterranean area and the results of climate models. Indeed, they criticized temperatures reconstructions based on pollen-data for this region, due to the fact that Mediterranean vegetation is mainly controlled by effective precipitation rather than temperatures. Our results are in agreement with Samartin et al. (2017) over northern Italy and the Alpine region despite the presence of heterogeneous colder conditions over southern Europe. Indeed, we believe that this complicated puzzle is not solved yet and more efforts should be put into the joint cooperation between different proxy experts and between them and climate modelers. We emphasize the fact that, data assimilation could be considered as very useful tool for syntheses of different climate records and physical representation of the climate system. This could offer a more reliable picture than individual proxy data-sets or a single climate simulation.

4 Conclusions and discussions

Using a computationally fast DA approach, we assimilated pseudo and real observations within an ensemble of precomputed RCM simulations. The ensemble is created by slightly perturbing the boundary and initial conditions of its members via the domain shifting method. In the framework of a perfect model experiment the performance of free ensemble and analysis quantities is evaluated. Such experiments facilitate the estimation of observation, background and analysis error. In the first set of experiments, we conducted an ensemble of 20 simulations driven by ERAInterim for the duration of 10 years. The nearest-neighbor interpolation is applied as the observation operator plus a random white noise with known standard deviation

to create a set of pseudo-observations from the nature run. Pseudo-observations are assimilated within the ensemble of RCM runs by means of the OI approach. By conducting a set of simulations using 4 perturbed members and a nature run, we repeated the perfect model experiment for a time expansion of 36 years. This allowed us to draw conclusions on the time evolution of the DA skill for a typical climatological period (more than 30 years). In a final step, we assimilated pollen-based temperature reconstructions of the mid-Holocene precomputed RCM simulation of 25 years duration at 0.44° horizontal resolution and compared the results with a test data-set.

The comparison of ensemble mean of COSMO-CLM model outputs and the pseudo-observations shows that the model seems to be well tuned for Central Europe. A region of significant model bias for both winter and summer seasons is located over the East side of the domain. This area is located far from the ocean where the ERAInterim data is prescribed (no coupled ocean was implemented). Therefore, we speculate that the model generates more variabilities and is free to evolve over this region (answer to question (i) in Introduction). Furthermore, we iterated the DA experiment on different values of correlation length for the summer and winter to find the optimal correlation length quantity. The optimum radius of correlation is found to be $1.7°$ ($\sim 190\ Km$) for summer (JJA) and $2.1°$ ($\sim 230\ Km$) for winter (answer to question (ii) in Introduction). Afterward, we showed that the skill of OI is linearly influenced by the SNRs used to create the noisy observations (answer to question (iv) in Introduction).

Our experiments showed that the ensemble OI is useful for conducting analysis of seasonally averaged quantities (answer to question (iii) in Introduction). Despite inhomogeneity of observation distribution over the domain, the analysis presents error reduction over most of the domain. For a small ensemble with longer integration period of 36 years, the analysis significantly outperforms the ensemble mean. However, for the winter season, the analysis error is increasing with time and it consists the same rising trend as in the error of the ensemble mean. For the summer time the trend was removed in the analysis. This was previously observed in the study of Acevedo et al. (2017) by assimilating the summer-time tree ring width pseudo observations in an AGCM using the EnKF. However, our simulations are very short compared to the one of Acevedo et al. (2017). Using a random climatology from a large climate simulation pool as the background instead of using an ensemble simulation at the observation period, following Steiger et al. (2014) and Dee et al. (2016), has removed the winter RMSE trend. However, the mean RMSE value was not significantly reduced compared to the background. For the real world experiment, the magnitude of the analysis skills might have been influenced by using the climate of a single run for the background state. By now available long RCM runs (ie., this study, Russo and Cubasch (2016), Fallah et al. (2016) and the Coordinated Downscaling Experiment-European Domain (EURO-CORDEX)), which can serve as a large climate analog for the background state, we suggest further DA experiments to reconstruct high resolution climate fields given the time-averaged observations.

In this paper the inverse model outputs (temperature reconstructions) are exerted directly. In a real world experiment, it is recommended to use the proxy system models (PSMs) to remove the simplistic assumptions of inverse climate modeling and let the assimilation to take place at the observation space instead of model space (Dee et al., 2016; Dolman and Laepple, 2018). However, there is a huge gap in the recent knowledge of PSMs and more efforts shall be made to move the knowledge of forward proxy modeling one step forward (Acevedo et al., 2017). For example, a previous study (Acevedo et al., 2017) using

a proxy forward model showed that the error reduction for other model values like precipitation was not significant. In such set-ups, the success of the paleo-DA is largely affected by the skill of the proxy forward model.

A major culprit of our experiment is linearity assumption of the forecast model and Gaussianity of the observation and model errors. In OI the background error covariance is usually prescribed and calculated once during the entire assimilation procedure. Our experiment showed that although the spread of the ensemble is increasing slightly in time, each individual member and the ensemble mean show a similar trend. This similar behavior of the members might be due to the systematic behavior of the CCLM. We suggest a multi-model ensemble approach to account for a wider range of internal variabilities. However, conducting such experiments are prohibitively expensive with today's computational powers.

Code and data availability. The experiment code and its full description are available upon the request to the corresponding author. A test set of climate model simulations for temperature quantities is available at http://users.met.fu-berlin.de/~BijanFallah/NETCDFS_CCLM/. The OI code and its description is public at http://modb.oce.ulg.ac.be/mediawiki/index.php/Optimal_interpolation_Fortran_module_with_Octave_interface by GeoHydrodynamics and Environment Research (GHER) research group.

Author contributions. BF and WA designed the experiment. ER has conducted the Holocene RCM simulations. BF has done the scientific coding and performed the simulations. All authors analysed the results and drafted the manuscript.

Acknowledgements. The authors gratefully acknowledge German Federal Ministry of Education and Research (BMBF) as Research for Sustainability initiative (FONA); through PalMod, a German-funded (DFG) project with the general goal of analyzing a complete glacial cycle, from the Last Interglacial to the Anthropocene. The computational resources were made available by the German Climate Computing Center (DKRZ). We acknowledge GeoHydrodynamics and Environment Research (GHER) research group for making their OI code available for public. Our special thanks go to Ingo Kirchner for his helpful comments on simulation set-up as well as the CCLM community.

References

Acevedo, W.: Towards Paleoclimate Reanalysis via Ensemble Kalman Filtering, Proxy Forward Modeling and Fuzzy Logic, Ph.D. thesis, Geosim, http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000099753, 2015.

Acevedo, W., Reich, S., and Cubasch, U.: Towards the assimilation of tree-ring-width records using ensemble Kalman filtering techniques, Climate Dynamics, pp. 1–12–, http://dx.doi.org/10.1007/s00382-015-2683-1, 2015.

Acevedo, W., Fallah, B., Reich, S., and Cubasch, U.: Assimilation of Pseudo−tree−ring−width Observations into an Atmospheric General Circulation Model, Clim. Past, 13, 545–557, https://www.clim-past.net/13/545/2017/, 2017.

Annan, J. D. and Hargreaves, J. C.: Identification of climatic state with limited proxy data, Clim. Past, 8, 1141–1151, http://www.clim-past.net/8/1141/2012/, 2012.

Asharaf, S., Dobler, A., and Ahrens, B.: Soil Moisture-Precipitation Feedback Processes in the Indian Summer Monsoon Season, J. Hydrometeor., 13, 1461–1474, doi:10.1175/JHM-D-12-06.1, http://dx.doi.org/10.1175/JHM-D-12-06.1, 2012.

Barth, A., Alvera-Azcárate, A., and Weisberg, R. H.: Assimilation of high-frequency radar currents in a nested model of the West Florida Shelf, J. Geophys. Res., 113, n/a–n/a, http://dx.doi.org/10.1029/2007JC004585, 2008a.

Barth, A., Azcárate, A., A., Joassin, P., Beckers, J.-M., and Troupin, C.: Introduction to Optimal Interpolation and Variational Analysis, Tech. rep., GeoHydrodynamics and Environment Research, 2008b.

Becker, N., Ulbrich, U., and Klein, R.: Systematic large-scale secondary circulations in a regional climate model, Geophysical Research Letters, 42, 4142–4149, 2015.

Bhend, J., Franke, J., Folini, D., Wild, M., and Brönnimann, S.: An ensemble-based approach to climate reconstructions, Clim. Past, 8, 963–976, http://www.clim-past.net/8/963/2012/, 2012.

Bonfils, C., de Noblet-Ducoudré, N., Guiot, J., and Bartlein, P.: Some mechanisms of mid-Holocene climate change in Europe, inferred from comparing PMIP models to data, Climate Dynamics, 23, 79–98, https://doi.org/10.1007/s00382-004-0425-x, 2004.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale features, Clim. Past, 3, 261–277, https://www.clim-past.net/3/261/2007/, 2007.

Brönnimann, S.: Towards a paleoreanalysis?, ProClim-Flash, 1, 16, 2011.

Cheddadi, R., Yu, G., Guiot, J., Harrison, S. P., and Prentice, I. C.: The climate of Europe 6000 years ago, Climate Dynamics, 13, 1–9, https://doi.org/10.1007/s003820050148, 1996.

Chen, X., Xing, P., Luo, Y., Nie, S., Zhao, Z., Huang, J., Wang, S., and Tian, Q.: Surface temperature dataset for North America obtained by application of optimal interpolation algorithm merging tree-ring chronologies and climate model output, Theoretical and Applied Climatology, pp. 1–17–, http://dx.doi.org/10.1007/s00704-015-1634-4, 2015.

Colin, J., Déqué, M., Radu, R., and Somot, S.: Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique, Tellus A, 62, 591–604, 2010.

Davis, B., Brewer, S., Stevenson, A., and Guiot, J.: The temperature of Europe during the Holocene reconstructed from pollen data, Quaternary Science Reviews, 22, 1701–1716, http://www.sciencedirect.com/science/article/pii/S0277379103001732, 2003.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L.,

Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597, http://dx.doi.org/10.1002/qj.828, 2011.

Dee, S. G., Steiger, N. J., Emile-Geay, J., and Hakim, G. J.: On the utility of proxy system models for estimating climate states over the
5   common era, J. Adv. Model. Earth Syst., 8, 1164–1179, http://dx.doi.org/10.1002/2016MS000677, 2016.

Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, Clim. Past Discuss., 2018, 1–31, https://www.clim-past-discuss.net/cp-2018-13/, 2018.

Evans, M., Tolwinski-Ward, S., Thompson, D., and Anchukaitis, K.: Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Science Reviews, 76, 16 – 28, doi:http://dx.doi.org/10.1016/j.quascirev.2013.05.024, http://www.sciencedirect.com/
10   science/article/pii/S0277379113002011, 2013.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, doi:10.1029/94JC00572, http://dx.doi.org/10.1029/94JC00572, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, http://dx.doi.org/10.1007/s10236-003-0036-9, 2003.

15   Fallah, B. and Cubasch, U.: A comparison of model simulations of Asian mega-droughts during the past millennium with proxy reconstructions, Clim. Past, 11, 253–263, https://www.clim-past.net/11/253/2015/, 2015.

Fallah, B., Sodoudi, S., and Cubasch, U.: Westerly jet stream and past millennium climate change in Arid Central Asia simulated by COSMO-CLM model, Theoretical and Applied Climatology, 124, 1079–1088, doi:10.1007/s00704-015-1479-x, 2016.

Fischer, N. and Jungclaus, J. H.: Evolution of the seasonal temperature cycle in a transient Holocene simulation: orbital forcing and sea-ice,
20   Clim. Past, 7, 1139–1148, https://www.clim-past.net/7/1139/2011/, 2011.

Franke, J., Brönnimann, S., Bhend, J., and Brugnara, Y.: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations, Scientific Data, 4, 170 076–, http://dx.doi.org/10.1038/sdata.2017.76, 2017.

Goosse, H.: An additional step toward comprehensive paleoclimate reanalyses, J. Adv. Model. Earth Syst., 8, 1501–1503, http://dx.doi.org/10.1002/2016MS000739, 2016.

25   Goswami, P., Shivappa, H., and Goud, S.: Comparative analysis of the role of domain size, horizontal resolution and initial conditions in the simulation of tropical heavy rainfall events, Meteorological Applications, 19, 170–178, 2012.

Hakim, G., Annan, J., Broennimann, S., Crucifix, M., Edwards, T., Goosse, H., Paul, A., van der Schrier, G., and Widmann, M.: Overview of data assimilation methods, PAGES, 21, 2013.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A.: The last millennium climate
30   reanalysis project: Framework and first results, J. Geophys. Res. Atmos., 121, 6745–6764, http://dx.doi.org/10.1002/2016JD024751, 2016.

Hamill, T. M.: Ensemble-based atmospheric data assimilation, in: Predictability of Weather and Climate, edited by Palmer, T. and Hagedorn, R., Cambridge University Press, web: http://dx.doi.org/10.1017/CBO9780511617652.007, 2006.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, J. Geophys. Res., 113, n/a–n/a, http://dx.doi.org/10.1029/2008JD010201, 2008.

35   Hollweg, H., Boehm, U., Fast, I., Hennemuth, B., Keuler, K., Keup-Thiel, E., Lautenschlager, M., Legutke, S., Radtke, K., Rockel, B., Schubert, M., Will, A., Woldt, M., and Wunram, C.: Ensemble Simulations over Europe with the Regional Climate Model CLM forced with IPCC AR4 Global Scenarios, Tech. rep., M und D Technical Report 3, 2008.

Hughes, M., Guiot, J., and Ammann, C.: An emerging paradigm: Process-based climate reconstructions, PAGES news, 18, 87–89, 2010.

Huntley, H. and Hakim, G.: Assimilation of time-averaged observations in a quasi-geostrophic atmospheric jet model, 35, 995–1009–, http://dx.doi.org/10.1007/s00382-009-0714-5, 2010.

Jones, P. D. and Mann, M. E.: Climate over past millennia, Rev. Geophys., 42, RG2002, http://dx.doi.org/10.1029/2003RG000143, 2004.

Kaspar, F. and Cubasch, U.: Simulation of East African precipitation patterns with the regional climate model CLM, Meteorologische Zeitschrift, 17, 511–517, doi:10.1127/0941-2948/2008/0299, 2008.

Larsen, M. A., Thejll, P., Christensen, J. H., Refsgaard, J. C., and Jensen, K. H.: On the role of domain size and resolution in the simulations with the HIRHAM region climate model, Climate dynamics, 40, 2903–2918, 2013.

Latif, M., Claussen, M., Schulz, M., and Brücher, T., eds.: Comprehensive Earth system models of the last glacial cycle, vol. 97, Eos, doi:10.1029/2016EO059587, https://eos.org/project-updates/comprehensive-earth-system-models-of-the-last-glacial-cycle, 2016.

Masson, V., Cheddadi, R., Braconnot, P., Joussaume, S., and Texier, D.: Mid-Holocene climate in Europe: what can we infer from PMIP model-data comparisons?, Climate Dynamics, 15, 163–182, https://doi.org/10.1007/s003820050275, 1999.

Matsikaris, A., Widmann, M., and Jungclaus, J.: On-line and off-line data assimilation in palaeoclimatology: a case study, Clim. Past, 11, 81–93, http://www.clim-past.net/11/81/2015/, 2015.

Matsikaris, A., Widmann, M., and Jungclaus, J.: Influence of proxy data uncertainty on data assimilation for the past climate, Clim. Past, 12, 1555–1563, http://www.clim-past.net/12/1555/2016/, 2016.

Mauri, A., Davis, B., Collins, P., and Kaplan, J.: The climate of Europe during the Holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation, Quaternary Science Reviews, 112, 109–127, http://www.sciencedirect.com/science/article/pii/S0277379115000372, 2015.

Mazza, E., Ulbrich, U., and Klein, R.: The tropical Transition of the October 1996 Medicane in the Western Mediterranean Sea: a warm seclusion event., Monthly Weather Review, 2017.

Miguez-Macho, G., Stenchikov, G. L., and Robock, A.: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations, Journal of Geophysical Research-atmospheres, 109, D13 104, doi:10.1029/2003JD004495, 2004.

Okazaki, A. and Yoshimura, K.: Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction, Clim. Past, 13, 379–393, http://www.clim-past.net/13/379/2017/, 2017.

Pardowitz, T., Befort, D. J., Leckebusch, G. C., and Ulbrich, U.: Estimating uncertainties from high resolution simulations of extreme wind storms and consequences for impacts, Meteorologische Zeitschrift, 25, 531–541, doi:10.1127/metz/2016/0582, http://dx.doi.org/10.1127/metz/2016/0582, 2016.

Perkins, W. A. and Hakim, G. J.: Reconstructing paleoclimate fields using online data assimilation with a linear inverse model, Clim. Past, 13, 421–436, https://www.clim-past.net/13/421/2017/, 2017.

Prömmel, K., Cubasch, U., and Kaspar, F.: A regional climate model study of the impact of tectonic and orbital forcing on African precipitation and vegetation, Palaeogeography, Palaeoclimatology, Palaeoecology, 369, 154–162, http://www.sciencedirect.com/science/article/pii/S0031018212005834, 2013.

Renssen, H., Isarin, R., Jacob, D., Podzun, R., and Vandenberghe, J.: Simulation of the Younger Dryas climate in Europe using a regional climate model nested in an AGCM: preliminary results, Global and Planetary Change, 30, 41–57, http://www.sciencedirect.com/science/article/pii/S0921818101000765, 2001.

Russo, E.: Mid-to-Late Holocene Climate and Ecological Changes over Europe, Ph.D. thesis, Freie Universität of Berlin, 2016.

Russo, E. and Cubasch, U.: Mid-to-late Holocene temperature evolution and atmospheric dynamics over Europe in regional model simulations, Clim. Past, 12, 1645–1662, http://www.clim-past.net/12/1645/2016/, 2016.

Samartin, S., Heiri, O., Joos, F., Renssen, H., Franke, J., Brönnimann, S., and Tinner, W.: Warm Mediterranean mid-Holocene summers inferred from fossil midge assemblages, Nature Geoscience, 10, 207–, http://dx.doi.org/10.1038/ngeo2891, 2017.

Steiger, N. and Hakim, G.: Multi-time scale data assimilation for atmosphere-ocean state estimates, Clim. Past Discuss., 11, 3729–3757, http://www.clim-past-discuss.net/11/3729/2015/, 2015.

5 Steiger, N. and Smerdon, J.: A pseudoproxy assessment of data assimilation for reconstructing the atmosphere-ocean dynamics of hydroclimate extremes, Clim. Past Discuss., 2017, 1–22, https://www.clim-past-discuss.net/cp-2017-69/, 2017.

Steiger, N., Hakim, G., Steig, E., Battisti, D., and Roe, G.: Assimilation of Time-Averaged Pseudoproxies for Climate Reconstruction, Journal of Climate, 27, 426–441, doi:10.1175/JCLI-D-12-00693.1, http://dx.doi.org/10.1175/2011JCLI4094.1, 2014.

Talagrand, O.: Assimilation of observations, an introduction, Journal-Meteorological Society of Japan Series 2, 75, 81–99, 1997.

10 Verbunt, M., Walser, A., Gurtz, J., Montani, A., and Schär, C.: Probabilistic flood forecasting with a limited-area ensemble prediction system: selected case studies, Journal of Hydrometeorology, 8, 897–909, 2007.

Wang, Y., Bellus, M., Wittmann, C., Steinheimer, M., Weidle, F., Kann, A., Ivatek-Šahdan, S., Tian, W., Ma, X., Tascu, S., et al.: The Central European limited-area ensemble forecasting system: ALADIN-LAEF, Quarterly Journal of the Royal Meteorological Society, 137, 483–502, 2011.
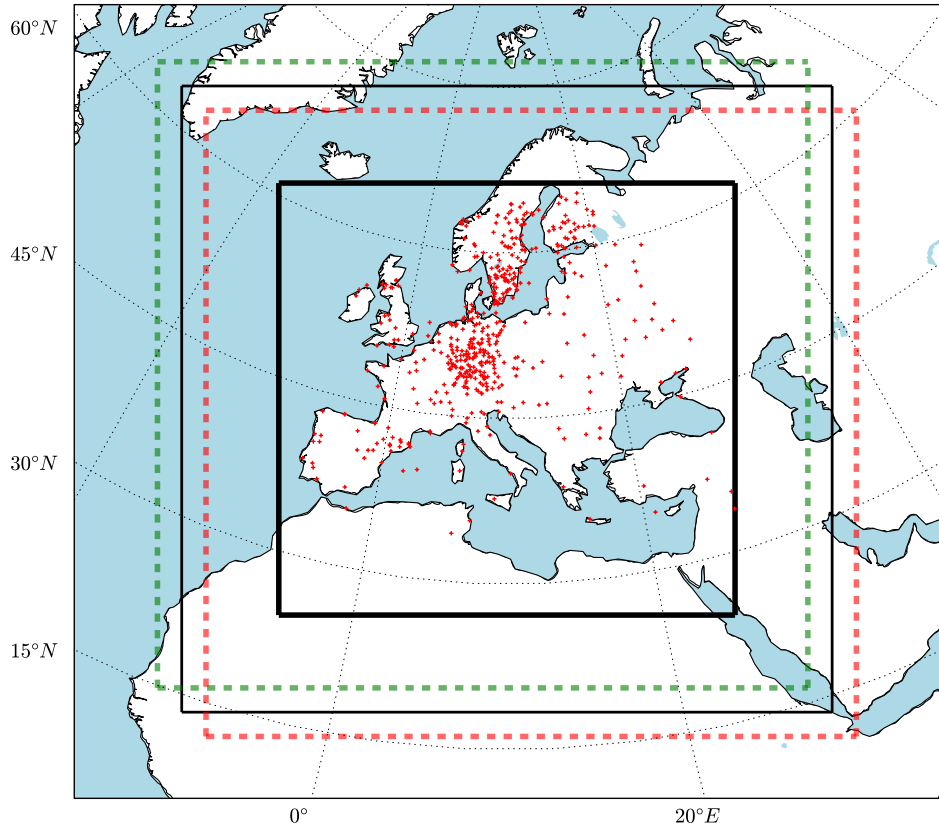
Figure 1. rcm! (**rcm!**) domains: black thin box shows the nature and dashed boxes show the shifted members (only two are shown for North-west and Southeast with 5 grid points distance from nature) and red pluses show 500 random meteorological stations of the "ENSEMBLES daily gridded observational dataset for precipitation, temperature and sea level pressure in Europe called E-OBS" (Haylock et al., 2008). The thick black box shows the evaluation domain.
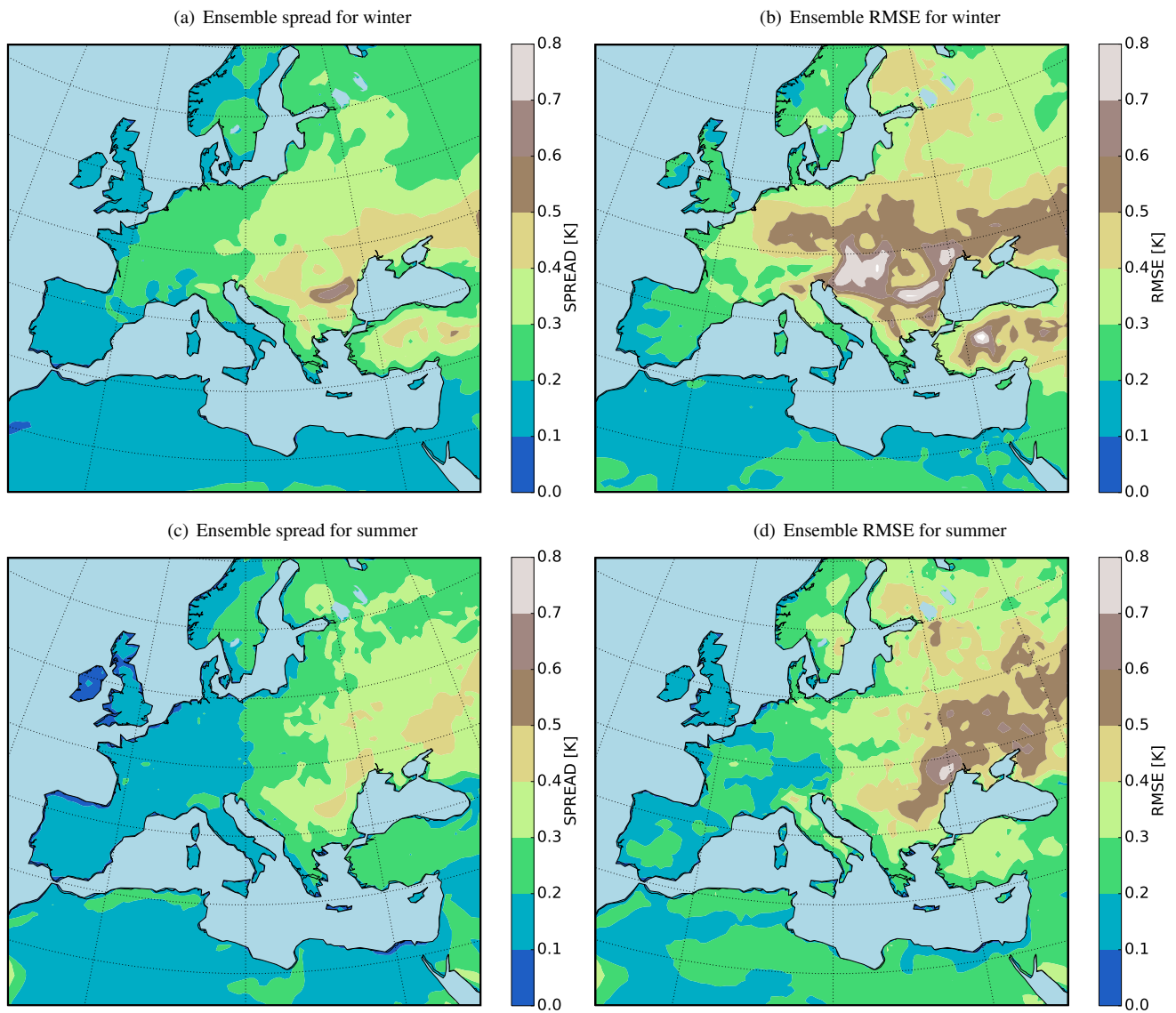
Figure 2. 10-year ensemble spread and RMSE for seasonal mean of 2 meter temperature (K).
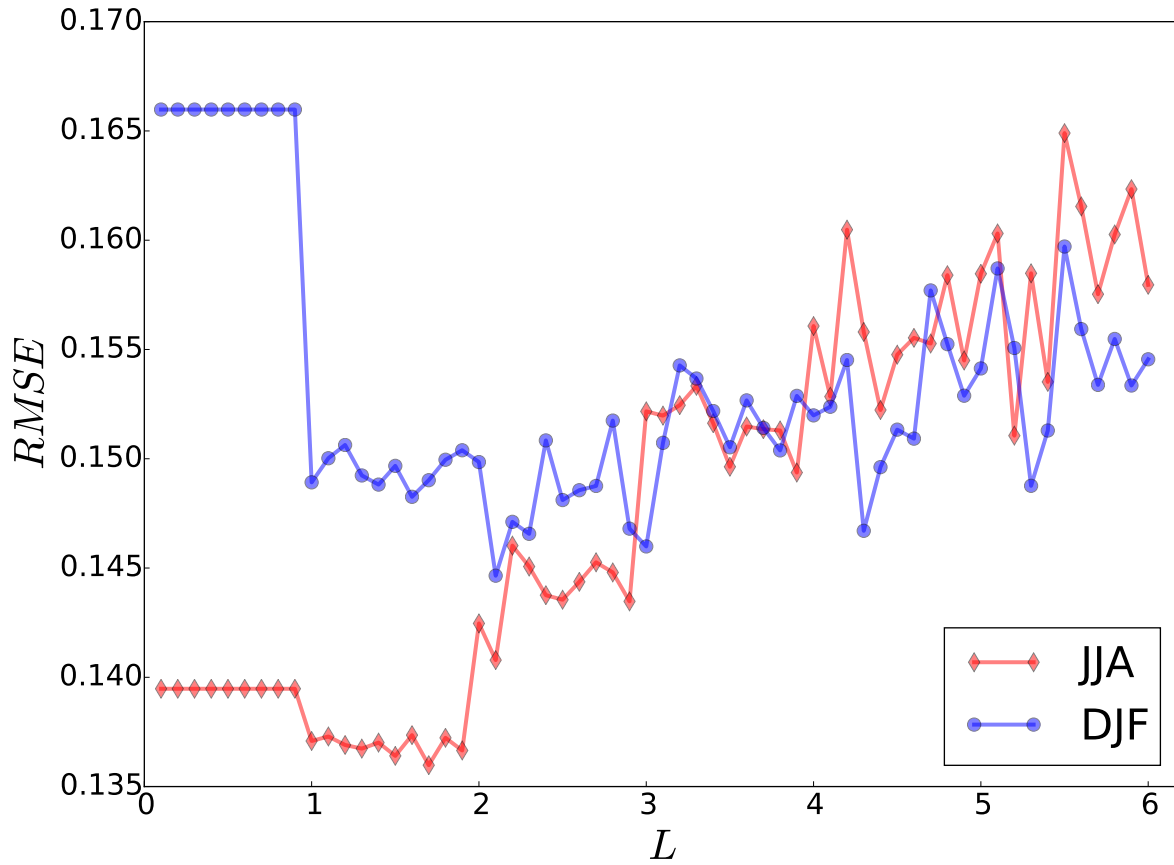
Figure 3. Field-mean of RMSE for near surface temperature (K) analysis over the evaluation domain for different correlation length ($L[°]$) for Winter (blue) and Summer (red).
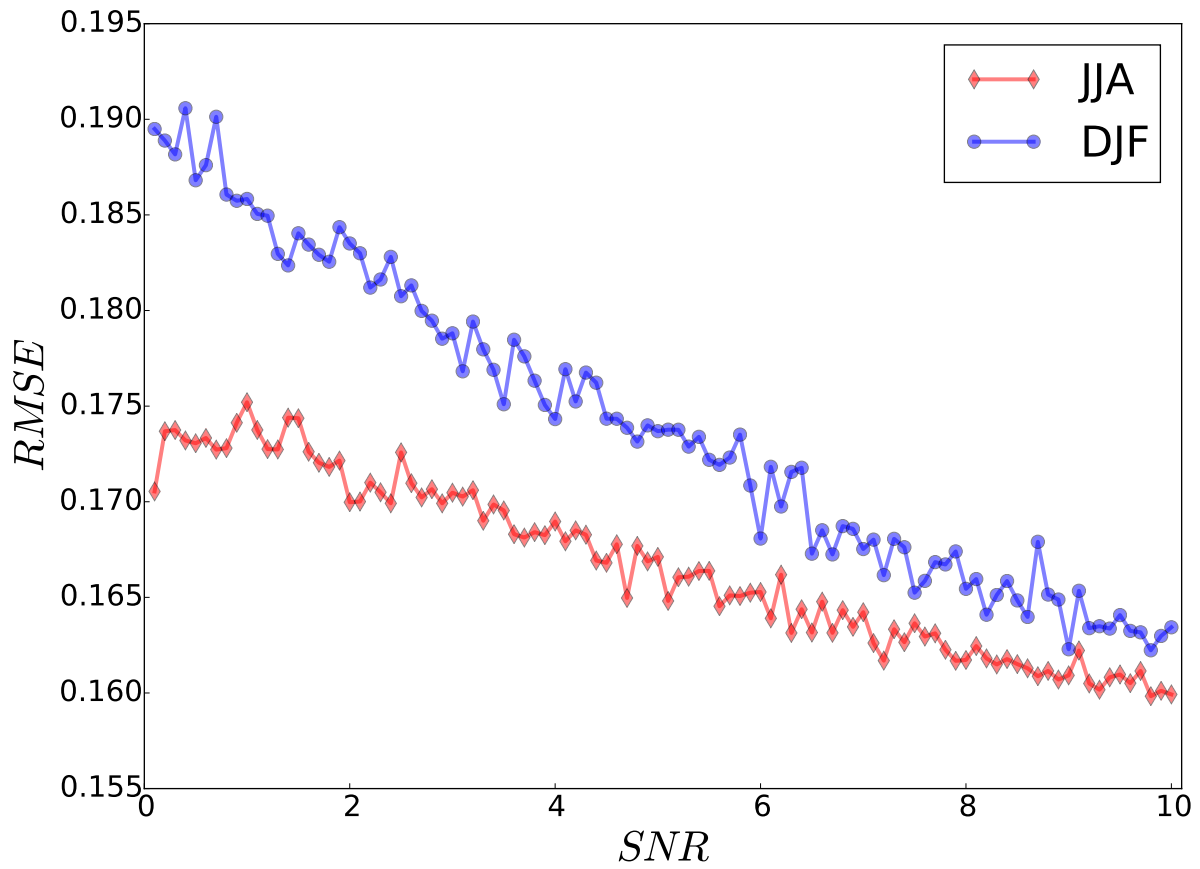
Figure 4. Field-mean of seasonal near-surface temperature RMSE [K] of analysis vs SNR.
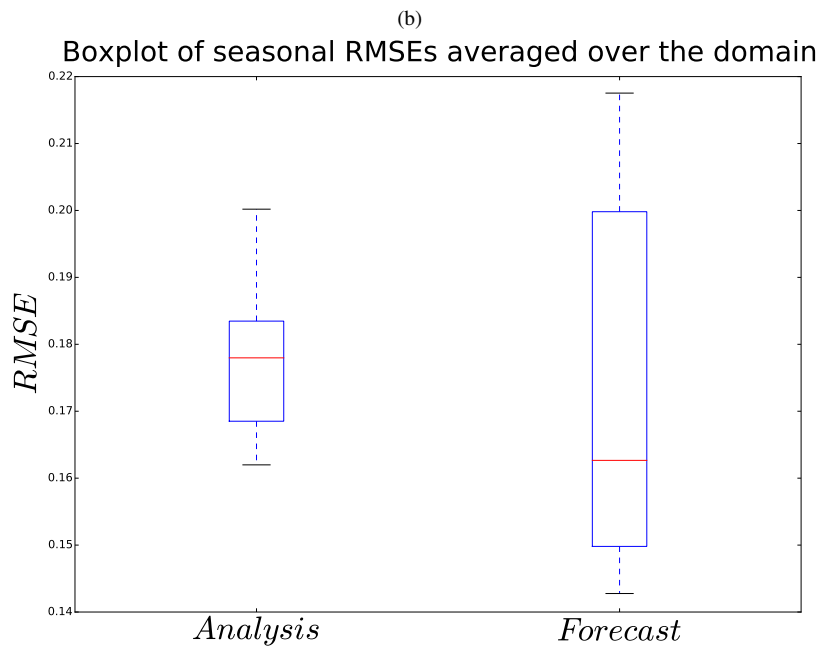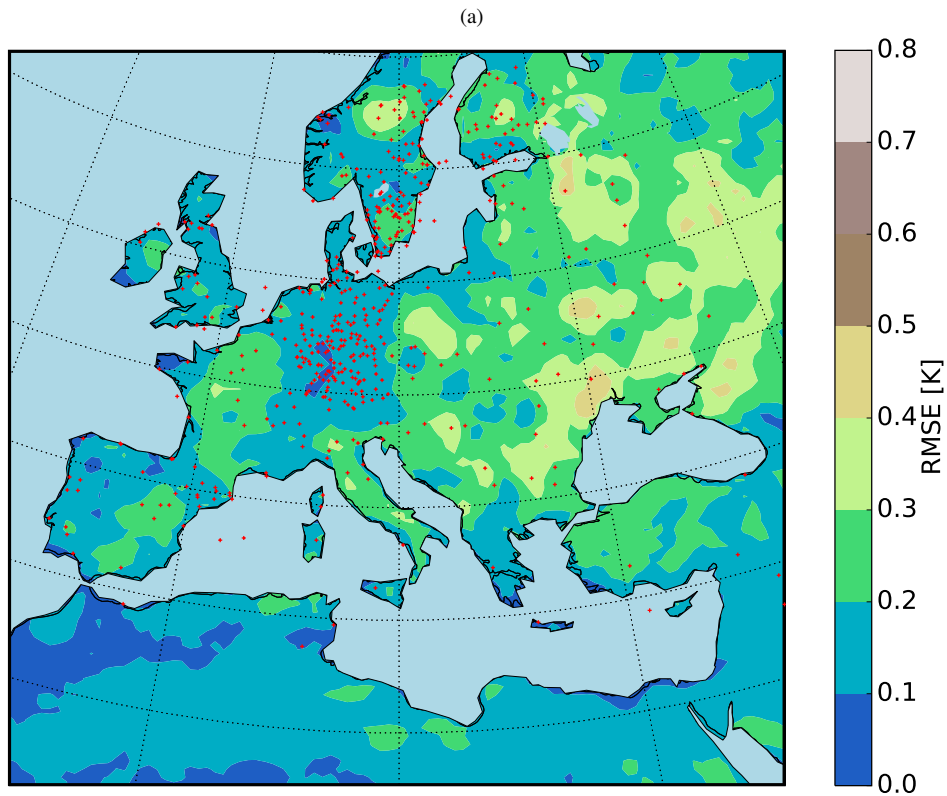
Figure 5. (a): 10-year RMSE [K] of analysis for summer (JJA) and **(b)**: the field-mean RMSE [K] of analysis and free ensemble run for summer (JJA).
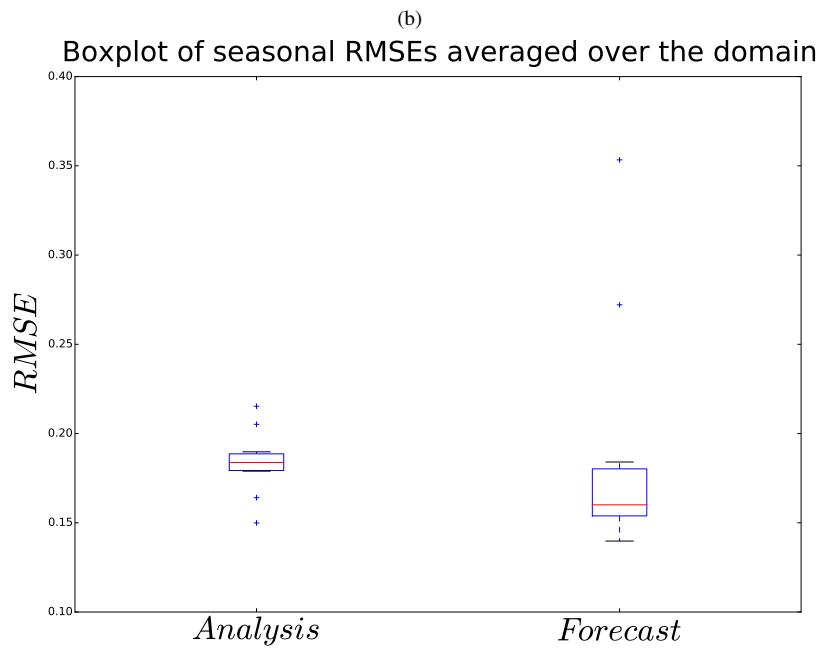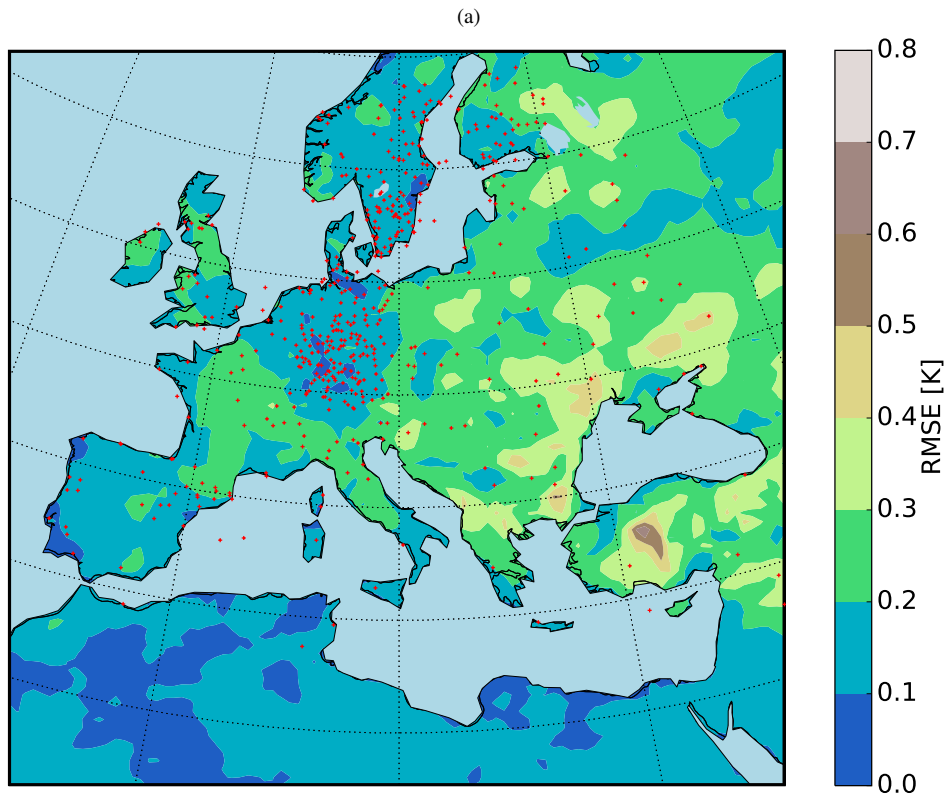
Figure 6. (a): 10-year RMSE [K] of analysis for winter (DJF) with and **(b)**: the field-mean RMSE [K] of analysis and free ensemble run for winter (DJF).
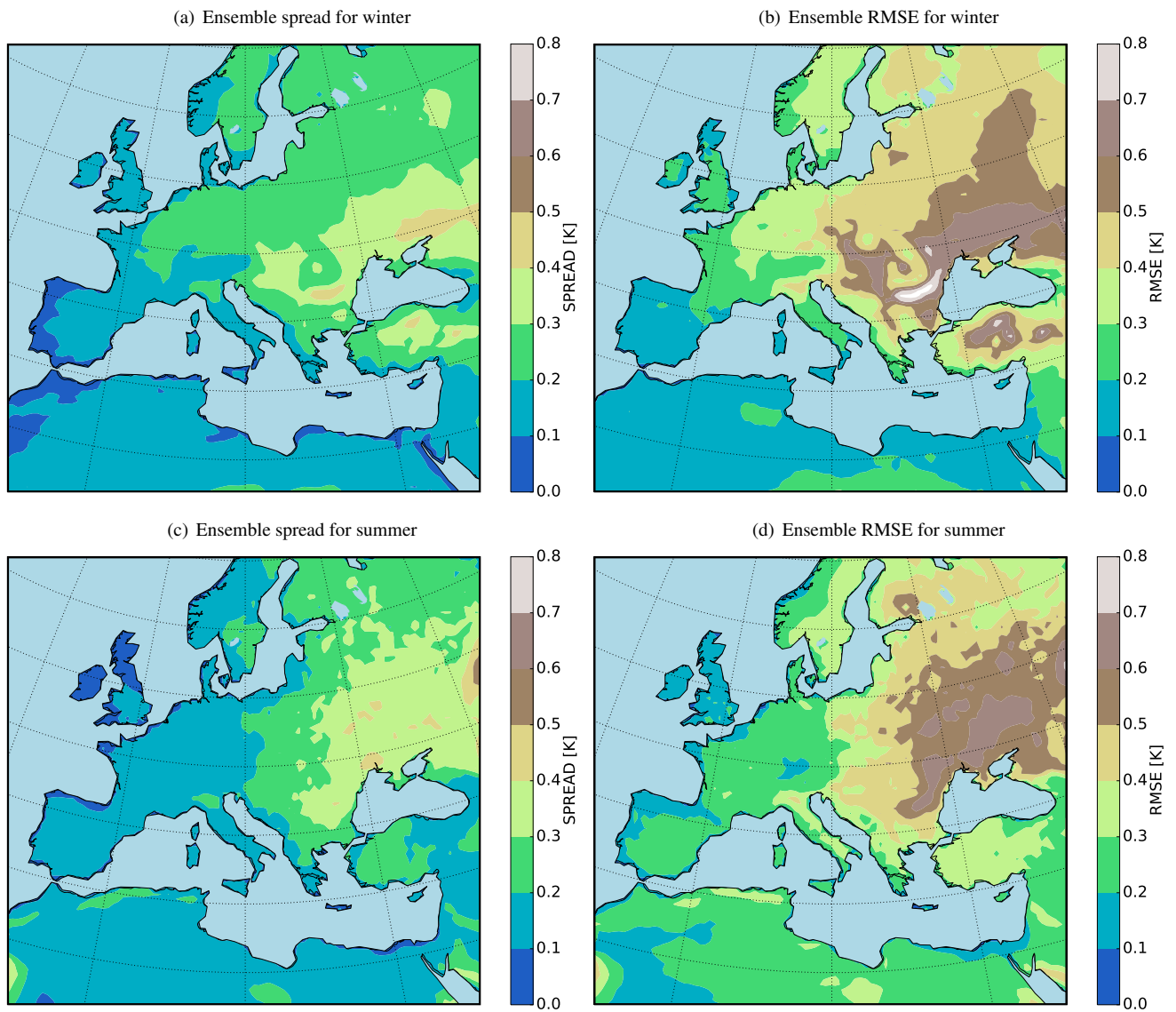
Figure 7. 36-year ensemble spread and RMSE for seasonal mean of 2 meter temperature (K).
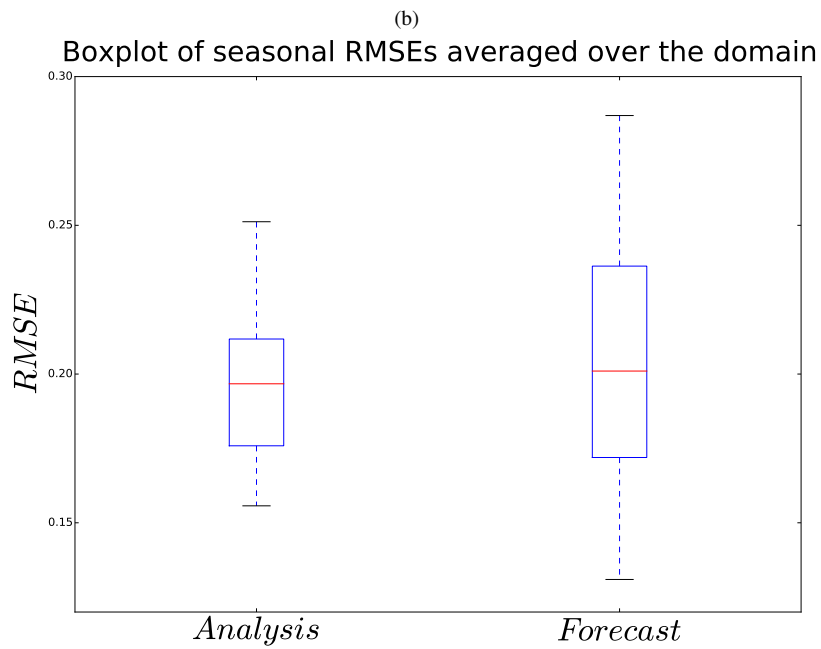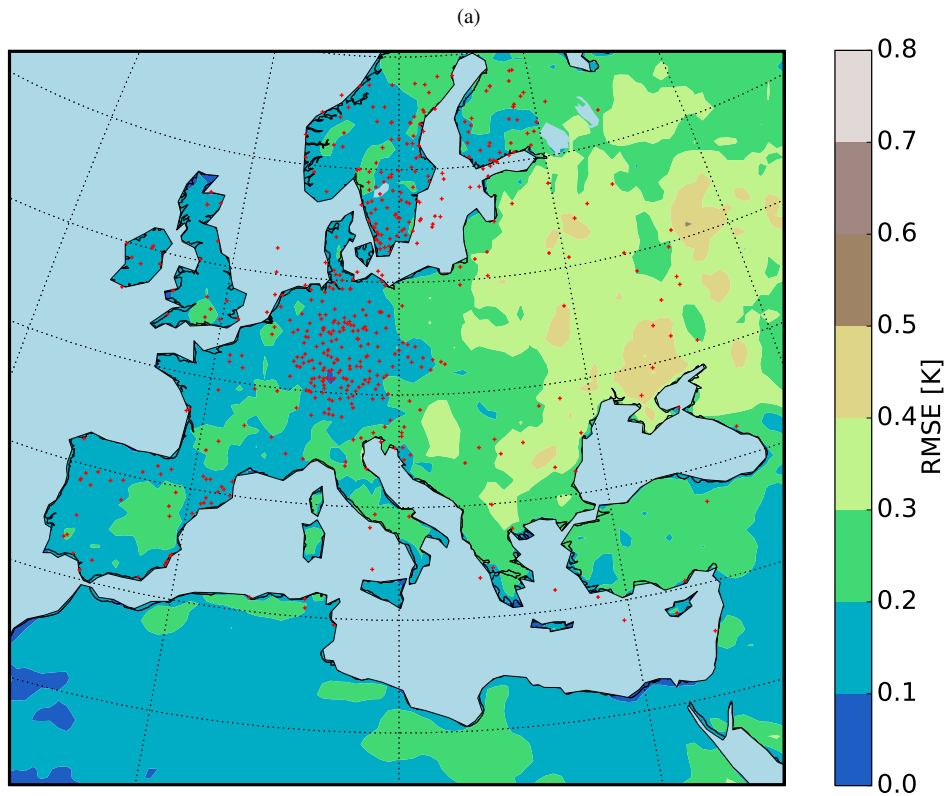
Figure 8. (a): 36-year RMSE [K] of analysis for summer (JJA) and **(b)**: the field-mean RMSE [K] of analysis and free ensemble for summer (JJA). Red squares indicate the mean of RMSEs.

(a)

(b)
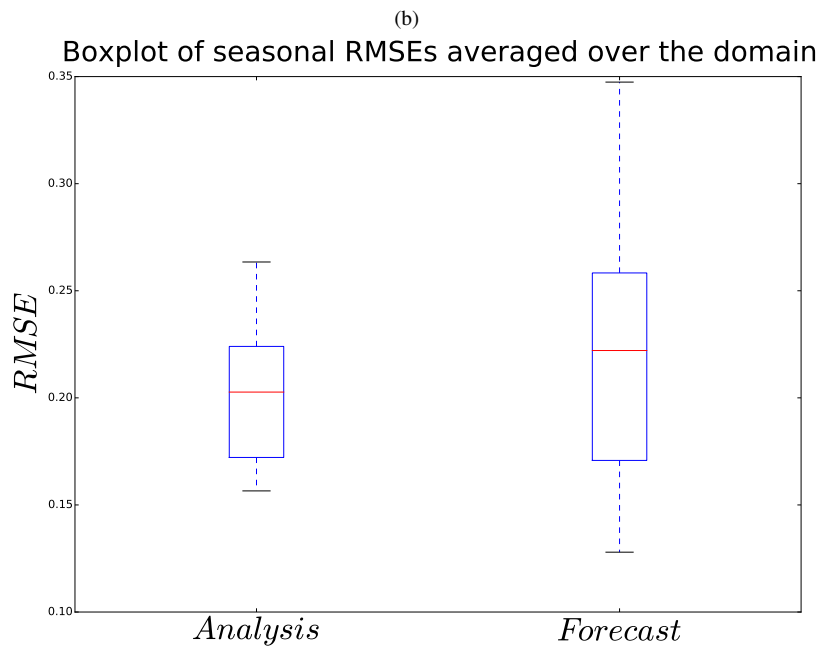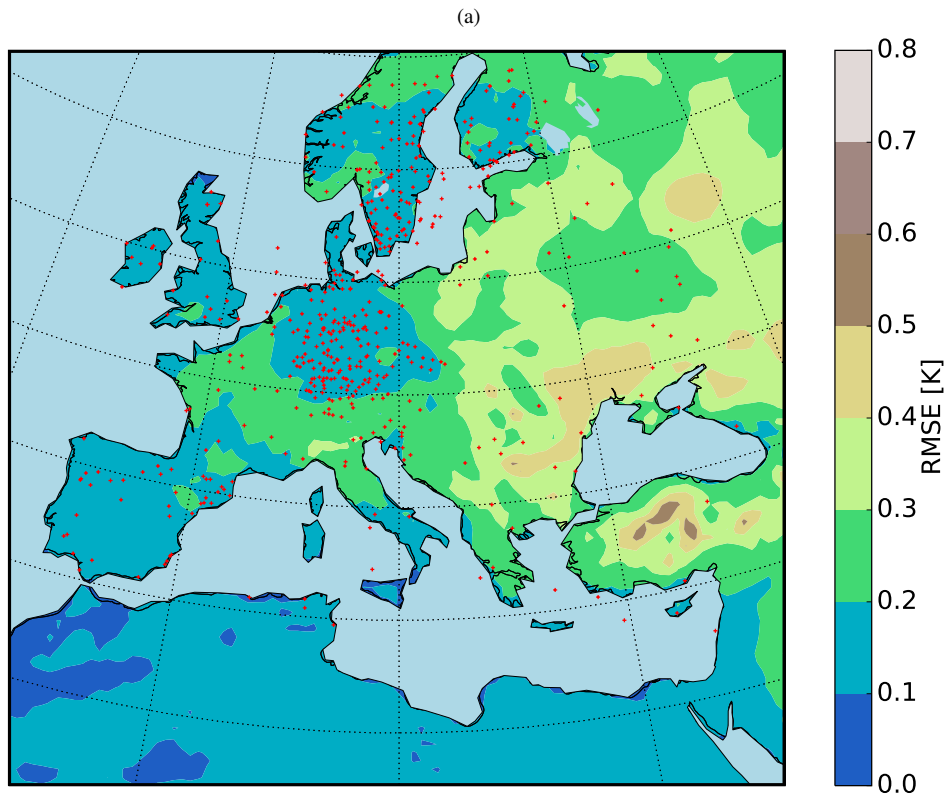
Figure 9. (a): 36-year RMSE [K] of analysis for winter (DJF) and **(b)**: the field-mean RMSE [K] of analysis and free ensemble run for winter (DJF). Red squares indicate the mean of RMSEs.
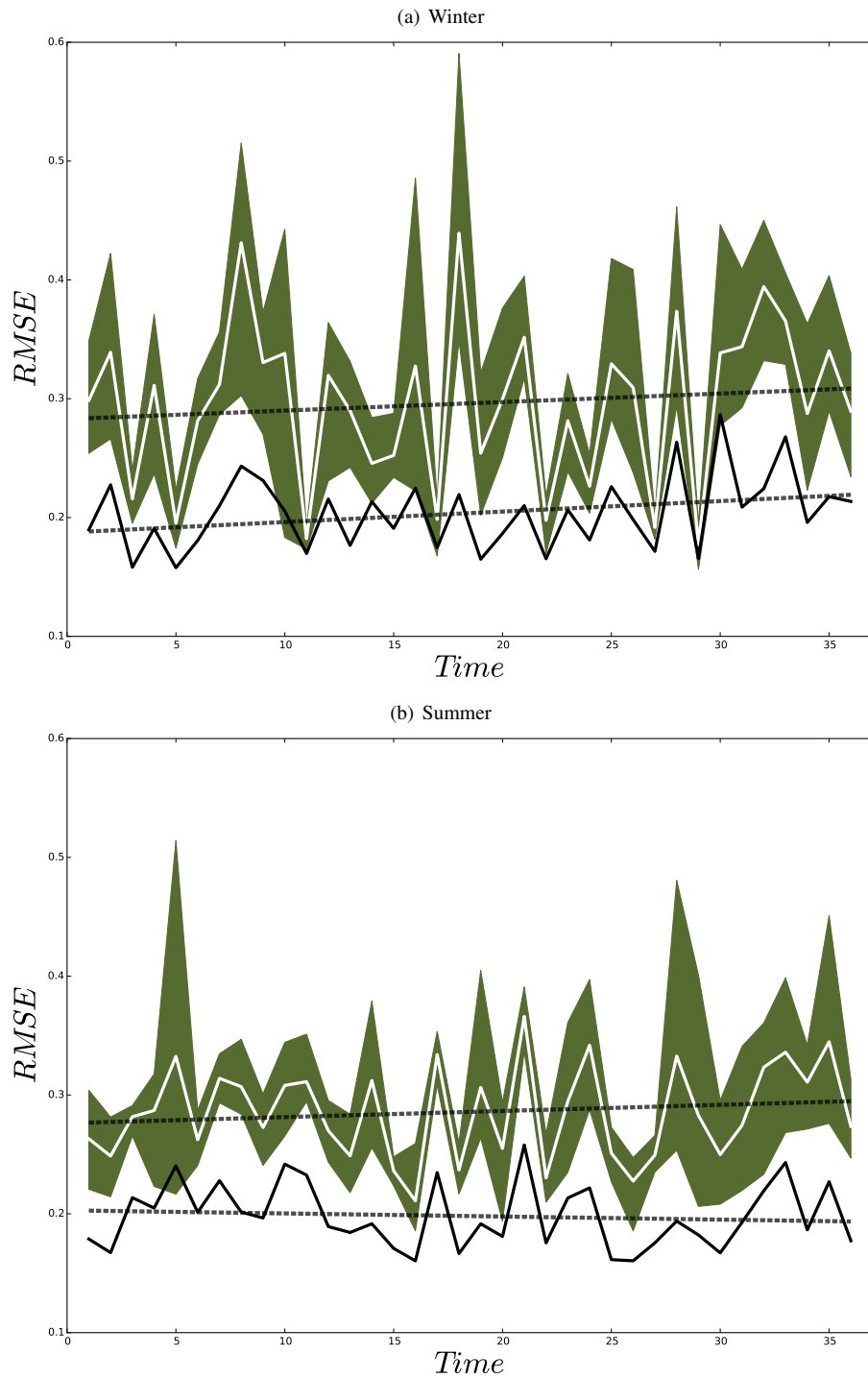
Figure 10. Yearly field-mean RMSE [K] of ensemble mean (white line) and analysis (black line) for winter **(a)** and for summer **(b)**. Shadings show the ensemble members and dashed lines indicate the linear trends.
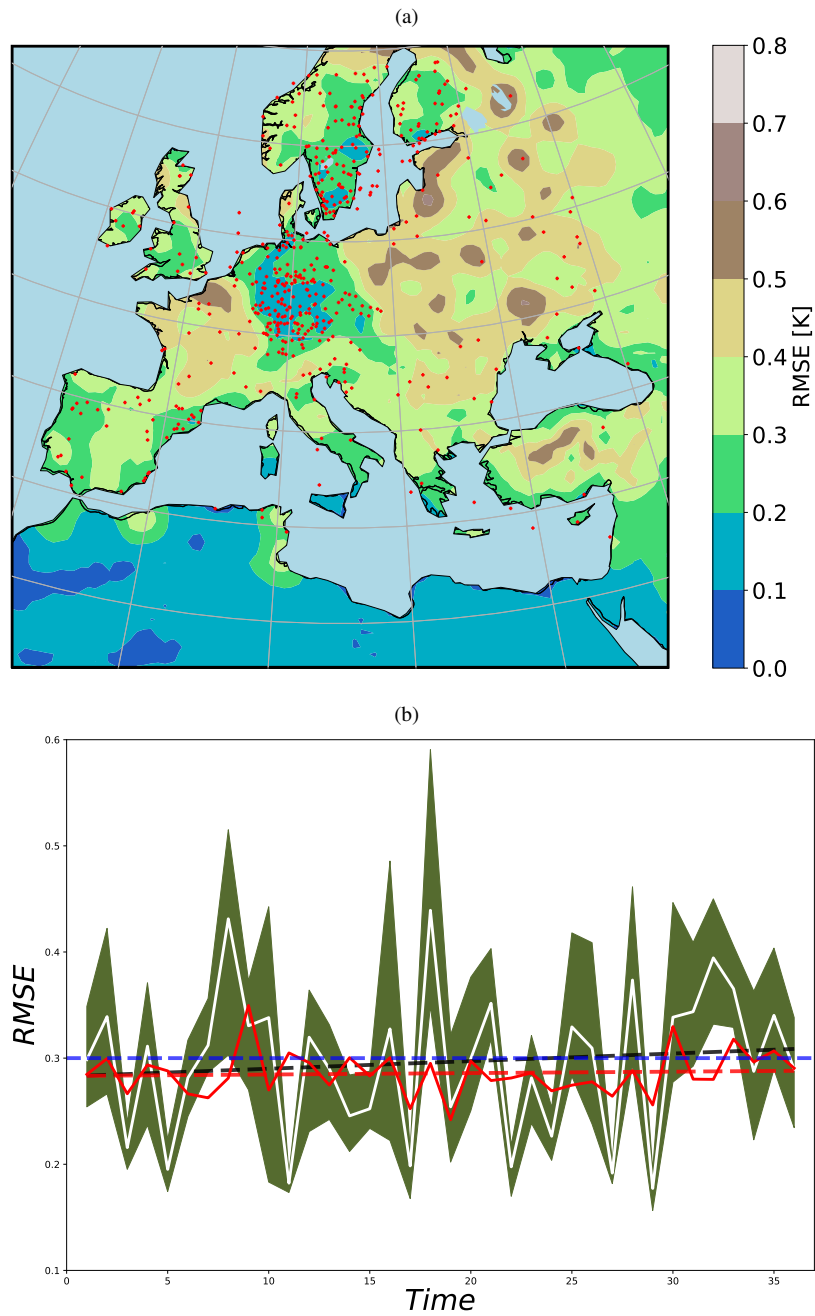
Figure 11. (a): 36 years averaged of RMSE of analysis using 40 random states as background, (b): fieldmean of RMSE from the free ensemble (shading shows the ensemble spread, the white line the mean) and from the analysis using random background states (red line). Dashed lines is the linear fits and the blue line shows the 0.3 K RMSE value (plotted only for comparison of the trends).
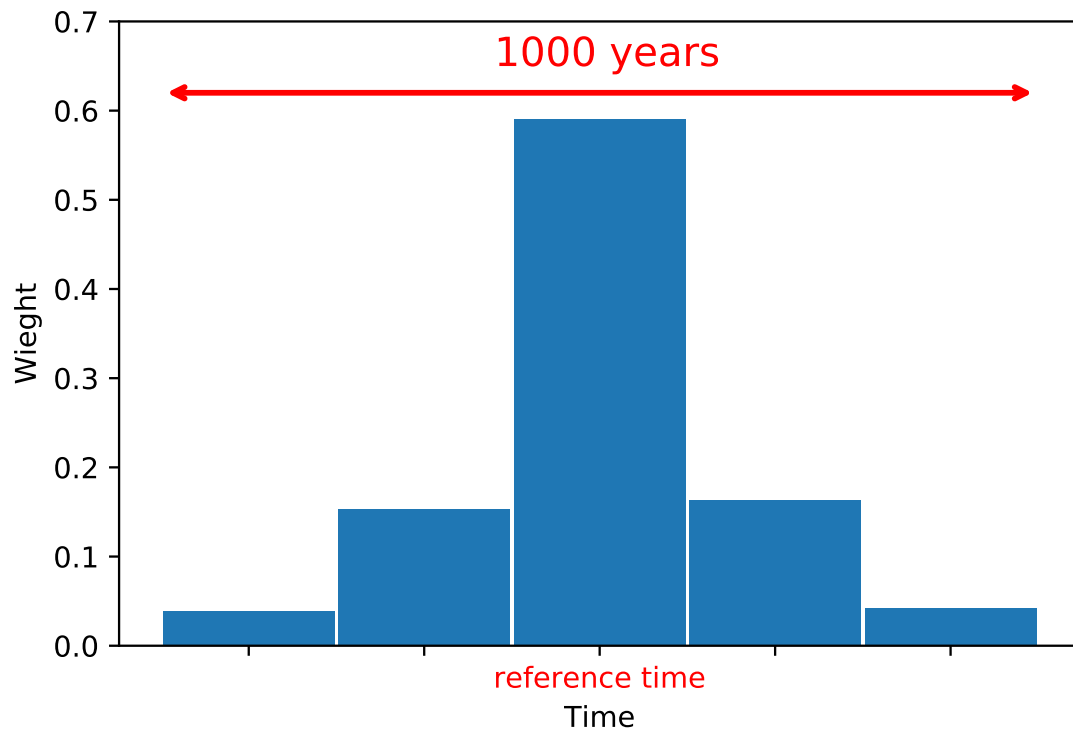
Figure 12. Schematic showing the weights for observations with respect to their distance to the target year. Time window of 1000 years is chosen. The observations are weighted depending on their distances to the reference time.
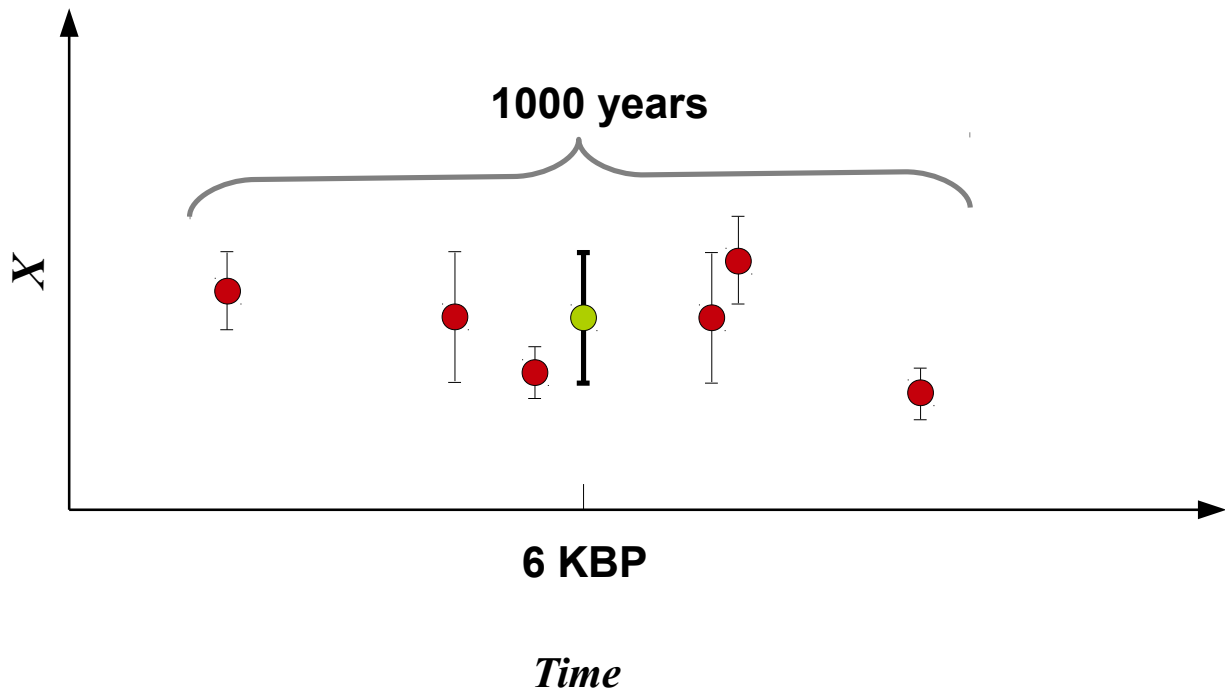
Figure 13. Schematic showing the weights for observations with respect to their standard error. Time window of 1000 years is chosen. The red dots resemble the proxies in the 1000-year time window and the green dot resembles the weighted mean.
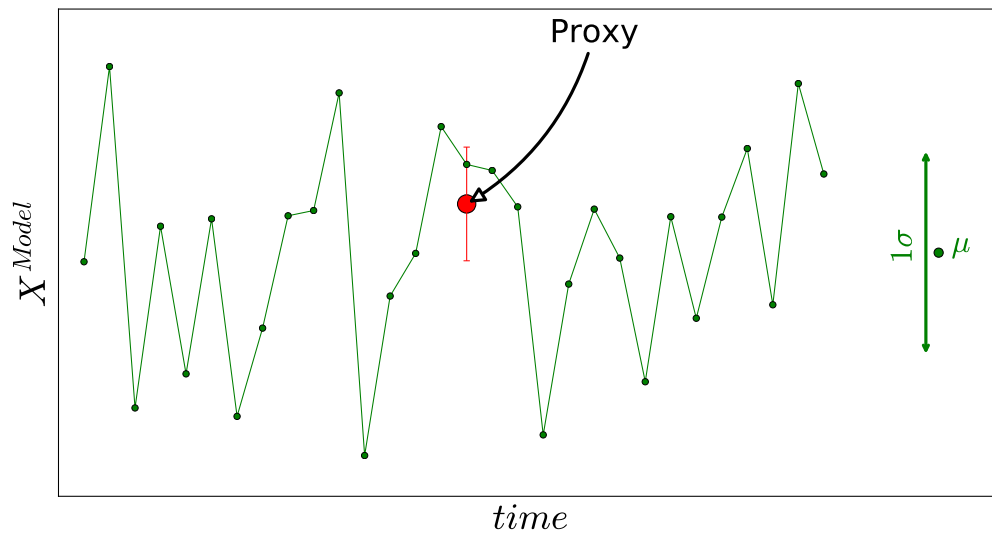
Figure 14. Schematic showing how the expected value (the mean) and the deviation from the mean for the time-slice simulation is selected. The model simulation is 25 years long. Green line resembles the model state.
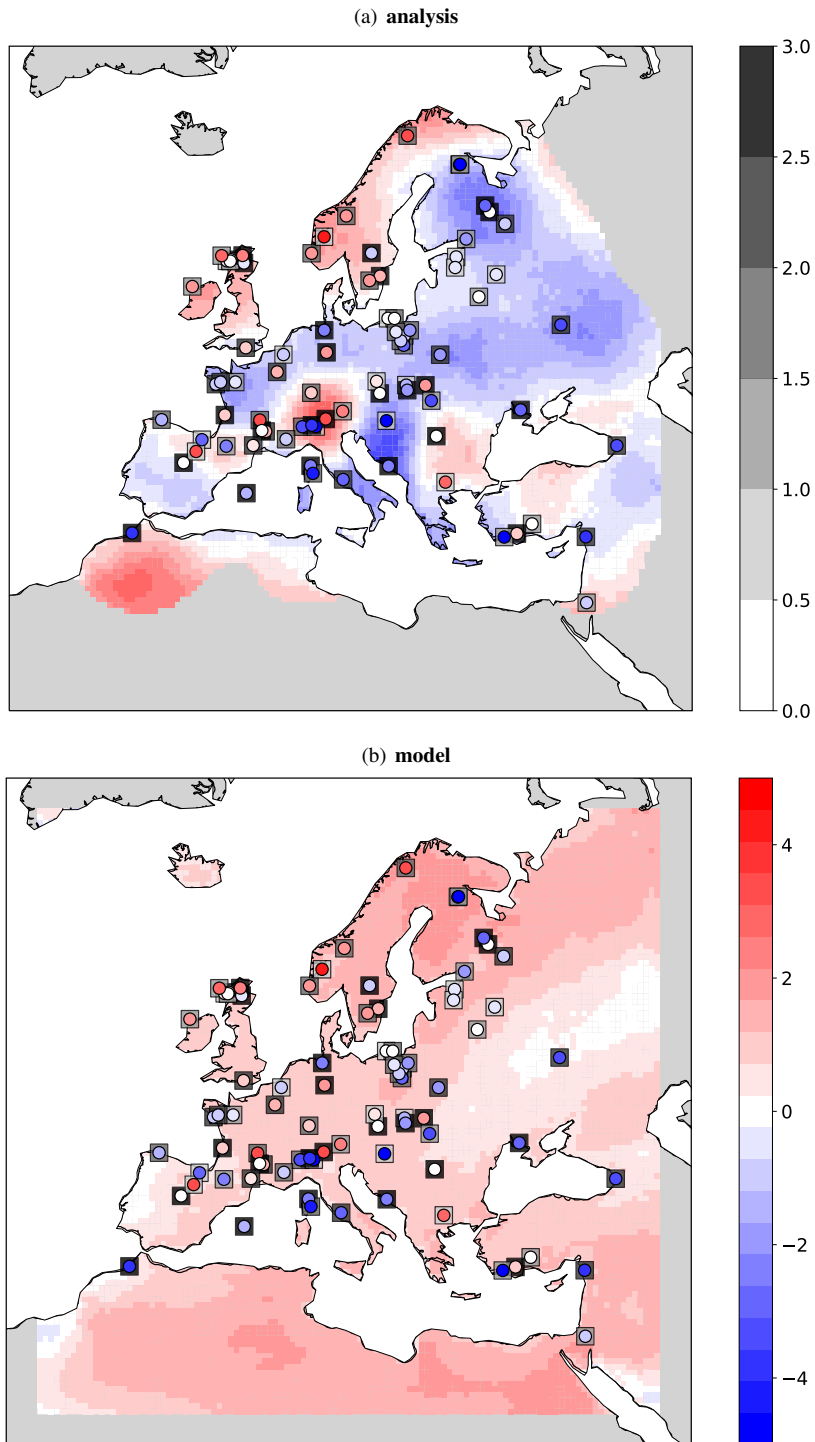
Figure 15. DA results of T2m during summer at 6KBP using weighted arithmetic mean by time distances: **(a)**: Analysis values along with the testing observations (circles) and their standard error (squares). Values with error covariance of analysis greater than 0.9 are masked out from analysis, and **(b)** the model forecast.
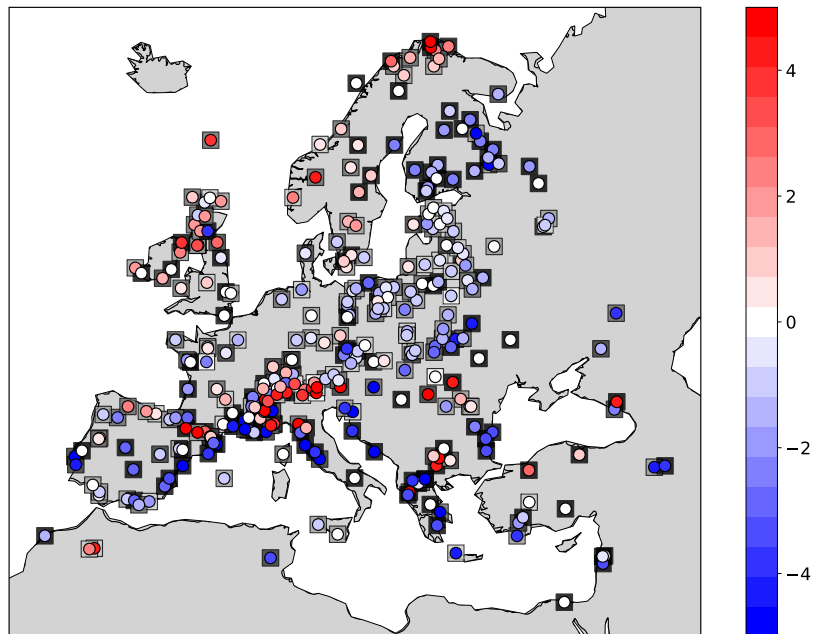
Figure 16. Weighted arithmetic mean using time distances: Anomalies ((6KBP - 0.2KBP)) of assimilated observations (circles) superposed on their standard errors(squares) with values in K. Color-bar of the standard errors as in Figure 15.a.
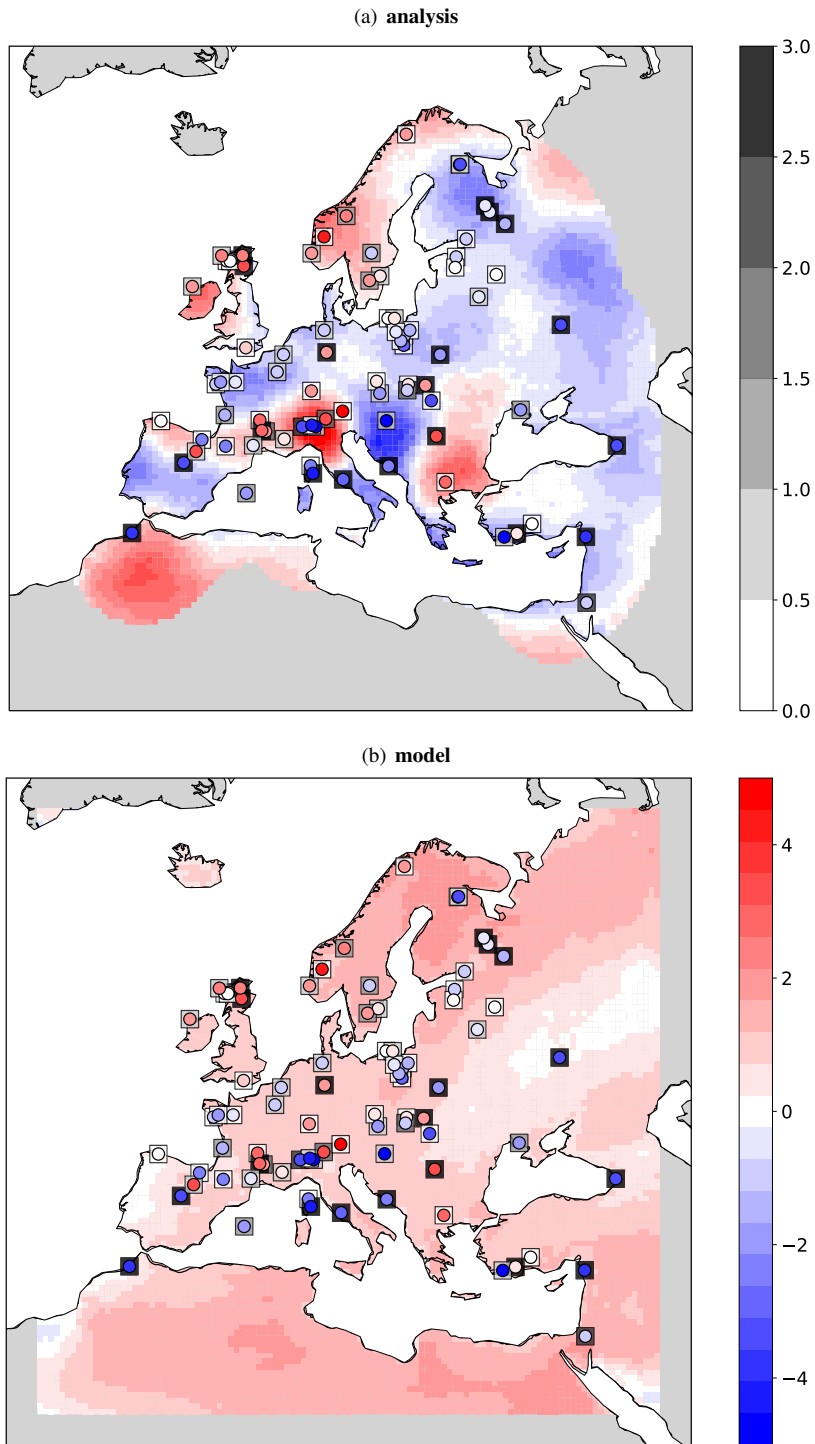
Figure 17. DA results of T2m during summer at 6KBP using weighted arithmetic mean by standard errors: **(a)**: Analysis values along with the testing observations (circles) and their standard error (squares). Values with error covariance of analysis greater than 0.9 are masked out from analysis, and **(b)** the model forecast.
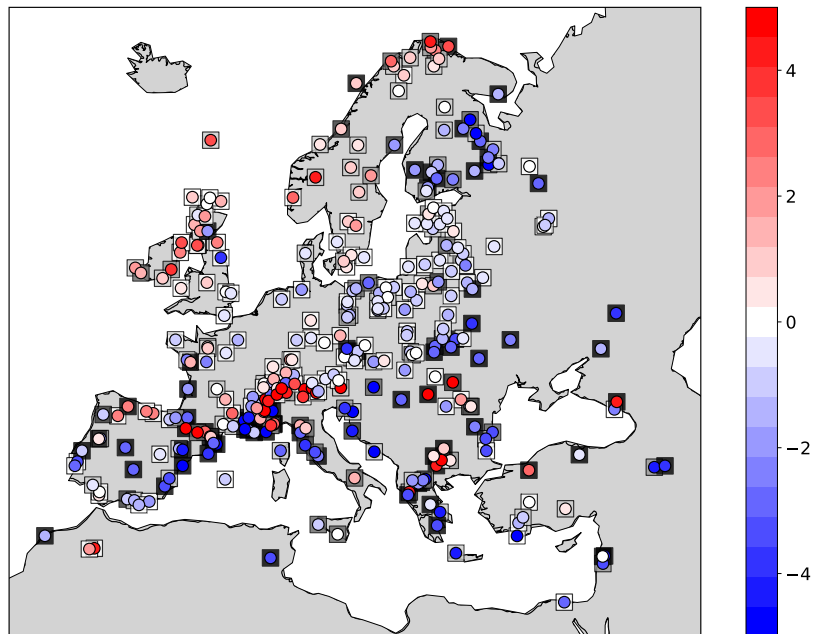
Figure 18. Weighted arithmetic mean using standard errors: Anomalies ((6KBP - 0.2KBP)) of assimilated observations (circles) superposed on their standard errors(squares) with values in K. Color-bar of the standard errors as in Figure 17.a.