Reconstructing winter climate anomalies in the Euro-Atlantic sector using circulation patterns

Erica Madonna[1], David S. Battisti[2], Camille Li[1], and Rachel H. White[3]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway
[2]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[3]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia

Correspondence: Erica Madonna (erica.madonna@uib.no)

Abstract. The efficacy of Euro-Atlantic circulation regimes for estimating wintertime climate anomalies (precipitation and surface temperature) over Europe is assessed. A comparison of seasonal climate reconstructions from two different regime frameworks (cluster analysis of the low-level zonal flow, and traditional blocking indices) is presented and contrasted with seasonal reconstructions using the North Atlantic Oscillation (NAO) index. The reconstructions are quantitatively evaluated

5 using correlations and the coefficient of efficiency, accounting for misfit in phase and amplitude. While all methods show skill in reproducing wintertime precipitation and temperature variability, they underestimate the magnitude of seasonal anomalies by a factor of two. The regimes are found to capture more spatial structure in seasonal precipitation anomalies over Europe than the NAO, with the jet framework showing generally better skill relative to the blocking indices. The reconstructions of temperature anomalies have lower skill than those for precipitation, with the best results for temperature obtained by the NAO

10 for high-latitude and by the blocking framework for southern Europe.

1 Introduction

Seasonal precipitation and temperature anomalies over Europe exhibit large year-to-year variability, with direct societal impacts such as on crop yields and renewable energy production (Grams et al., 2017; Jerez et al., 2013; Lesk et al., 2016). The seasonal

15 climate signal results from a succession of daily weather that is often organized by the large-scale flow into a finite number of preferred circulation patterns, also called regimes (Corte-Real et al., 1995; Molteni et al., 1990; Vautard, 1990). Therefore, one might expect that seasonal climate anomalies could be reconstructed from the frequency of the dominant atmospheric patterns over a season. This approach has been used to study future trends in European precipitation (Santos et al., 2016) and as a predictor for droughts (Lavaysse et al., 2018). Atmospheric variability patterns can be characterized in several ways,

20 such as by the North Atlantic Oscillation index (NAO, e.g. Hurrell, 1995), the occurrence of blocking (e.g. Pfahl and Wernli, 2012; Sillmann and Croci-Maspoli, 2009; Trigo et al., 2004), and the configuration of the North Atlantic jet stream (Woollings et al., 2010; Madonna et al., 2017). All these complementary classifications have been shown to be able to capture specific

aspects of the wintertime climate over Europe, but there has been to our knowledge no direct comparison of the skill of these classifications in reconstructing seasonal climate.

25 In this study we compare how well European winter conditions are described by indices of the NAO, blocking, and North Atlantic jet configurations. The NAO manifests as mean sea level pressure (MSLP) fluctuations with anticorrelated extrema between two poles, one over the Azores and one over Iceland. These fluctuations influence the strength of the prevailing westerly winds and the propagation path of storms into Europe (e.g., Hurrell, 1995; Hurrell et al., 2003; Qian et al., 2000; Rogers, 1997; Wanner et al., 2001). In contrast, blocking is the presence of a persistent and stationary high-pressure system

30 that obstructs the westerly flow (Rex, 1950). In the North Atlantic sector blocking occurs mainly over three regions: Greenland, Scandinavia, and the Iberian Peninsula (Treidl et al., 1981; Davini et al., 2014, also shown in Fig. 1). Precipitation is reduced within the blocked region (Sousa et al., 2017) while cold temperature extremes are often observed southeast of the blocked region (Sillmann and Croci-Maspoli, 2009). Lastly, jet configurations describe preferred flow paths of the North Atlantic jet stream, which acts as waveguide for midlatitude storms (Athanasiadis et al., 2010; Hoskins and Ambrizzi, 1993; Wettstein and

35 Wallace, 2010; Wirth et al., 2018). In winter the North Atlantic jet stream can assume five different configurations (southern, central, northern, tilted, and mixed, Madonna et al., 2017) with distinct patterns of storminess, all of which are associated with regional climate anomalies over Europe.

The three classification methods for circulation regimes that we use in this study are not independent. The strength and position of the jet are intrinsically linked to changes in MSLP and thus to the NAO and blocking. In fact, blocking over

40 Greenland, Scandinavia and the Iberian Peninsula correspond to a southern, mixed, and northern configuration of the jet stream, respectively (Madonna et al., 2017). However, the NAO does not map clearly onto distinct jet configurations or blocking patterns, with the positive phase being especially ambiguous (Woollings et al., 2010; Davini et al., 2014; Madonna et al., 2017).

These classifications of North Atlantic atmospheric variability thus offer closely related but alternative views of seasonal climate fluctuations. For example, an extreme season may occur due to the unusual persistence or frequency of a certain jet

45 configuration (Madonna et al., 2019), without a corresponding extreme value of the NAO index or blocking pattern. This study aims to compare the ability of three classification methods (NAO, blocking, jet configurations) to reconstruct seasonal climate anomalies over Europe. By knowing the frequency of each circulation pattern, we assess the skill of each method to reproduce the sign (i.e. correlation) and amplitude (i.e. ratio of standard deviations) of seasonal precipitation and temperature anomalies.

2   Methods and data

50 We focus on the low-level wind (900 to 700 hPa), two-metre temperature, and total precipitation. The analyses are conducted for winter (DJF; the 90 day period December 1 –February 28), with ERA-Interim reanalysis (Dee et al., 2011) for the period 1979-2014, interpolated to a 0.5° horizontal resolution. To identify blocking, 6 hourly data of geopotential height at 500 hPa are used. For the rest of the analysis, daily means are used.

2.1 Classifications

2.1.1 NAO

We use the daily NAO time series from NOAA (downloaded from ftp://ftp.cpc.ncep.noaa.gov/cwlinks/norm.daily.nao.index.
b500101.current.ascii), which is calculated using geopotential height at 500 hPa and covers the whole North Atlantic basin
north of 20°N (green box in Figure 1). A day is classified as a positive (negative) NAO day if its NAO value exceeds 0.5 (-0.5)
of the wintertime NAO standard deviation. The remaining days, about 35 days per winter (Table 1), are considered neutral
NAO days.

2.1.2 Blocking

Blocking events are identified on 6-hourly data following the criteria from Scherrer et al. (2006), which define a block to be
when there is a reversal in the meridional gradient of the geopotential height at 500 hPa in a 30° latitudinal band that lasts for
at least 5 days. Climatologically in the North Atlantic there are three main regions affected by blocking (shading in Figure 1),
one over Greenland, one over Northern Europe/Scandinavia, and one offshore of the Iberian Peninsula. Note that the enhanced
frequency at 30°N is an artefact of the detection method (see discussion in Davini et al., 2014). We define three boxes to capture
these three regions: Greenland (65-25°W, 60-75°N, GB, orange box in Figure 1), Scandinavia (15°W-25°E, 50-65°N, SBL,
red box) and offshore of the Iberian Peninsula (30°W-0°, 40-50°N, named as in Davini et al. (2014) Iberian wave breaking
(IWB), blue box).

A day is considered a blocked day if at least 10% of the gridpoints in the respective box satisfy the blocking criteria. Table 1
reports the average (standard deviation) number of days per winter with blocking at different locations. During the 35 winters,
we detect on average 13.5 days per winter of GB, 16.9 of SBL and 13.0 of IWB while 51.9 days are considered as "non
blocked" (NB). It can occasionally happen that during one day several regions are simultaneously blocked, therefore the sum
of blocking and NB days do not sum to exactly 90 days (i.e. one winter).

2.1.3 Jet clusters

We calculate mass-weighted average zonal wind (U) between 900-700 hPa in the sector 60°W-0°, 15-75°N (black box, Figure 1) and calculate jet clusters as described in Madonna et al. (2017). We perform an EOF analysis on the low-level wind to
reduce the dimensions and apply a k-mean cluster algorithm to the first five principal components, which explains up to 80%
of the wintertime variability in that sector. The number of clusters has to be chosen *a priori*. Madonna et al. (2017) showed that
four clusters produce jet configurations that correspond to the four classical weather regimes in the Euro-Atlantic sector; however, following their discussion and Dorrington and Strommen (2020), we use five clusters here, which leads to more distinct
jet configurations over the North Atlantic.

Every day is associated with a cluster depending on the normalized inverse distance (d) from the cluster centroid (i.e. $d = 1$
at the centroid location and $d = 0$ far from the centroid) in the five-dimensional space of the principal components. Since some

days can be close to more than one centroid, in particular during transition days from one cluster to the other, we keep only

days whose d to the respective centroid is larger than 0.5. Approximately 54% of the 3150 (=35 x 90) days are unequivocally

attributed to a specific cluster (see Table 1, 46% of days are not assigned to any cluster).

The five clusters represent a southern jet (S-jet), a central jet (C-jet), a northern jet (N-jet), a tilted jet (T-jet), and a mixed jet

(M-jet with a split structure). On average each cluster occurs between 7.3 (S-jet) and 12.0 (T-jet) days per winter, as reported

in Table 1, with large winter-to-winter variability (standard deviation).

2.1.4 Associated weather anomalies

Each NAO phase, blocking category and jet cluster is characterized by different circulation, precipitation and temperature

anomalies. We compute daily anomalies of zonal wind, precipitation and temperature and then construct composites for each

NAO phase, blocking category and jet cluster by averaging the daily fields of all (defined) days belonging to the corresponding

category. For wind and precipitation the daily anomalies are calculated by subtracting the climatological 35-year winter mean,

while for temperature, we remove the daily 35-year average, smoothed with a 30-day running mean, to account for the strong

seasonal cycle. The resulting anomalies are presented in section 3.1.

2.2 Seasonal Reconstructions

To reconstruct the seasonal anomalies we count the number of days in each NAO phase, blocking category and jet cluster for

each winter. Similar to Cortesi et al. (2019), we then reconstruct the seasonal precipitation and temperature anomalies maps

$A_{rec}(\phi, \lambda, t)$ for each season (t) as follows:

$$A_{rec}(\phi, \lambda, t) \quad = \quad \sum_i (Y_i(\phi, \lambda) \cdot f_i(t)) \tag{1}$$

where ϕ, λ and t are latitude, longitude and time, respectively; i represents the two NAO phases, three blocking categories or

five jet clusters; $Y_i(\phi, \lambda)$ are the maps of seasonal average precipitation/temperature anomalies associated with the considered

pattern; and $f_i(t)$ is the fraction of time the pattern occurs in the given season t (i.e. $f_i(t) = \frac{\#\text{days of pattern } i}{\#\text{days per season (=90)}}$).

For example for the jet clusters in DJF 2013/2014 (cf. Figure 5) there are 0 days classified as S-jet and M-jet, 11 days as

T-jet, 7 days as N-jet and 60 days as C-jets. The remaining 12 days belong to the undefined category and are not used. The

reconstructed anomalies using the jet clusters for DJF 2013/2014 reads:

$$A_{rec}(\phi, \lambda, 2013/2014) \quad = \quad \frac{0 \cdot Y_S(\phi, \lambda) + 0 \cdot Y_M(\phi, \lambda) + 11 \cdot Y_T(\phi, \lambda) + 7 \cdot Y_N(\phi, \lambda) + 60 \cdot Y_C(\phi, \lambda)}{90} \tag{2}$$

where $Y(\phi, \lambda)$ are the composite maps of precipitation or temperature anomalies and the subscript stands for the jet type.

To compare the ability of each classification method to reconstruct seasonal anomalies we first compute at each grid point

the correlation between the reconstructed and the observed (ERA-Interim) seasonal anomalies. This simple metric does not

take into account mismatches in the magnitude of the anomalies, unlike the Coefficient of Efficiency (CE) metric described in the next subsection.

2.2.1 Coefficient of Efficiency (CE)

Assume o is the observed quantity and p is the reconstructed quantity. The Coefficient of Efficiency (CE) (Nash and Sutcliffe (1970), see also Bürger (2007); Briffa et al. (1992); Wang et al. (2014)) is given by

$$CE \equiv 1 - \frac{\sum\limits_{t}(o_t - p_t)^2}{\sum\limits_{t}(o_t - \overline{o})^2} \;, \tag{3}$$

where $\overline{(\;)}$ denotes the mean of a quantity and the sum is over all winter seasons (t). If we consider only anomalies and therefore assume that the mean of o and p are both zero, the equation can be simplified such that

$$CE \equiv 2ra - a^2 \;, \tag{4}$$

where a is an amplitude ratio of the standard deviations of the time series,

$$a = \frac{\sigma(p)}{\sigma(o)} \tag{5}$$

and r is the correlation between o and p.

When applied to reconstructions and observations, the CE is a measure of skill in reconstruction that is more restrictive than a simple correlation because it penalizes for both phase and amplitude misfits. For a perfect reconstruction, CE=1. For a reconstruction with the observed variance (a=1) that is correlated with the observed time series of seasonal anomalies at r=0.5, CE=0. For a reconstruction that is perfectly correlated with observations but with twice the observed amplitude, CE=0. In this study, we consider CE>0.25 to indicate a good reconstruction.

If the mean values of the time series $\overline{o}$ and $\overline{p}$ are not zero, then

$$CE \equiv 1 - \frac{\sigma^2(o') + \sigma^2(p') - 2 <o',p'>/N + (\overline{o} - \overline{p})^2}{\sigma^2(o') + \overline{o}^2} \;, \tag{6}$$

where $(\;')$ denotes the anomaly about the mean $\overline{(\;)}$.

In our case, o is the observed climatological anomalies of precipitation/temperature and $\overline{o}$ is by definition zero, while p is the reconstructed anomaly and $\overline{p}$ is not necessarily zero.

2.2.2 Scaling factor/maximising the reconstruction ability

The reconstruction described above assumes that the composite mean precipitation or temperature field for each category is representative of all the days falling into the composites. This assumption works well for variables that follow a Gaussian

distribution and have a small standard deviation. This does not have to be the case, especially for a field such as precipitation, which is known to be skewed. A simple approach that removes the influence of extreme values is to use the median instead
140 of the mean of each composite. A more complex option is to estimate the representative values from a random sample within each category as done by Fereday et al. (2018).

We opt for a different method to estimate the mean values that best represent the composite *a posteriori*, by determining the scaling factor that maximizes the CE. It is not possible to improve the correlation r between the two time series but it is possible to better estimate the amplitude a (i.e. the ratio of standard deviations). We use the approximation of the CE expressed
145 as a function of correlation and amplitude (Eq. 4) and calculate the centre of mass of the CE in the correlation-amplitude (x-y) space. The scaling factor β is obtained by moving the centre of mass in the y-direction so that $a = r$, which maximizes the CE. Thus, equation (1) can be rewritten as:

$$A_{rec}(\phi, \lambda, t) \quad = \quad \beta \quad \sum_i (Y_i(\phi, \lambda) \cdot f_i(t)) \tag{7}$$

Only points with a correlation of above 0.5 (which represents approximately the 1% significance level of a two-tailed t-test
150 with 34 degrees of freedom) are used to calculate the scaling factors; this prevents amplitude biases from points with weak correlations, and thus little skill, from affecting the reconstruction ability of points with higher skill. The values of β are very similar for each reconstructed fields (precipitation, temperature, and wind) and classification methods, with values around 2 (see Table 2).

2.2.3 Uncertainty estimation

155 In order to estimate the uncertainty on the calculation of the CE, the reconstructions of the seasonal anomalies are re-calculated 100 times using different combinations of 35 years of the original time series (bootstrap with replacement). The results are not sensitive to the number of iterations used (i.e. 1000 instead of 100).

3 Results

3.1 Classification anomalies and interannual variability

160 Distinct wind, temperature and precipitation anomalies are associated with each NAO phase, blocking category and jet cluster (Figs. 2-4). Figure 2 shows the zonal wind anomalies for each pattern (colours) and the climatological DJF zonal wind distribution (black contours). Figures 3 and 4 show precipitation and temperature anomalies respectively (colours), and the composite zonal wind (black contours). The panels are placed to allow easy comparison between similar patterns identified by the NAO (left column), blocking (centre column) and the jet clusters (right column). In general, wind and temperature patterns across
165 the different methods (i.e. along the same row) are similar despite the different number of days in the composites for patterns from different metrics, e.g. for the top row, there are on average 25.4 days per season corresponding to the NAO- group, 13.5

days to the GB, and 7.3 days to S-jet (see Table 1). Despite this similarity in the spatial structures of the composites, there can be large differences in the strength of the anomalies for the different classification methods.

The top row of Figures 2-4 shows a clear correspondence between the negative NAO phase, Greenland blocking (GB), and
170 the S-jet cluster, consistent with previous work (Woollings et al., 2010; Madonna et al., 2017). In all three composites the jet is located southwards of its climatological position (Fig. 2). This southerly shifted jet is also zonally oriented, which can be seen in the full composite wind fields shown in black contours in Figure 3. In all three composites, precipitation is enhanced in the jet core and at its ends, e.g. over the Iberian Peninsula (Fig. 3, blue shading). Although the patterns look fairly similar, they differ in intensity, with the highest precipitation anomalies found in the S-jet cluster. When the jet is shifted south (NAO-, GB,
175 S-jet), Greenland is warmer than usual, while northern Europe and the Barents Sea are colder than usual (Fig. 4, shading).

There is less correspondence between the positive NAO phase and any other patterns. As the positive NAO phase has often been referred to as the unblocked or unperturbed state (e.g. Woollings et al., 2008, 2010) it does not resemble any of the blocking patterns. The positive NAO composite has similar wind and temperature anomaly as both the tilted and central jet clusters (Figs. 2 and 4, shadings), with a jet shifted to the north (and tilted) and warmer temperatures over central and Northern
180 Europe. The warming over the Barents Sea observed during the positive NAO phase is seen in both the tilted and central jet composites, with a stronger warming signal in the latter, while the cooling over Greenland is linked to the tilted jet rather than the central jet. Also, precipitation anomalies of the positive NAO phase are more similar to the tilted jet cluster than the central jet cluster, indicative of the large change in precipitation pattern associated with the relatively small shifts in jet position.

Blocking over the Iberian Peninsula (IWB) is associated with a northward shift of the jet (N-jet), while the jet splits during
185 Scandinavian blocking (SBL), resulting in a M-jet cluster, in agreement with Madonna et al. (2017). The blocked region is drier than climatology for both cases, with less precipitation offshore and over the Iberian Peninsula associated with the N-jet cluster and IWB, and drier central Europe during M-jet and SBL (Fig. 3). During IWB southern Europe is colder, while northern Europe is warmer than normal, while during N-jet, only the cold anomaly is evident. During SBL and M-jet, most of Europe is cold and north Scandinavia is warm.

190 The correspondence between the jet and blocking composites extends to seasonal time scales. Figure 5 shows the time series of the occurrence (i.e. the number of days per winter) of each NAO phase, blocking and jet cluster. The time series of the S-jet cluster has a correlation of 0.64 with GB and 0.67 with the negative NAO phase. However, the S-jet is less frequent than the other two, with on average only 7.3 days per winter, compared to 13.5 days of GB and 25.4 days of the negative NAO phase (Table 1). The N-jet occurs on average 8.6 days per winter and IWB 13.0 days (Table 1) and their time series (Fig. 5b) are
195 correlated at 0.71. The M-jet occurs on average 9.7 days, SBL 16.9 days and their time series are correlated at 0.58. The C-jet (10.9 days) and T-jet (12.0 days) are the most frequent jet clusters, however, they show a relatively low correlation with the positive NAO time series (0.30 and 0.45, respectively).

The winter to winter differences in the number of days in each jet cluster and blocking type is large: standard deviations are of similar magnitude to the mean values (Table 1). Moreover, about 40-60% of the days are classified as NAO neutral days,
200 unblocked or undefined (for the jet cluster). The composites of precipitation and temperature for those categories are similar to climatology and therefore their patterns are characterized by almost no anomalies (not shown).

3.2 Seasonal reconstructions

Using the method described in section 2.2, we reconstruct seasonal anomalies of temperature and precipitation for each of the three classification methods: the NAO, blocking, and jet clusters. Based on our definitions, for each classification the average
205 number of days per season used for reconstruction is between 38.1 and 55.4 days (Table 1). Next, we compare our reconstructed seasonal anomalies to the observed anomalies to evaluate the skill of each method in reproducing precipitation and temperature.

3.2.1 Correlation

The ability of each method to reconstruct seasonal weather anomalies varies greatly with location. To assess the ability of the different classification methods to capture seasonal anomalies we calculate for each gridpoint the correlation coefficient
210 between the time series of seasonal reconstructed anomalies and that of the actual anomalies from the ERA-interim data. Figure 6 shows the spatial distribution of this correlation coefficient over Europe for precipitation (left column) and temperature (right column); regions are masked with white dots when the correlation coefficient is below 0.5. The spatial structure of correlations for temperatures is much smoother than that for precipitation; this is consistent with smoother variations in temperature fields relative to precipitation, which varies at much smaller spatial scales. For the reconstructions based on jet clusters, precipitation
215 agrees better (higher correlations) with observations than temperature, in particular in regions of high topography. In these regions it is likely that precipitation depends more uniquely on circulation that temperature does, as temperature can be affected by other mechanisms including cloud cover and land surface feedbacks.

Overall, the correlations between precipitation reconstructions and observations are higher in western Europe and Scandi-navia, and lower in central to south-east Europe. Regions with low correlations also show little seasonal variability (i.e. small
220 seasonal standard deviations, Fig. S1) suggesting that large-scale circulation patterns have less impact on precipitation in these regions. There are relatively minor differences in the correlation of precipitation for the different methods, although correlations over France are noticeably worse in the NAO reconstruction (cf. figures 6a with 6c and e).

The skill of the temperature reconstructions depends greatly on the classification method. Over Spain, the blocking method does substantially better than the NAO and slightly better than the jet clusters. Conversely, the NAO performs much better in a
225 band from 50 to 65°N than the other methods, but substantially worse south of 50°N. This is consistent with the temperature anomalies in Figure 4 - neither positive nor negative NAO is associated with strong temperature anomalies across southern Europe. Winter temperature variability exhibits a southwest-northeast gradient (Fig. S1) and regions with poor correlation skill mostly have low temperature variability.

3.2.2 Coefficient of Efficiency

230 Having shown strong correlations between reconstructed and observed seasonal anomalies for many regions of Europe, we now examine the coefficient of efficiency (CE), which takes into account both the correlation (i.e. the phase) and the magnitude of the reconstructed values relative to observations. The spatial pattern of the CE (Fig. 7) generally follows that of the correlation (Fig. 6); but the absolute values are lower – less than 0.25 across much of the domain. While precipitation and temperature

showed a similar degree of correlation across the region as a whole, the CE is typically substantially lower for temperature

235   than for precipitation (cf. Figs. 6 and 7). Similar results are obtained by bootstrapping (with replacement) 100 time series of 35 years (Fig. S2) showing that the results are robust to time series re-sampling. For purposes of applicability, we focus on land regions over Europe. In the supplementary material (Fig. S3) we show plots that extend westward into the North Atlantic where values of CE are typically higher over the ocean off the west coast of Europe and the differences between the methods become even more evident, e.g. over the North Atlantic the NAO can not reconstruct precipitation anomalies in the 45-50°N

240   latitudinal band.

Given the relatively high correlation between reconstructed and observed anomalies, the low magnitude of the CE suggests that the reconstruction does not reproduce the observed amplitude of the seasonal anomalies well (see Eq. 4). For each gridbox in our domain (land points not grey in Fig. 6) we calculate the correlation and amplitude (ratio of standard deviations) of the reconstructed time series relative to the observed. Figure 8 shows the distribution of these values in correlation-amplitude space

245   for the unscaled data (i.e. that shown in Figure 7) as black contours. The amplitudes are relatively small, with almost all points falling below the $a = r$ line (maximum CE, shown in red). In the following section, we consider methods to maximise the reconstruction skill (CE).

3.3   Maximising the Reconstruction Skill

The black contours in Figure 8 demonstrate that the amplitudes of our reconstructed anomalies tend to be underestimated,

250   despite relatively good correlation. A possible reason for the underestimation is the large variability within each composite. For example for a S-jet day, we expect precipitation to be enhanced over the Iberian Peninsula (cf. Fig. 3), but the exact location where the precipitation peaks, which is likely linked to the passage of a cyclone, varies from case to case (i.e. cyclones do not have the exact same path). Therefore, at each grid point the standard deviation within each composite can be quite large (not shown).

255   A simple method to improve our reconstructions is to apply a scaling factor to the reconstructed anomalies to increase their amplitude (see Eq. 7). Our method for determining a scaling factor to maximise the CE is discussed in section 2.2.2; note that we only scale gridboxes that have a correlation greater than 0.5. The effect of this scaling can be seen in Figure 8, contrasting between the unscaled (black contours) and scaled (blue shading) values. In all cases the centroid of the scaled points now lies on the maximum CE, the $r = a$ line (in red). Table 2 gives the scaling factors for each of the classification techniques and

260   variables. Overall, the amplitudes for all variables (precipitation, temperature, and wind) are underestimated by half ($\beta \sim 2$), with scaling factors comparable between the methods. The CE skill score for unscaled ($\beta = 1$) precipitation and temperature reconstruction are shown in Figure 7, while those for the scaled reconstructions are shown in Figure 9. The CEs recalculated using the scaling factors substantially improve seasonal temperature reconstructions over most part of the domain, while the improvements are more localized for precipitation (cf. Fig. 7 with Fig. 9).

265   For four locations across Europe (Lisbon, Berlin, Bergen and London, indicated by X1-4 in Fig. 6a-b) we consider the temporal evolution of the seasonal reconstructions. Figure 10 shows the observed precipitation and temperature anomalies (black, ERA-Interim) and reconstructions using the scaling factors for these locations. Precipitation can be well reconstructed

for Lisbon (10a), Bergen (10e) and London (10g) using the five jet clusters (red) or the three blocking categories (blue), with time series of comparable amplitudes and high correlation values with ERA-Interim, while only for Bergen the NAO
270 (grey) leads to comparable reconstructions. For Berlin (10c), none of the methods shows skill in reconstructing precipitation (low correlations). However, we can also see that at this location the precipitation anomalies (black curve) are small. The correlations between the observed and reconstructed temperature for the chosen locations are comparable across the methods, but not for Lisbon (10 b), where the reconstruction using the NAO shows no skill. For the five jet clusters (red), the reconstructed temperature time series have similar amplitudes as the observed (black), while for the reconstruction with the three blocking
275 categories (blue) the amplitudes are often still underestimated.

Finally, in addition to the scaling of the reconstructions, we briefly consider whether alternative methods of defining the anomalies for each classification could be improved. For example, whether the NAO reconstruction would be improved if we construct the anomaly patterns based on linear regression of positive NAO and negative NAO values (treated separately to take account of the skewness of the NAO e.g. Nakamura and Wallace, 1991). This method produces positive and negative NAO
280 precipitation and temperature anomaly patterns very similar to those in Figure 3 and 4 but there is no improvement in the skill (correlation or amplitude) of the reconstruction (not shown).

4 Discussion and Conclusion

In this study, we investigated how well seasonal anomalies in precipitation and temperature over Europe can be reconstructed based on the frequency of circulation patterns defined using three different classifications: the NAO index, blocking, and the
285 configuration of the North Atlantic jet stream.

In general, precipitation is better reconstructed than temperature. This is not surprising, as precipitation changes in winter are directly linked to the propagation of storms (Hawcroft et al., 2012; Pfahl et al., 2014), which is to first-order set by the large-scale circulation. Nevertheless, there is still considerable spread in precipitation within each circulation pattern due to variations in the exact path and intensity of individual storms (Pfahl and Sprenger, 2016). For temperatures, the circulation
290 influences the horizontal and vertical advection of air, but other factors, such as radiative forcing (e.g. clear vs. cloudy, Trigo et al., 2004), soil moisture coupling (Fischer et al., 2007) and snow-albedo feedback may also be important.

The skill of the seasonal reconstructions using the three different classifications depends very much on the region and the climate variable of interest. For example, for Scandinavia (e.g., Bergen, Norway), the three reconstruction methods show comparable skill for precipitation, while for Lisbon, the jet clusters perform much better. The NAO has overall lower skill for
295 precipitation but provides the most skilful reconstructions of temperature in Northern Europe. There are also regions, such as central Europe (e.g., Berlin, Germany), where none of the methods provides skilful reconstruction of either the precipitation or temperature patterns.

All the reconstructions underestimate the amplitude of the observed precipitation and temperature anomalies. This is why the skill score (coefficient of efficiency, CE) in many regions in Europe is low, despite relatively high correlations between the
300 reconstruction and reanalysis data. The CE skill score for precipitation and temperature reconstructions for all classification

methods improves when we introduce a scaling factor (β) of $\sim$2 to the composite mean anomaly patterns. This could indicate that mean composite values of the seasonal anomalies used for the reconstruction might not be representative of the distribution (e.g. if the distribution of the variable to be reconstructed is skewed) or that extreme values have large impacts on the seasonal signal.

305 No classification metric emerges as providing a better reconstruction of both precipitation and temperature across all regions. One might expect the skill of a reconstruction to improve as the number of basis functions (patterns) increases. We would thus expect to capture more interannual variability and have better skills using jet clusters (five patterns) than either blocking (three patterns) or the NAO (only two). While this is mostly true for precipitation, it does not always apply for temperature. In northeast Europe and Scandinavia, the NAO performs better than the other classification methods. A possible explanation

310 might lie in the different domains used to define the classification patterns: the region used for the jet is much smaller than the one for the NAO (Fig. 1), with an eastern limit at the Greenwich meridian (0 °). Thus, it is not so surprising that the NAO captures better the seasonal anomalies over central and eastern Europe, as circulation variability over these regions is integrated into the NAO definition. In fact, it is rather remarkable that so much of the seasonal precipitation and temperature signal over Europe can be inferred just by knowing the circulation over the North Atlantic ocean (i.e. using the jet clusters).

315 One might also expect the skill of a reconstruction to depend on the amount of information included. In our reconstructions, we include only the days that distinctly belong to a certain basis function within each classification method. Interestingly, this sums to about half of the total days in the record, regardless of method. One could of course use more information, but this does not necessarily improve the reconstruction skill. For example, using 30 MSLP patterns (i.e., many more basis functions) and all days to reconstruct winter precipitation, Fereday et al. (2018) obtain correlations of about r=0.8 with observations for northern

320 and southern Europe. Averaging our results over the same regions, we obtain lower but comparable correlations between observed and reconstructed precipitation (0.54-0.78 for the northern region (Fig. S4, blue) and 0.53-0.66 for the southern region (Fig. S4, red)). Moreover, our study shows that reconstructions are not uniformly skilful across Europe, meaning that area-averaged skill scores hide important spatial information. It also suggests that some regions have a larger potential for seasonal (c.f. Kim et al., 2016; Scaife et al., 2014) and decadal predictions (e.g. Smith et al., 2019).

325 Finally, the ability to reconstruct precipitation and temperature might be affected by extreme events. For example, if a single extreme precipitation event is responsible for the lion's share of precipitation in a specific season, we would not expect a skilful reconstruction of the seasonal anomaly using the mean precipitation signals associated with each basis function. The effect of extreme events varies regionally, as shown for summer temperatures by Röthlisberger et al. (2020). It would therefore be interesting to investigate to what extent extreme events influence the seasonal precipitation and temperature anomalies over

330 Europe and how these events are related to circulation anomalies.

Code and data availability. ERA-Interim data can be downloaded from the ECMWF page https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/. The NAO index was downloaded from NOAA ftp://ftp.cpc.ncep.noaa.gov/cwlinks/norm.daily.nao.index.b500101.current.ascii.

The method to identify blocking is described in Scherrer et al. (2006) and for jet clusters in Madonna et al. (2017). The winter time series used in this study are available at DOI: 10.5281/zenodo.4011886.

335 *Author contributions.* DB and EM designed the study. EM performed most of the analysis. All authors contributed to the interpretation and discussion of the results and the writing of the paper.

Competing interests. CL is a member of the editorial board of the journal.

Acknowledgements. EM and CL acknowledge funding from the Research Council of Norway (Nansen Legacy grant #276730). RHW received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant
340 Agreement 797961, and from the Tamaki Foundation. ECMWF and NOAA are acknowledged for providing the ERA-Interim reanalyses and NAO data, respectively.

References

Athanasiadis, P. J., Wallace, J. M., and Wettstein, J. J.: Patterns of wintertime jet stream variability and their relation to the storm tracks, J. Atmos. Sci., 67, 1361–1381, 2010.

Briffa, K. R., Jones, P. D., and Schweingruber, F. H.: Tree-Ring Density Reconstructions of Summer Temperature Patterns across Western North America since 1600, J. Climate, 5, 735–754, 1992.

Bürger, G.: On the verification of climate reconstructions, Climate Past, 3, 397–409, 2007.

Corte-Real, J., Zhang, X., and Wang, X.: Large-scale circulation regimes and surface climatic anomalies over the Mediterranean, Int. J. Climatol., 15, 1135–1150, 1995.

Cortesi, N., Torralba, V., González-Reviriego, N., Soret, A., and Doblas-Reyes, F. J.: Characterization of European wind speed variability using weather regimes, Climate Dyn., 53, 4961–4976, 2019.

Davini, P., Cagnazzo, C., Fogli, P. G., Manzini, E., Gualdi, S., and Navarra, A.: European blocking and Atlantic jet stream variability in the NCEP/NCAR reanalysis and the CMCC-CMS climate model, Climate Dyn., 43, 71–85, 2014.

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. Roy. Meteor. Soc., pp. 553–597, 2011.

Dorrington, J. and Strommen, K.: Jet Speed Variability Obscures Euro-Atlantic Regime Structure, Geophys. Res. Lett., 47, e2020GL087 907, https://doi.org/10.1029/2020GL087907, 2020.

Fereday, D., Chadwick, R., Knight, J., and Scaife, A. A.: Atmospheric Dynamics is the Largest Source of Uncertainty in Future Winter European Rainfall, J. Climate, 31, 963–977, 2018.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D., and Schär, C.: Soil moisture–atmosphere interactions during the 2003 European summer heat wave, J. Climate, 20, 5081–5099, 2007.

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H.: Balancing Europe's wind-power output through spatial deployment informed by weather regimes, Nature Climate Change, 7, 557–562, 2017.

Hawcroft, M., Shaffrey, L., Hodges, K., and Dacre, H.: How much Northern Hemisphere precipitation is associated with extratropical cyclones?, Geophys. Res. Lett., 39, 2012.

Hoskins, B. J. and Ambrizzi, T.: Rossby wave propagation on a realistic longitudinally varying flow, J. Atmos. Sci., 50, 1661–1671, 1993.

Hurrell, J. W.: Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation, Science, 269, 676–679, 1995.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M.: An overview of the North Atlantic Oscillation, Geophys. Monograph-American Geophys. Union, 134, 1–36, 2003.

Jerez, S., Trigo, R. M., Vicente-Serrano, S. M., Pozo-Vázquez, D., Lorente-Plazas, R., Lorenzo-Lacruz, J., Santos-Alamillos, F., and Montávez, J.: The impact of the North Atlantic Oscillation on renewable energy resources in southwestern Europe, J. Appl. Meteorol. Climatol., 52, 2204–2225, 2013.

Kim, G., Ahn, J.-B., Kryjov, V. N., Sohn, S.-J., Yun, W.-T., Graham, R., Kolli, R. K., Kumar, A., and Ceron, J.-P.: Global and regional skill of the seasonal predictions by WMO Lead Centre for Long-Range Forecast Multi-Model Ensemble, Int. J. Climatol., 36, 1657–1675, 2016.

Lavaysse, C., Vogt, J., Toreti, A., Carrera, M. L., and Pappenberger, F.: On the use of weather regimes to forecast meteorological drought over Europe, Nat. Hazards Earth Syst. Sci., 18, 3297–3309, 2018.

380 Lesk, C., Rowhani, P., and Ramankutty, N.: Influence of extreme weather disasters on global crop production, Nature, 529, 84–87, 2016.

Madonna, E., Li, C., Grams, C. M., and Woollings, T.: The link between eddy–driven jet variability and weather regimes in the North Atlantic–European sector, Q. J. Roy. Meteor. Soc., 143, 2960–2972, https://doi.org/10.1002/qj.3155, 2017.

Madonna, E., Li, C., and Wettstein, J. J.: Suppressed eddy driving during southward excursions of the North Atlantic jet on synoptic to seasonal time scales, Atmos. Sci. Lett., 20, e937, 2019.

385 Molteni, F., Tibaldi, S., and Palmer, T.: Regimes in the wintertime circulation over northern extratropics. I: Observational evidence, Q. J. Roy. Meteor. Soc., 116, 31–67, 1990.

Nakamura, H. and Wallace, J. M.: Skewness of low-frequency fluctuations in the tropospheric circulation during the Northern Hemisphere winter, J. Atmos. Sci., 48, 1441–1448, 1991.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion of principles, J. Hydrol., 10, 282–290,
390   1970.

Pfahl, S. and Sprenger, M.: On the relationship between extratropical cyclone precipitation and intensity, Geophys. Res. Lett., 43, 1752–1758, 2016.

Pfahl, S. and Wernli, H.: Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-) daily time scales, Geophys. Res. Lett., 39, 2012.

395 Pfahl, S., Madonna, E., Boettcher, M., Joos, H., and Wernli, H.: Warm conveyor belts in the ERA-Interim dataset (1979–2010). Part II: Moisture origin and relevance for precipitation, J. Climate, 27, 27–40, 2014.

Qian, B., Corte-Real, J., and Xu, H.: Is the North Atlantic Oscillation the most important atmospheric pattern for precipitation in Europe?, J. Geophys. Res.: Atmos., 105, 11 901–11 910, 2000.

Rex, D. F.: Blocking action in the middle troposphere and its effect upon regional climate, Tellus, 2, 275–301, 1950.

400 Rogers, J. C.: North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe, J. Climate, 10, 1635–1647, 1997.

Röthlisberger, M., Sprenger, M., Flaounas, E., Beyerle, U., and Wernli, H.: The substructure of extremely hot summers in the Northern Hemisphere, Weather Climate Dyn., 1, 45–62, 2020.

Santos, J. A., Belo-Pereira, M., Fraga, H., and Pinto, J. G.: Understanding climate change projections for precipitation over western Europe
405   with a weather typing approach, J. Geophys. Res. Atmos., 121, 1170–1189, 2016.

Scaife, A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R., Dunstone, N., Eade, R., Fereday, D., Folland, C., Gordon, M., et al.: Skillful long-range prediction of European and North American winters, Geophys. Res. Lett., 41, 2514–2519, 2014.

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C., and Appenzeller, C.: Two–dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region, Int. J. Climatol., 26, 233–249, 2006.

410 Sillmann, J. and Croci-Maspoli, M.: Present and future atmospheric blocking and its impact on European mean and extreme climate, Geophys. Res. Lett., 36, 2009.

Smith, D., Eade, R., Scaife, A. A., Caron, L.-P., Danabasoglu, G., DelSole, T., Delworth, T., Doblas-Reyes, F., Dunstone, N., Hermanson, L., et al.: Robust skill of decadal climate predictions, npj Climate Atmos. Sci., 2, 1–10, 2019.

Sousa, P. M., Trigo, R. M., Barriopedro, D., Soares, P. M., Ramos, A. M., and Liberato, M. L.: Responses of European precipitation distri-
415   butions and regimes to different blocking locations, Climate Dyn., 48, 1141–1160, 2017.

Treidl, R., Birch, E., and Sajecki, P.: Blocking action in the Northern Hemisphere: A climatological study, Atmos.-ocean, 19, 1–23, 1981.

Trigo, R., Trigo, I., DaCamara, C., and Osborn, T.: Climate impact of the European winter blocking episodes from the NCEP/NCAR Reanalyses, Climate Dyn., 23, 17–28, 2004.

Vautard, R.: Multiple weather regimes over the North Atlantic: Analysis of precursors and successors, Mon. Weather Rev., 118, 2056–2081, 1990.

Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E., and Rajaratnam, B.: Evaluating climate field reconstruction techniques using improved emulations of real-world conditions, Clim. Past, 10, 1–19, 2014.

Wanner, H., Brönnimann, S., Casty, C., Gyalistras, D., Luterbacher, J., Schmutz, C., Stephenson, D. B., and Xoplaki, E.: North Atlantic Oscillation – concepts and studies, Surv. Geophys., 22, 321–381, 2001.

Wettstein, J. J. and Wallace, J. M.: Observed patterns of month-to-month storm-track variability and their relationship to the background flow, J. Atmos. Sci., 67, 1420–1437, 2010.

Wirth, V., Riemer, M., Chang, E. K., and Martius, O.: Rossby wave packets on the midlatitude waveguide - A review, Mon. Weather Rev., 146, 1965–2001, 2018.

Woollings, T., Hoskins, B., Blackburn, M., and Berrisford, P.: A new Rossby wave–breaking interpretation of the North Atlantic Oscillation, J. Atmos. Sci., 65, 609–626, 2008.

Woollings, T., Hannachi, A., and Hoskins, B.: Variability of the North Atlantic eddy-driven jet stream, Q. J. Roy. Meteor. Soc., 136, 856–868, 2010.

Table 1. Average and (in brackets) standard deviation of the number of days per winter in each category. The positive (negative) NAO phase is defined as days with NAO values above (below) 0.5 of the time series standard deviation (see Section 2 for more information), and the remaining days are considered as neutral. We differentiate between blocking over Greenland (GB), over the Iberian Peninsula (IWB), Scandinavia (SBL) and non-blocked days (NB). The categories for the jet are: south (S-jet), tilt (T-jet), north (N-jet), mixed (M-jet) and central (C-jet). We also reported the number of days in the undefined category using the jet classification (undef).

	NAO		blocking		jet
NAO-	25.4 (16.8)	GB	13.5 (11.4)	S-jet	7.3 (7.7)
NAO+	30.0 (13.9)			T-jet	12.0 (9.8)
		IWB	13.0 (12.3)	N-jet	8.6 (7.5)
		SBL	16.9 (9.3)	M-jet	9.7 (7.4)
				C-jet	10.9 (11.0)
neutral	34.6 (11.1)	NB	51.9 (14.0)	undef	41.4 (10.0)

Table 2. Scaling factors β for two NAO phases, three blocking categories and five jet clusters for gridpoints over Europe (only land) and only with correlation larger than 0.5.

scaling factor β	precipitation	temperatures	zonal wind
two NAO phases	1.8	1.9	1.9
three blocking types	1.9	2.2	1.8
five jet clusters	1.8	2.2	1.7
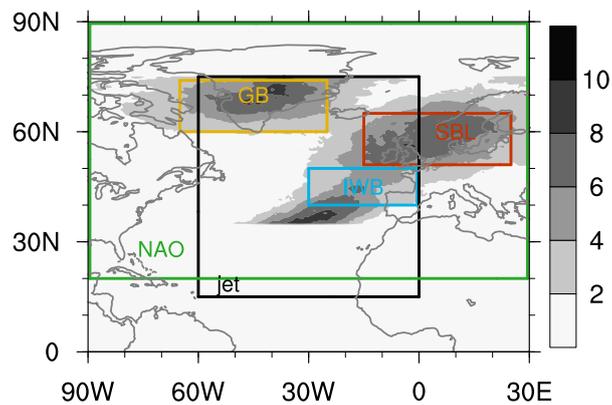
Figure 1. Blocking climatology (shading, as % of time). The green and black boxes show the domain used for the definition of the NAO and jet clusters, respectively. The orange, blue, and red boxes denote the regions used for Greenland blocking (GB), Iberian wave breaking (IWB), and Scandinavian Blocking (SBL), respectively.
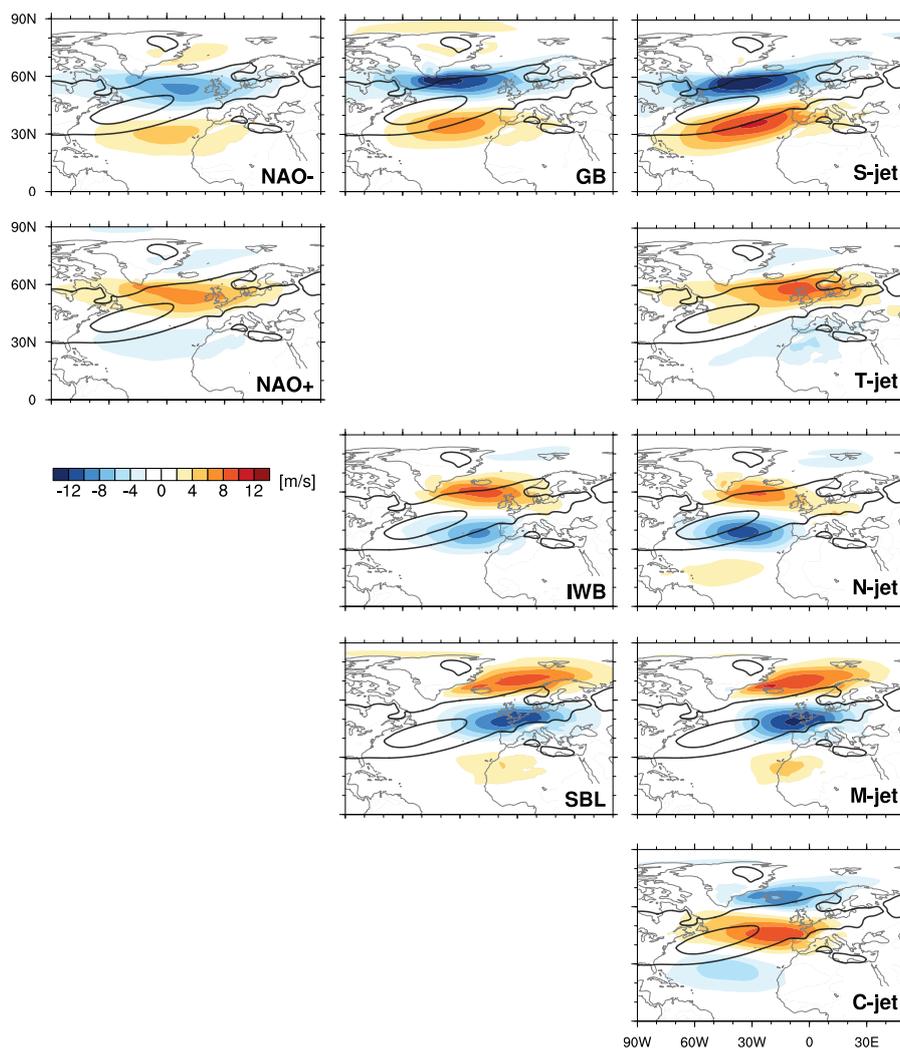
Figure 2. Zonal wind anomalies (shading, in $m\,s^{-1}$) at 850 hPa for two NAO phases (1st column), three blocking categories (2nd column) and five jet clusters (3rd column). Black contours show the climatological zonal wind at 850 hPa (contours at 5 and 10 $m\,s^{-1}$). The figure is organised such that maps in the same row represent similar circulation patterns identified by more than one method (NAO, blocking, jet regime).
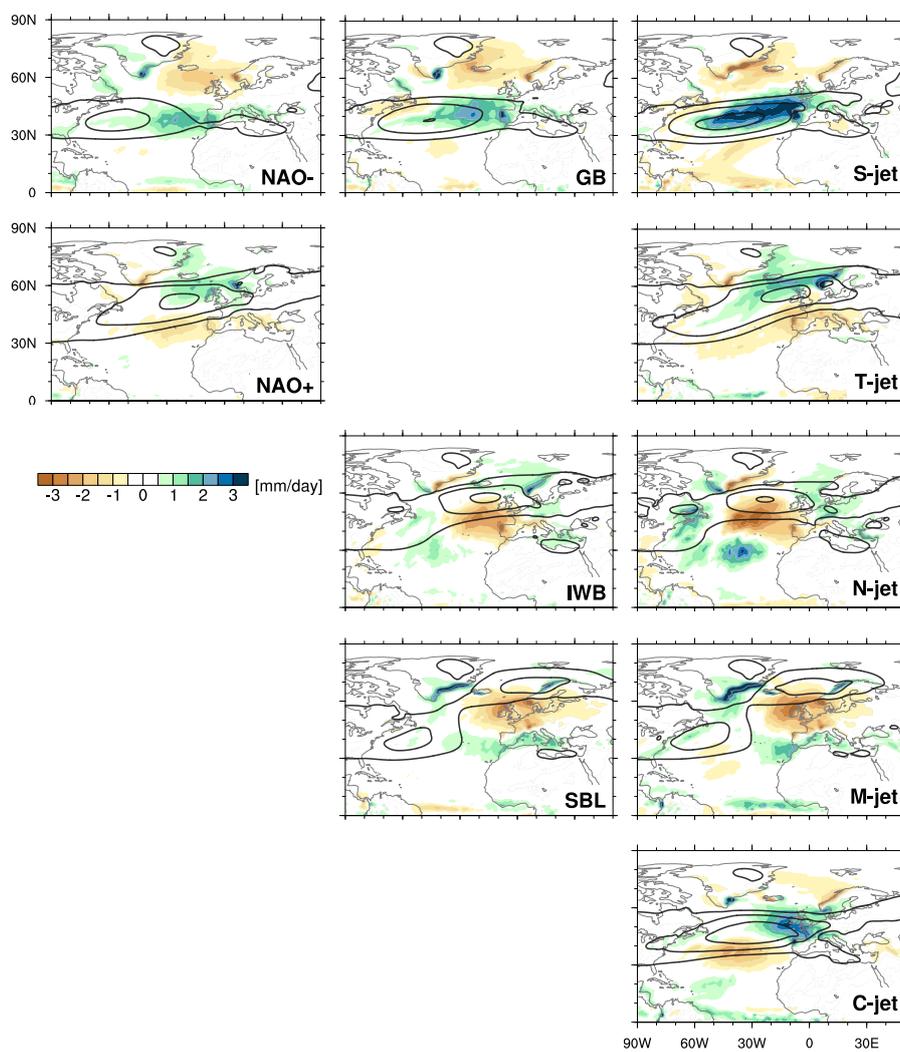
Figure 3. Precipitation anomalies (shading, in $\mathrm{mm\,day^{-1}}$) and zonal wind at $850\,\mathrm{hPa}$ (contours at 5, 10, and $15\,\mathrm{m\,s^{-1}}$) for two NAO phases (1st column), three blocking categories (2nd column) and five jet clusters (3rd column).
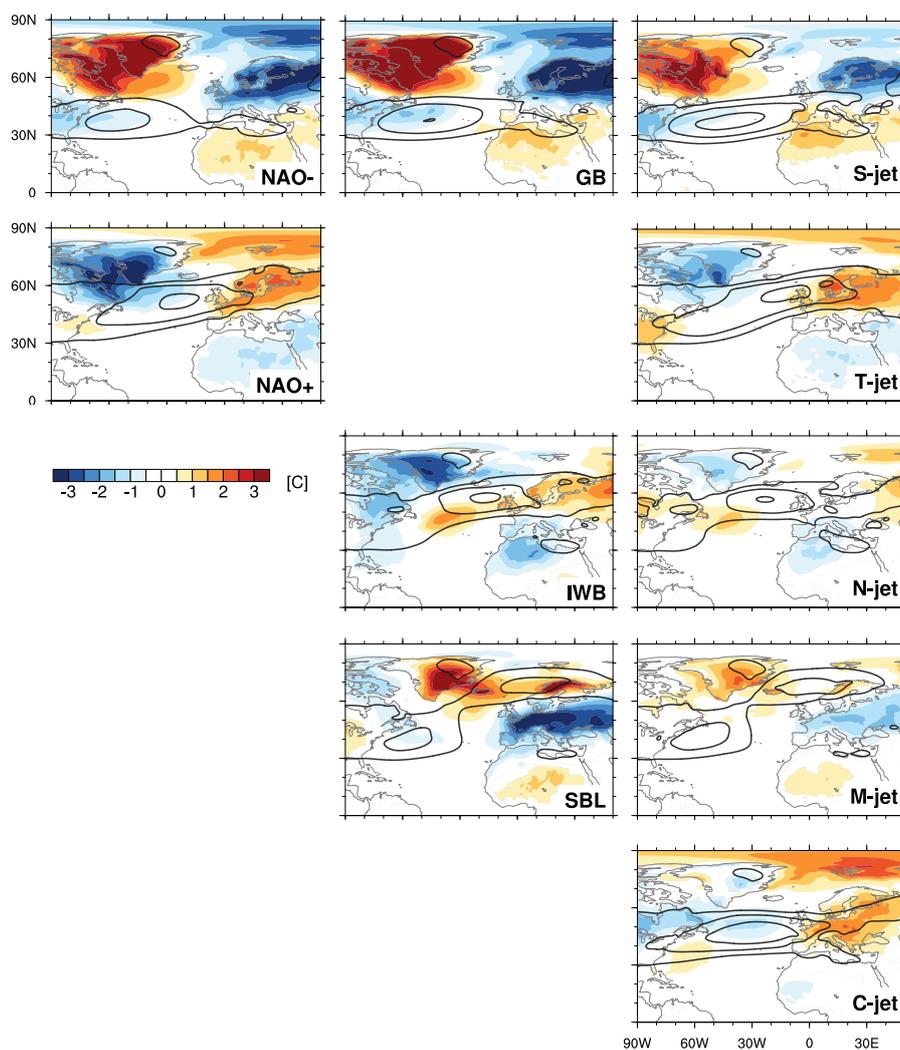
Figure 4. Two metre temperature anomalies (shading, in °C) and zonal wind at 850 hPa (black contours at 5, 10, and 15 m s^{-1}) for two NAO phases (1st column), three blocking categories (2nd column) and five jet clusters (3rd column).
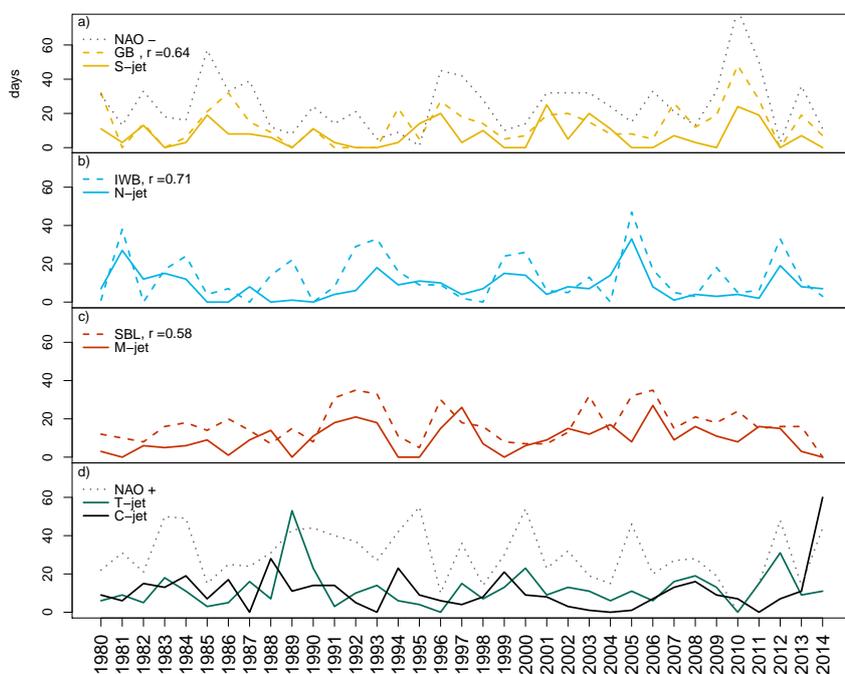
Figure 5. Time series of the number of days per winter for different categories. (a) negative NAO phase (NAO-), Greenland blocking (GB) and S-jet, (b) blocking over Iberia (IWB) and N-jet, (c) blocking over Scandinavia (SBL) and M-jet (d) positive NAO phase (NAO+), T-jet and C-jet. For a-c the correlation (*r*) between each time series of blocking and jet is shown in the plot. The negative NAO phase has a correlation of 0.67 with S-jet, while the positive NAO phase has correlations of 0.45 with T-jet, 0.30 with C-jet and 0.31 with N-jet. The year denotes the December-February period; e.g. 2013 is the average for December 2012 to February 2013.
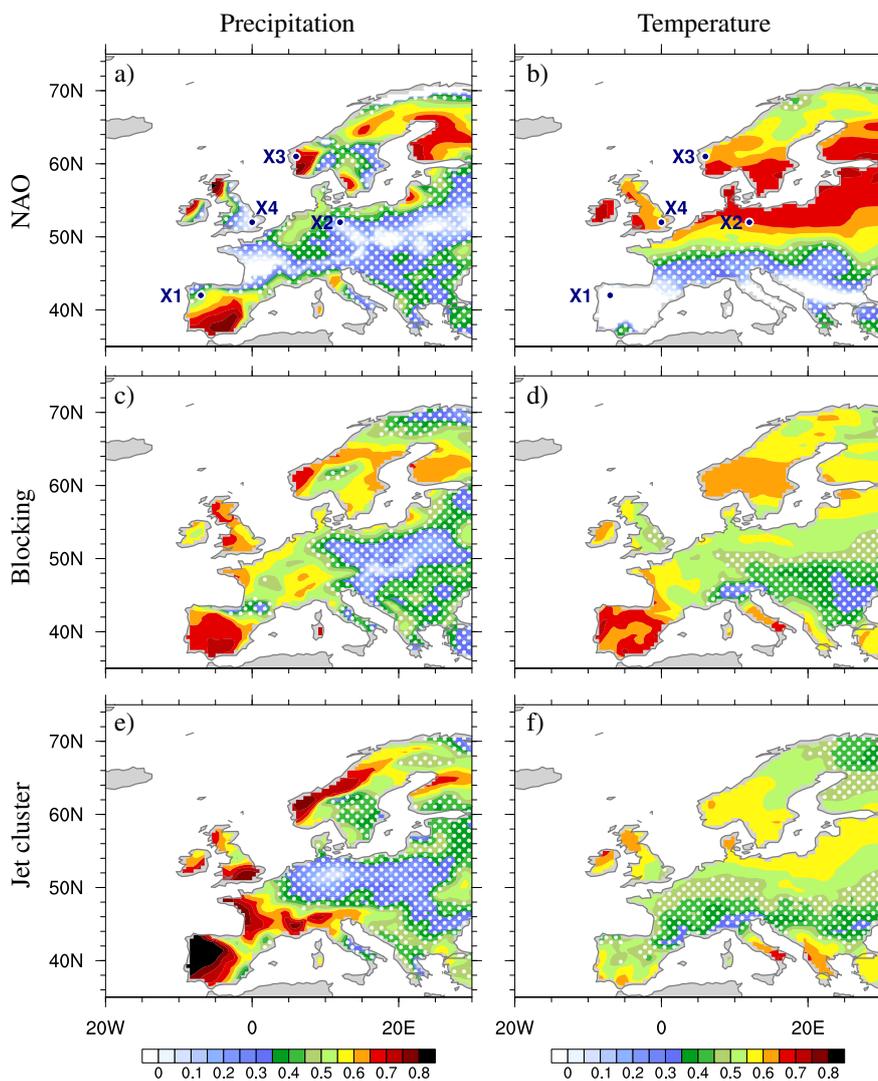
Figure 6. Correlation coefficient between seasonal anomalies of observed and reconstructed precipitation (left) and temperature (right) over Europe for two NAO phases (a-b), three blocking categories (c-d) and five jet clusters (e-f). White dots mark regions with correlation below 0.5. The blue dots labeled X1-4 in (a-b) indicate the four locations shown in Figure 10.
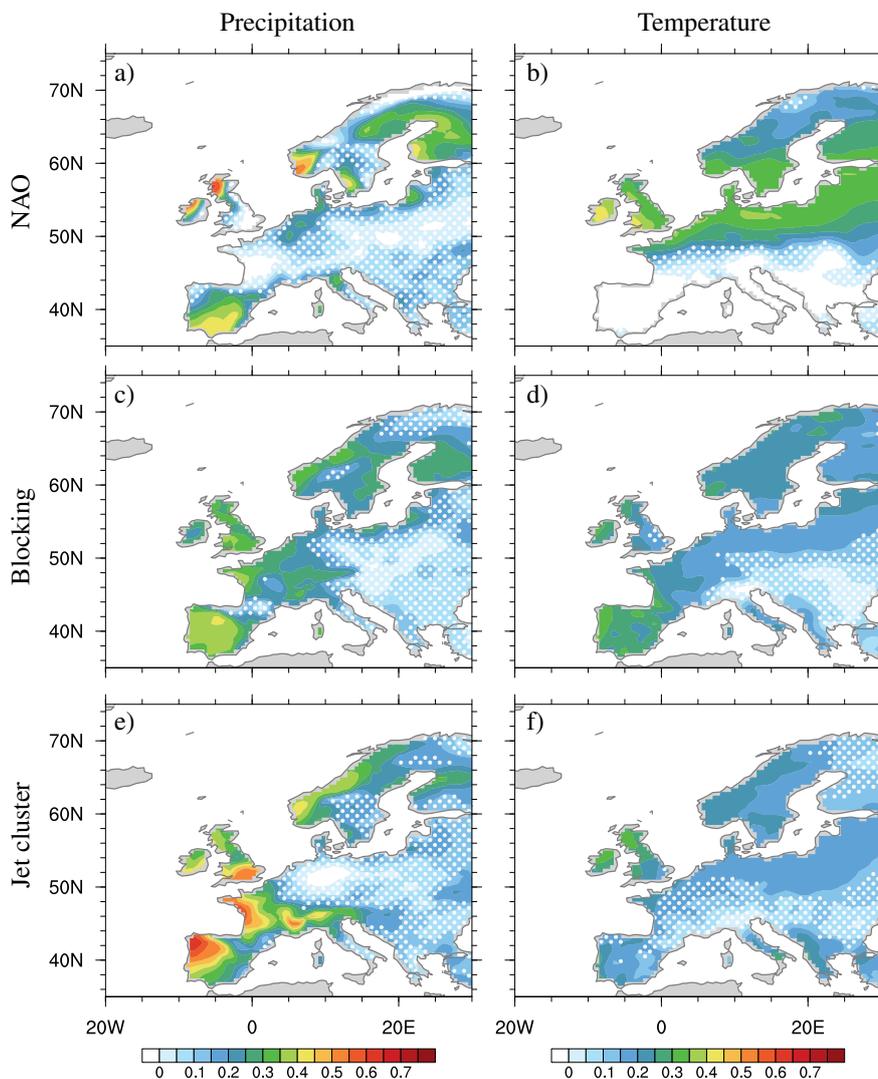
Figure 7. Coefficient of efficiency (CE) for Europe for precipitation (left) and temperature (right) for two NAO phases (a-b), there blocking categories (c-d), and five jet clusters (e-f). White dots mark regions with correlation below 0.5 (as in Figure 6).
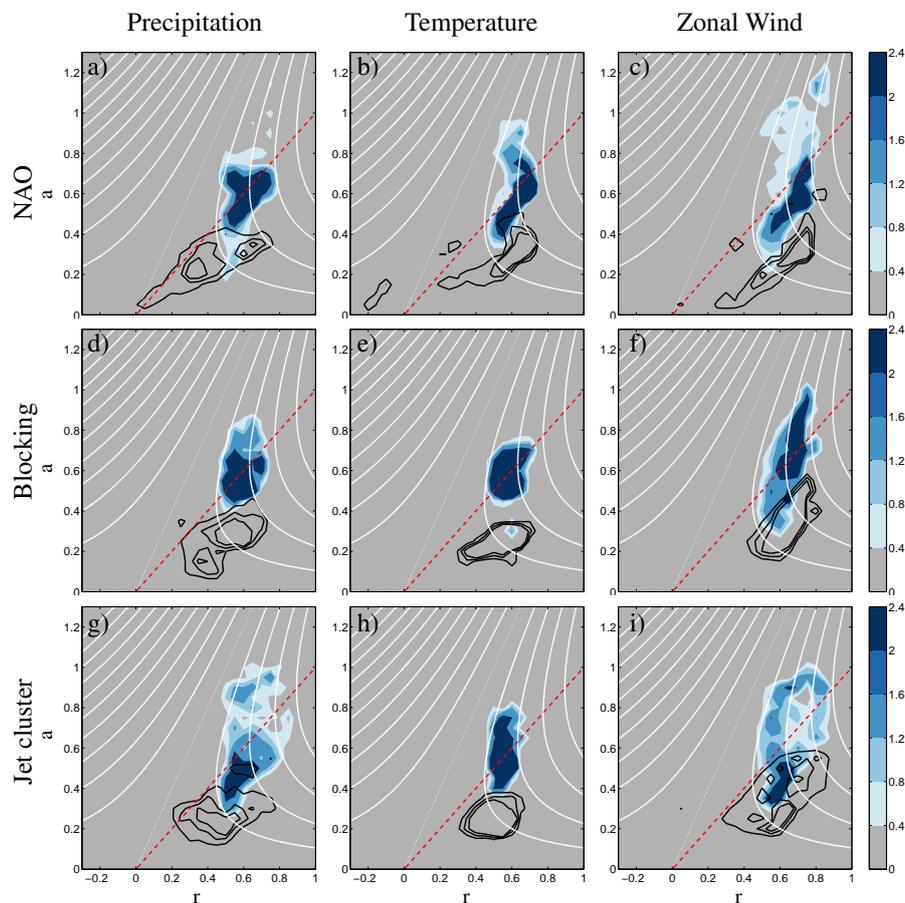
Figure 8. Frequency distribution of correlation (r) versus amplitude (a) of the Coefficient of Efficiency (CE) over Europe (land only) for precipitation (first column), temperatures (second column) and zonal wind (third column) for two NAO phases (a-c), three blocking categories (d-f) and five jet clusters (g-i). Black contours (0.8, 1.6 and 2.4%) are for unscaled reconstructed time series, while the shading are the shifted (scaled) values applied only to grid points with correlation greater than 0.5. Data are normalized by the number of points in each distribution and units are given in percent. White lines are CE values following equation (4) in 0.2 intervals (0 line dotted), while the red line shows the maximisation of CE as function of r (i.e. the $r = a$ line).
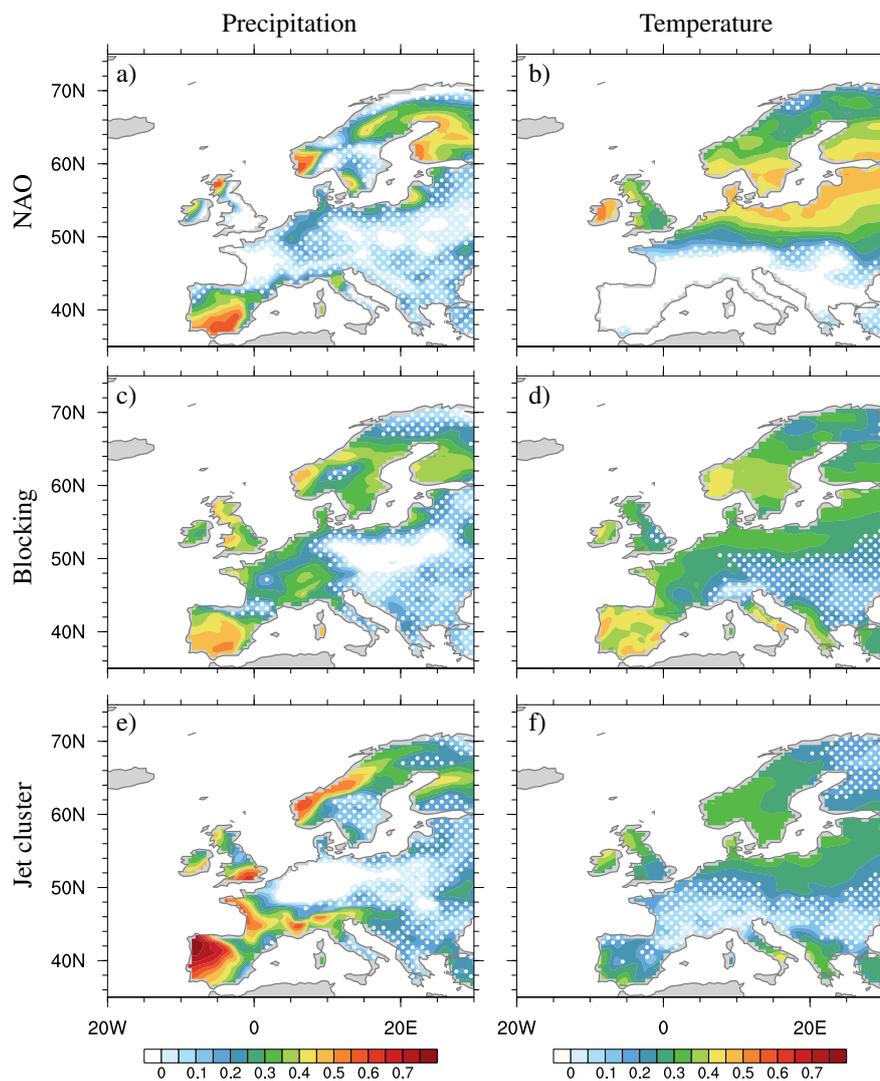
Figure 9. Scaled CE for Europe for precipitation (left) and temperature (right) for two NAO phases (a-b), there blocking categories (c-d) and five jet clusters (e-f). Scaling factors are calculated using only points with correlation $r > 0.5$ (where dots marks regions with $r < 0.5$). The scaling coefficient are shown in Table 2.
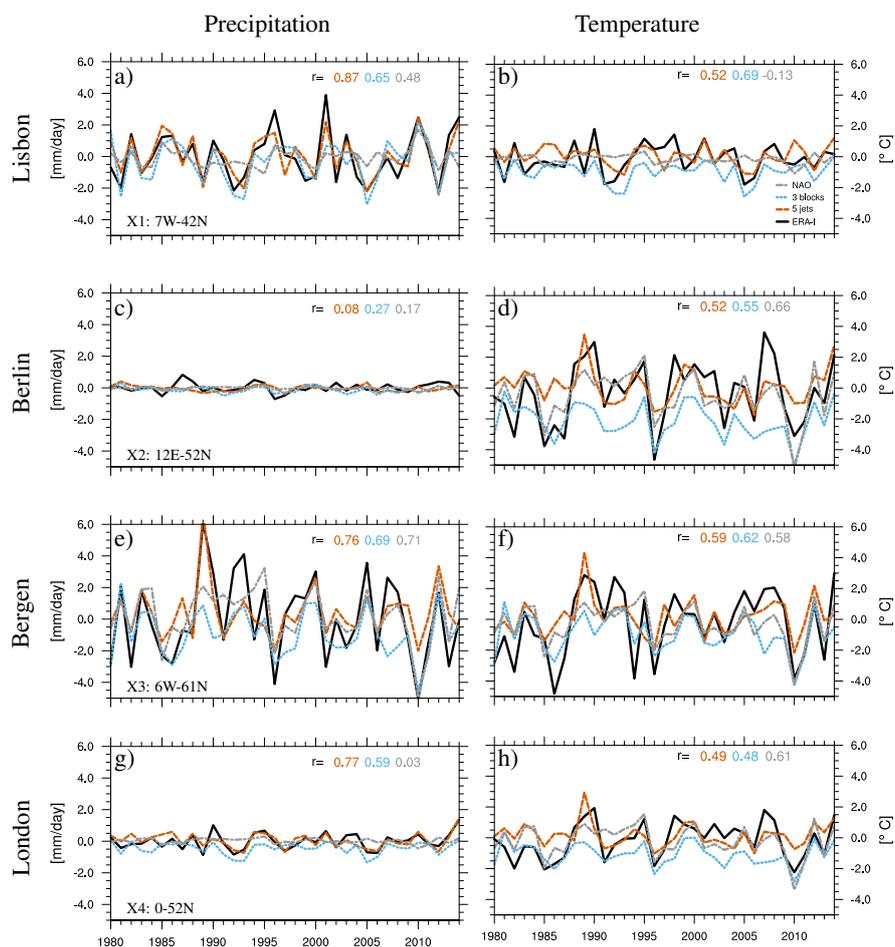
Figure 10. Time series of DJF precipitation (left, in $\mathrm{mm\,day^{-1}}$, averaged over a season) and temperatures anomalies (right, in °C) for four locations shown in Figure 6: Lisbon (X1, a-b), Berlin (X2, c-d), Bergen (X3 e-f), and London (X4, g-h). Correlation (r) between the observed anomalies (ERA-Interim, in black) and the reconstructed anomalies (with scaled amplitudes β) for the different methods are shown using the colour legend in (b). The same y-scale is used to highlight the large differences in magnitude and variability across the different locations.