Reconstructing Late Holocene North Atlantic atmospheric circulation changes using functional paleoclimate networks

Jasper G. Franke[1,2], Johannes P. Werner[3], and Reik V. Donner[1]

[1]Potsdam Institute for Climate Impact Research, Telegrafenberg A31, 14473 Potsdam, Germany
[2]Department of Physics, Humboldt University, Newtonstraße 15, 12489 Berlin, Germany
[3]Bjerknes Centre for Climate Research and Department of Earth Science, University of Bergen, Postboks 7803, 5020 Bergen, Norway

Correspondence to: Jasper G. Franke (jasper.franke@pik-potsdam.de)

Abstract. Obtaining reliable reconstructions of long-term atmospheric circulation changes in the North Atlantic region presents a persistent challenge to contemporary paleoclimate research, which has been addressed by a multitude of recent studies. In order to contribute a novel methodological aspect to this active field, we apply here evolving functional network analysis, a recently developed tool for studying temporal changes of the spatial co-variability structure of the Earth's climate system, to a set of Late Holocene paleoclimate proxy records covering the last two millenia. The emerging patterns obtained by our analysis are related to long-term changes in the dominant mode of atmospheric circulation in the region, the North Atlantic Oscillation (NAO). By comparing the time-dependent inter-regional linkage structures of the obtained functional paleoclimate network representations to a recent multi-centennial NAO reconstruction, we identify co-variability between Southern Greenland, Svalbard and Fennoscandia as being indicative of a positive NAO phase, while connections from Greenland and Fennoscandia to Central Europe are more pronounced during negative NAO phases. By drawing upon this correspondence, we use some key parameters of the evolving network structure to obtain a qualitative reconstruction of the NAO long-term variability over the entire Common Era (last 2000 years) using a linear regression model trained upon the existing shorter reconstruction.

1 Introduction

The increasing availability of high-resolution paleoclimate archives and resulting proxy records allows to not only study local climate variability before the beginning of the instrumental period, but also associated spatial structures at least at a regional level. Corresponding studies have commonly been performed using linear multivariate statistical methods like empirical orthogonal function (EOF) analysis (Gouirand et al., 2008; Mann et al., 1998) or, more recently, Bayesian Hierarchical Modeling (Luterbacher et al., 2016). However, at the conceptual level, many of the classical statistical approaches have considerable problems in analyzing paleoclimate data. On the one hand, traditionally used estimators are often inappropriate for coping with spatially sparse and unevenly sampled time series. On the other hand, the appealing alternative of data interpolation can lead to a systematic bias and large uncertainties in the resulting reconstructions of spatial patterns of past climate variability (Rehfeld et al., 2011). Furthermore, many previously applied methods rely on some kind of linearity and/or orthogonality assumption, which might result in some unrealistic representation of the climatic processes or phenomena under study.

Some of the aforementioned challenges can be (at least partially) addressed by the concept of functional climate networks (Tsonis et al., 2006; Donner et al., 2017), a recently developed nonlinear approach to studying climate dynamics that can also be employed for evaluating spatial co-variability among paleoclimate archives (Rehfeld et al., 2013). Here, each time series from a set of climate observations associated with different geographical locations is represented as a node of an abstract network embedded in geographical space. Pairs of such nodes are connected by links if the observed dynamics is sufficiently similar, which is referred to as *functional connectivity* to highlight that similar, mutually dependent physical processes are commonly reflected by statistical co-variability. Climate networks provide an intuitive way to quantitatively account for the full complexity of co-variability and teleconnection patterns. Furthermore, instead of considering a spatially homogeneous data coverage, they simply ignore regions without data, which is particularly important in the case of sparse paleoclimate data.

Beyond the viewpoint of time-independent or average spatial co-variability patterns, evolving functional networks are constructed from the available data covering different time windows and thus allow for studying the evolution of such spatial patterns in time. While evolving functional climate networks have become a widespread tool to analyze modern climate data (Donner et al., 2017; Radebach et al., 2013), applications to paleoclimate data sets have been much less common so far (Rehfeld et al., 2013; McRobie et al., 2015; Oster and Kelley, 2016). Although such network representations rely on the (potentially questionable) assumption that the underlying spatio-temporally continuous climate system has been coarse-grained and represented in some meaningful way by the considered data sets, they take only the existing information into account and do not make any explicit statements on regions not covered by these data.

In this study, we highlight the potential of evolving functional paleoclimate networks for investigating climate variability during the Common Era (last 2 ka) in the European North Atlantic region. Climate dynamics within this region is of crucial importance not only at regional scales (Scaife et al., 2008; Trigo et al., 2002), but also as a pacemaker for the whole Northern Hemisphere (Delworth et al., 2016). Inter-annual to multi-decadal climate variability in the North Atlantic sector is strongly influenced by large-scale variability patterns like the ~~Atlantic Multidecadal Circulation (AMO) (Knight et al., 2006) and~~ North Atlantic Oscillation (NAO) (Hurrell et al., 2003). ~~In the following, we will focus on spatial patterns of inter-annual climate variability, which will be attributed to the NAO even though we are considering averages over multi-decadal periods. Other relevant regional variability modes, in particular the AMO, appear not to be consistently represented in the proxy data studied in this work to the same extent as the NAO. Thus, efforts to relate functional paleoclimate networks to the AMO by making use of the approach followed in this work for the NAO did not yet yield any interpretable results and will therefore not be reported here.~~ The NAO is related with the persistent redistribution of air masses between the Arctic and the Central Atlantic (Hurrell and Deser, 2010) and is commonly defined as a pressure dipole over the North Atlantic, consisting of a predominant low pressure system over Iceland and a high pressure system close to the Azores. The strength of the gradient between both varies in time and provides a basis for the quantitative description of the NAO based on an index where high (low) values correspond to a strong (weak) gradient. This pressure gradient has severe consequences for climate variability in Europe, Greenland and North America. A positive phase of the NAO is commonly associated with more moderate temperatures and higher precipitation sums during winter in Northern Europe and the Eastern United States, whereas Greenland, Canada and

Southern Europe often exhibit opposite characteristics. While the influence of the NAO phase is strongest during boreal winter, it also affects summer conditions (Ogi et al., 2003; Folland et al., 2009).

As the NAO is a key aspect of European climate variability, long-term changes in the dominant phase of this atmospheric variability mode should also be reflected in the co-variability structure of existing paleoclimate records. To this end, there are various types of archives available that could be utilized for reconstructing such changes. Specifically, most high-resolution archives in the region are sensitive tracers of inter-annual temperature variability (commonly seasonal or annual mean values) and, hence, should have been influenced to a certain degree by the NAO. For example, ice core data from Southern Greenland mostly reflect winter temperatures and, thus, strongly follow winter NAO conditions (Appenzeller et al., 1998; Vinther et al., 2010). In turn, tree ring chronologies mainly trace summer temperatures, but can be additionally influenced by winter conditions, especially in terms of extreme **positive** precipitation anomalies (see, e.g., Lindholm et al., 2001; Linderholm and Chen, 2005; Vaganov et al., 1999, and references therein). While these facts have been utilized to reconstruct NAO indices prior to the instrumental period (Cook et al., 1998), caution has to be taken since the corresponding relationship is non-stationary and, thus, the principle of uniformitarianism can be violated (Schmutz et al., 2000; Zorita and González-Rouco, 2002; Lehner et al., 2012). Furthermore, the actual effect may not be the same at different locations **and at different phases of the NAO. For example, a persistent positive phase of the NAO can enhance winter precipitation in Northern Europe, which in turn has an indirect influence on tree growth during the subsequent summer. The corresponding opposite effect of a negative NAO phase is expected to be much smaller. A similar relationship is expected to be present in Central and Southern Europe, but here increased precipitation is commonly associated with negative NAO phases, while positive NAO phases foster dry conditions and even droughts.** ~~while extreme snowfall during winter is of lower importance than in Fennoscandia. Hence, we expect tree ring records from Central and Southern Europe to be more strongly affected by negative NAO phases than by positive ones.~~

Following upon these considerations, we anticipate that different types of terrestrial archives available in different parts of the North Atlantic region have in common that they all reflect the leading mode of regional climate variability at inter-annual to multi-decadal time scales in one way or another. Hence, the main idea of this study is that by exploiting the temporary presence or absence of similar variability patterns between different paleoclimate records, especially in Southern Greenland versus the rest of the study region, one can draw conclusions about changing commonalities between the main atmospheric drivers in different sub-regions and, thus, the mean state of the atmospheric circulation in the North Atlantic region.

We emphasize that the aforementioned perspective differs markedly from previous efforts to reconstruct NAO variability over the past millennia. The general approach of existing studies, as used by Ortega et al. (2015); Trouet et al. (2009) and others, has been to find a reasonable set of records and use linear regression techniques, to infer NAO variability from proxy variability. In contrast to this, the evolving functional network approach does not assume any stationarity of the relationship between each individual proxy and the NAO. In fact, we explicitly utilize the observed non-stationarity and thus offer a complementary view on the same climatological target variable to common linear regression models. This fact makes it also challenging to test our proposed procedure by making use of pseudo-proxies, ~~as the latter commonly only represent a primary signal (e.g., temperature) in a stationary manner.~~ **which present a common framework for applying statistical methodologies based**

upon, e.g., climate model outputs, where the "true" target variable (in our case, some NAO index) is known. However, in the present case, the actual relationship between the NAO and the considered multiplicity of terrestrial paleoclimate archives is mediated by multiple climate variables, for example temperature, extreme winter precipitation and others. The influence of multiple variables on paleoclimate archives is in many cases not sufficiently well constrained to infer a particular statistical model. While pseudo-proxies exhibiting the corresponding level of complexity could probably be constructed, their application would present a rather novel methodological aspect that deserves further studies on its own.

The remainder of this paper is structured as follows: In Sec. 2, we describe the ensemble of paleoclimate proxy records used in this study. Section 3 presents the methods used to construct evolving functional paleoclimate networks and derive from them a scalar index variable describing the time-dependent dominant mode of North Atlantic climate variability over the Common Era. In Sec. 4, we discuss the emerging structures and how well they reflect associated changes in the dominant NAO phase at multi-decadal to multi-centennial time scales. Our results are compared to findings from other studies in Sec. 5, including a discussion on the possibilities and possible shortcomings of our approach. The paper ends with concluding remarks in Sec. 6.

2   Data

The North Atlantic region comprises a large variety of well-studied high-resolution paleoclimate archives for the Late Holocene. Existing data sets include several ice core records from the Greenland ice shelf and Svalbard, tree ring chronologies from the Scandinavian mountains, the Alps and other mountain ranges, and varved lake sediments, especially in Southern Finland. In addition, there exist also some very long historical temperature records based upon early instrumental records (see references in the Supplementary Material Tab. S1 and S2). Many of the available proxies are strongly correlated to seasonal or annual temperature variability and thus have been used as key input for existing regional (Werner et al., 2017a; Luterbacher et al., 2016; PAGES 2k Consortium, 2013) and hemispheric (Ljungqvist et al., 2012; Mann et al., 2008) temperature reconstructions. While there are particularly many archives covering the last millennium, a considerably lower number spans the full Common Era at high resolution.

In this study, we concentrate on changes in the inter-annual to multi-decadal co-variability of temperature-sensitive proxies throughout the Common Era. Thus, we only include records into our analysis that span at least 300 years and have close to annual resolution. This leaves us with 37 time series, which are described in detail in the Supplementary Material Tab. S1 and S2. They are shown in Fig. 1. Note that only 12 of these records cover the full Common Era, all of them being located in either Greenland, Fennoscandia or the Alps.

We note that the considered selection criteria exclude some existing records, which have been related to past NAO variability in previous works, but are either not temperature sensitive or not of (approximately) annual resolution (for example, the records used in the recent study of Deininger et al. (2016)). Specifically, most of the records studied in this work have been derived from either tree rings, varved lake sediments or ice cores, all of which have been dated rather precisely (for estimates of the dating uncertainty of some of the selected records, see, e.g., Werner et al. (2017b)). Thus, age uncertainty is considered to be

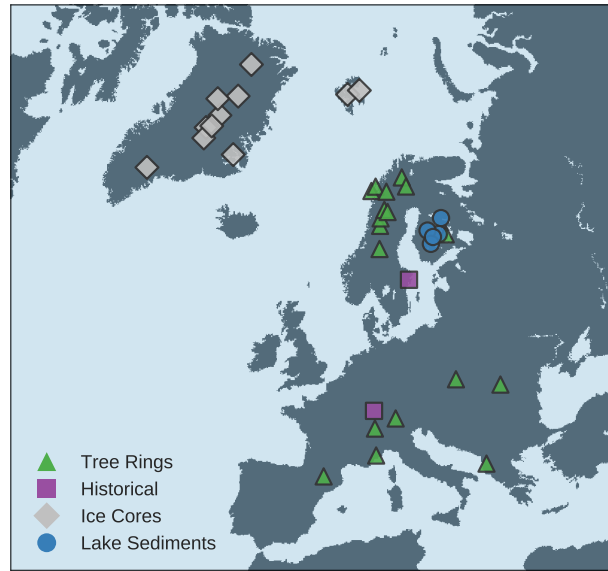



Figure 1. Locations and types of the paleoclimate archives used in this study. Different markers indicate different paleoclimate archives. Detailed information on the individual data sets can be found in the Supplementary Material Tab. S1 and S2.

5 negligible in the following. Unfortunately, there is no detailed information on proxy uncertainty that is consistently available for all records. Consequently, this uncertainty will not be further addressed explicitly in this following analysis.

3 Methods

As mentioned in the introduction, functional climate network analysis has recently become an established tool for studies on climate dynamics. Following upon the success of this approach, a few initial studies have transferred the corresponding idea to
10 the analysis of spatial co-variability patterns among paleoclimate archives in a defined region (Rehfeld et al., 2013; McRobie et al., 2015; Oster and Kelley, 2016). Beyond the original framework, we aim here at studying the statistical interdependence structure between subsets of archives from different appropriately defined regions and relate the information inferred from this analysis to a macroscopic index tracing the dominating mode of inter-annual North Atlantic climate variability at multi-decadal time scales. Accordingly, the methodological approach followed in this work comprises the following four steps:

15 1. Identify an ensemble of paleoclimate proxy records that have been influenced by a common climate variable (in our case, temperature).

2. Construct evolving functional networks based upon this ensemble according to the mutual similarity between individual records. Temporal changes in the network's connectivity structure represent changes of co-variability between the studied proxies.

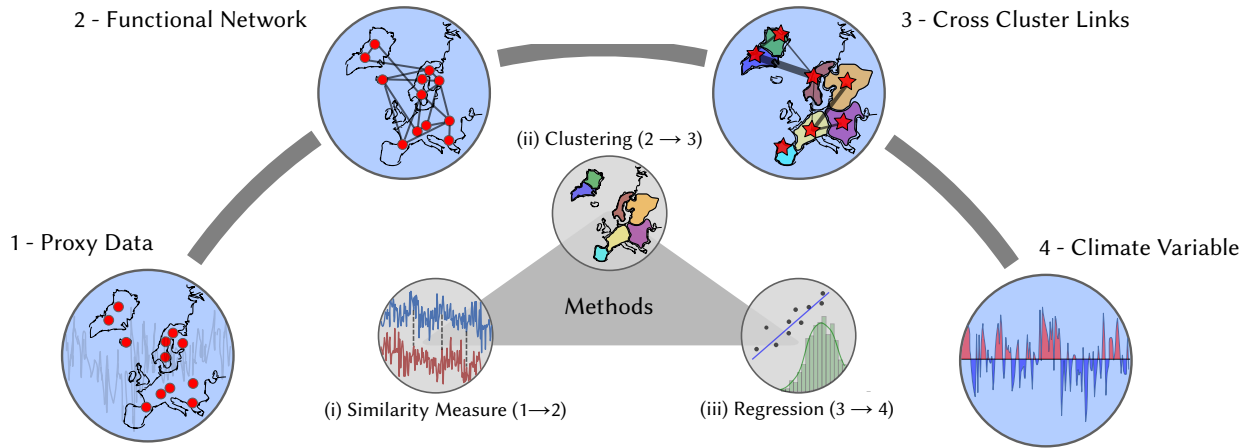



Figure 2. Schematic overview on the methodological approach of this study. Based upon a network of paleoclimate proxy records, we construct evolving functional networks encoding the co-variability among the different time series. Using cluster analysis, we simplify the emerging network structures and obtain quantitative measures of the inter-cluster linkages as key characteristics of the obtained networks. These variables are then related to an established long-term reconstruction of the NAO index via linear regression.

5   3. Introduce a meaningful grouping of the records to reduce the complexity and increase robustness of the obtained information in the presence of proxy uncertainty and a varying number of records.

    4. Establish a statistical relationship between the characteristics of the paleoclimate network and some climate variable or index (in our case, an existing long-term NAO reconstruction).

This general workflow of functional paleoclimate network analysis is illustrated in Fig. 2. While we have already discussed the
10 first step of this workflow in the previous section, we will highlight in the following the methodological realization of the three remaining steps.

3.1 Functional network construction

In general, an evolving functional network represents the time-dependent statistical co-variability structure among a set of time series. In the case of a perfect climate recording, this information can be used to infer teleconnections between distant regions. For paleoclimate archives however, one has to recall the existence of additional influencing factors represented in the recorded proxies. Thus, co-variability might not just indicate direct physical linkages, but possibly also secondary (external) effects that manifest in a similar way in different archives. For example, extremal winter precipitation can affect both, tree ring records and lake sediments, even though the archives considered in this work primarily record temperature variations.

5 More specifically, a functional network is a graph-theoretical representation of the mutual similarity structure among a set of time series $\{x_t^i\}$ $(i = 1, \ldots, N^r)$, in our case a set of $N^r = 37$ paleoclimate proxy records, which are considered as nodes of the network. Here, the effective number of available records N_t^r varies in time, with $\max_t N_t^r \leq N^r$, and each node is identified with the respective geographical location of the underlying paleoclimate archive. Links between different nodes are established if the corresponding time series are significantly similar according to some corresponding measure of pair-wise 10 statistical dependency. The details of this similarity assessment are discussed in Sec. 3.2.

 Evolving functional networks describe a time-ordered sequence of such functional networks, each being constructed from data within a given time window of length $\mathcal{W}$, which span different periods of time. Here, a time coordinate is assigned to each window according to its respective end point. Thus, any statistical measure calculated from a subset of a time series $\{x^i\}_{t^{\mathcal{W}}}$ corresponding to the time window of length $\mathcal{W}$ ending at time t is associated with the set of time indices $t^{\mathcal{W}} := \{t' | 0 \leq$ 15 $t - t' \leq \mathcal{W}\}$. The window size $\mathcal{W}$ determines the temporal resolution of the analysis. Thus, we obtain a sequence of networks associated with different time intervals. Within this sequence, changes in the network structure allow tracing changes in the spatio-temporal co-variability among the studied proxies through time.

 Each individual network is described by the $N^r \times N^r$ *adjacency matrix* $\mathbf{A}$, defining the connections between nodes,

$$
A_{t^{\mathcal{W}}}^{i,j} = \begin{cases} 1 & \text{if } \{x^i\}_{t^{\mathcal{W}}} \text{ and } \{x^j\}_{t^{\mathcal{W}}} \text{ are similar} \\ 0 & \text{else.} \end{cases} \tag{1}
$$

20 3.2 Similarity assessment

In general, functional networks can be constructed based upon different types of similarity measures, including classical linear approaches like the Pearson correlation, but also other measures suitable for detecting non-linear or event-based relationships, like mutual information or event synchronization (see, e.g., Rehfeld and Kurths, 2014, for a comparison of possible measures). Specifically, to determine the strength of co-variability between two paleoclimate records, a suitable similarity measure has 25 to be able to cope with unevenly sampled and/or discontinuous time series. Here, we use a Gaussian kernel-based variant of the Pearson correlation coefficient (gXRF) (Rehfeld et al., 2011). Given two **(normalized)** time series $\{x_i\}_{i=1}^{N^x}$ and $\{y_j\}_{j=1}^{N^y}$ **(exhibiting zero mean and unit variance)** with observation times $\{t_i^x\}_{i=1}^{N^x}$ and $\{t_j^y\}_{j=1}^{N^y}$, this correlation is defined as

$$
\rho(x,y) = \frac{\sum_{i=1}^{N^x} \sum_{j=1}^{N^y} x_i y_j K(t_j^y - t_i^x)}{\sum_{i=1}^{N^x} \sum_{j=1}^{N^y} K(t_j^y - t_i^x)}. \tag{2}
$$

Here, the kernel function $K(\cdot)$ is given by

$$
K(t_j^y - t_i^x) := \frac{1}{\sqrt{2\pi h}} e^{-(t_j^y - t_i^x)^2 / 2h} \tag{3}
$$

with $h = \max(\Delta t^x, \Delta t^y)/4$, where $\Delta t^{x,y}$ denote the mean sampling intervals of the corresponding time series. Specifically, as the same atmospheric driving variable can have qualitatively different effect on different proxies or at different locations, we take the absolute value of this correlation to quantify the strength of similarity.

5   Since our analysis makes use of different types of paleoclimate proxies, it is advisable to define similarity in a way that takes the different characteristics of the proxies into account. Time series originating from the same type of archive record climate variability in a similar fashion and thus might intrinsically exhibit stronger mutual correlations than such from different types of archives. Specifically, long-range auto-correlations and associated low-frequency variability can lead to spurious correlations, which have to be corrected for (Guez et al., 2014). To account for this problem, we apply a surrogate-based significance

10   test to calculate p-values corresponding to the probability that two records are similar just by chance, given their inherent auto-correlation structures. For this purpose, we use 1000 amplitude-adjusted Fourier transform (AAFT) surrogates (Schreiber and Schmitz, 2000), which leave the auto-correlation structure of each time series intact. Note that among the considered set of archives, all but four proxy records are complete and actually evenly distributed at annual time scale (one having lower sampling resolution and the three others containing gaps). In this regard, using the Gaussian kernel correlation (developed

15   for unevenly sampled data) instead of classical Pearson correlation coefficient accounts for these four records with different properties, while AAFT surrogates can still be generated with standard procedures (using linear interpolation for the very few missing data) for all records. The estimated p-values resulting from the surrogate ensembles provide a generally applicable measure of similarity, and two proxies are considered to be similar, if their p-value (rather than the estimated correlation value itself) is below a defined threshold value α_{pr}.

20   **3.3   Network analysis**

In the case of paleoclimate time series, there are some particular complications to be addressed in the context of functional network analysis. First, many records do not cover the full time span under study. Thus, the effective number of nodes N_t^r varies with time (see Supplementary Material Fig. S1). Second, while different archives might be significantly affected by the same climate variable, they still exhibit both local and proxy-specific effects, so that the outcomes of pair-wise similarity assessments

25   can be highly case-specific, even though the shared climatic influence might be the same. Furthermore, paleoclimate networks often already exhibit too many nodes to allow for an intuitive climatic interpretation (in particular, of individual links between pairs of proxies), but are at the same time too small to apply more sophisticated methods from complex network theory **like community detection or structural characterization by means of more complex measures like network transitivity or betweenness centrality, which have been recently applied in the context of functional networks constructed from recent**

30   **climate data (Donner et al., 2017).**

    In order to address these peculiarities and to simplify the interpretation of the resulting network structures, it appears reasonable to combine spatially close records into *clusters* (as will be further detailed below) to yield smaller networks with fewer, but weighted connections. In the present context, a *cluster* is a subset of records $C_{t\mathcal{W}}^M \subset \left\{x_\bullet^i\right\}_{i \in M}$ with $M \subset \{1, 2, \ldots N^r\}$ and $\left|C_{t\mathcal{W}}^M\right| \geq 2$ (here, $|\bullet|$ denotes the cardinality, the number of members of a set). Note that we will consider the assignment of

5   any archive to a specific cluster as being fixed over the entire analysis period (see below). Hence, the existence and size of a given cluster ~~size~~ vary only due to the (non-) availability of the given archives during different periods of time.

Having obtained a climatologically meaningful grouping of our archives into spatially connected clusters, we define the cross-link density (CLD) between any two clusters $C_{t\mathcal{W}}^K$ and $C_{t\mathcal{W}}^L$ as

$$CLD_{t\mathcal{W}}^{K,L} := \frac{\text{\# links between } C_{t\mathcal{W}}^K \text{ and } C_{t\mathcal{W}}^L}{\text{\# possible links between } C_{t\mathcal{W}}^K \text{ and } C_{t\mathcal{W}}^L}$$

$$= \frac{\sum_{i \in K} \sum_{j \in L} A_{t\mathcal{W}}^{i,j}}{|K_{t\mathcal{W}}| \cdot |L_{t\mathcal{W}}|}. \tag{4}$$

Note that as we are generally considering evolving (i.e. time-dependent) network structures, the CLD for each pair of clusters will commonly vary in time. However, the CLD values are expected to be more robust tracers of the essential network structure than other commonly used network characteristics, since they combine information from various links and are properly normalized by the (time-dependent) number of records. If the paleoclimate archives were perfect recorders of the same climate variable these simplified networks would then provide a coarse representation of the teleconnectivity structure between larger regions. In the case of real-world paleoclimate records, the attribution of any specific physical process to a link between clusters is less clear. However, we can still interpret the existence of such links (rather than the actual directionality of the represented correlations) as being indicative of joint dynamical patterns and/or influences of different regions.

By making use of the approach detailed above, for S_C denoting the number of clusters $\{C_M \ldots | |C_M| \geq 2\}$ we can define an $S = \binom{S_C}{2}$ dimensional vector of cross-link densities

$$\mathbf{X}_{t\mathcal{W}} = \left\{ CLD_{t\mathcal{W}}^{K,L} | K, L \in \{1, \ldots S_C\}, K \neq L \right\}.$$

This way, we effectively obtain coarse-grained networks with fewer nodes (associated with each cluster) and weighted links (CLD values). In turn, we disregard any information on the intra-cluster statistical linkages between individual archives, since we expect the latter to mainly reflect the intrinsic spatial correlation length of the influencing climate variable. Combining information on intra- and inter-cluster linkages would result in a paleoclimate "network of networks" approach; a framework that has already been employed in a few studies on recent climate variability (Donges et al., 2011; Wiedermann et al., 2016), but might suffer from the low number of nodes in a paleoclimate setting.

3.4 Spatial clustering of proxies

As discussed above, it is advisable to simplify the functional paleoclimate network by grouping several archives into spatially connected and climatologically meaningful clusters and study exclusively the temporal changes in the mutual similarity of proxies at the resulting cluster level. Specifically, the obtained clustering should meet the conditions that clusters (i) comprise spatially close archives and (ii) are large enough to reduce the impact of individual records and thus lead to a robust representation of the large-scale spatial co-variability structure.

Given the distinct individual characteristics of the different proxy time series and local as well as archive-specific effects, it is difficult to perform a cluster analysis directly at the set of time series originating from the archives, since such a procedure would likely result in a highly fragmented cluster structure. Instead, given that all proxies in our ensemble are temperature-sensitive, we define clusters as regions which have shown similar inter-annual temperature variability over the modern (instrumental) period. For this purpose, we make use of the gridded ERA-20C reanalysis summer temperature data spanning the

whole 20th century (Poli et al., 2016). Based upon this data set, we generate a functional climate network reflecting only the strongest absolute linear (Pearson) correlations (as determined by a threshold value α_C) among the time series of seasonal mean (annually averaged boreal summer, JJA) temperatures for each grid point over land in the study region. From this net-

10 work representation, we identify subsets of grid points with high intrinsic and low extrinsic connectivity (referred to as network communities) by applying the so-called Louvain algorithm (for details, see Blondel et al. (2008)).

We emphasize that the described spatial clustering procedure introduces an additional parameter α_C into the analysis. In general, we observe that small values of α_C yield more but smaller clusters, while larger values lead to a lower number of larger clusters (not shown). One of the main differences between the obtained clusters using different values of α_C (from

15 a reasonable range of values) is the division of Greenland and the border between Central and Eastern Europe. This is also the main difference in using different variables (e.g. ERA-20C winter mean temperature) for the clustering (Supplementary Material Fig. S2).

3.5 Statistical modeling by regression

Beyond simple visual analysis of the evolving network structures, we aim to statistically link the obtained time-dependent CLD

20 values with a climate-related variable (in our case, an existing NAO reconstruction by Ortega et al. (2015), in the following referred to as $\text{NAO}_{\text{Ortega}}$) that reflects a common influence on the co-variability between different regions. The simplest model to establish such a relationship between $\mathbf{X}_{t w}$ and some variable Y would be a linear model

$$Y_{t w} = \mathbf{D}_{t w} \mathbf{X}_{t w} + \epsilon_{t w} \tag{5}$$

with a coefficient vector $\mathbf{D}$ and a noise process $\epsilon_{t w}$. Such a model implies, that the connectivity between different regions is

25 linearly related to the strength and the sign of the NAO. Note that while in general the CLD values should rather be described as a superposition of different climatic influences, we take here the opposite approach, which is potentially useful for obtaining a reconstruction of the (unknown) climate driver based upon our evolving functional paleoclimate network properties. For a detailed discussion on different (regular vs. inverse) versions of this regression problem and their implications in the context of paleoclimate reconstructions, we refer to Christiansen and Ljungqvist (2017); Christiansen (2014) and references therein.

30 We emphasize that the analysis procedure described above has two free parameters, the threshold values α_{pr} (for generating the paleoclimate network) and α_C (for obtaining the spatial clustering). Given that our general aim is to maximize the inferred information about the mean state of the leading mode of North Atlantic climate variability at inter-annual to multi-decadal scales (which we assume here to largely reflect the respective NAO phase) from the simplified networks, we vary the values of α_{pr} and α_C to obtain an ensemble of sequences of evolving networks as well as geographical clusterings, each ensemble member corresponding to a different combination of both parameters. For each member, we individually perform a multiple linear ordinary least-square (OLS) regression of all associated CLD values to the 50-year averaged $\text{NAO}_{\text{Ortega}}$ reconstruction (Ortega et al., 2015) – in the following denoted as $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ – using the model in Eq. (5). Application of larger smoothing

5 windows essentially yielded results very similar to those presented in the following, while for shorter windows, the statistical uncertainty of the regression substantially increased (not shown). In this regard, our choice of averaging over 50-year periods

represents a reasonable trade-off. The parameter combination (α_{pr}, α_C) for which the resulting regression model describes the largest fraction of variability in $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ (see Supplementary Material Fig. S3) is then selected for further analysis. In general the results obtained are robust for different window sizes (not shown).

10   As a final step, we further investigate the linear model (Eq. 5) for $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ using a Bayesian approach, Markov Chain Monte Carlo (MCMC) regression (Gilks et al., 1995). Unlike OLS regression, this method results not in individual estimates of the different regression coefficients, but in joint distributions for all model parameters. Thereby, we implicitly account for the uncertainty in the description of the target variable. Since some of the considered clusters of paleoclimate archives do not cover the full Common Era, we can furthermore use the parameter distributions of the full set of clusters as priors to find the new

15 distributions of the reduced set of CLD values, thus utilizing the knowledge of the full data for cases of lower data availability. For performing the MCMC regression, we employ a NUTS sampler (Hoffman and Gelman, 2014) with 10^4 samples with one quarter of these as burn-in.

4   Results

We have followed the procedure described in Sec. 3 to study the evolving paleoclimate networks derived from the set of 37

20 paleoclimate records described in Sec. 2. The window size $\mathcal{W}$ of our analysis has been selected as 50 years, consistent with the averaging window in $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$. We have considered sliding windows with a mutual offset of 1 year, implying an overlap of 49 years between subsequent windows. The threshold values α_{pr} and α_C have been determined by maximizing the explained variance of the 50-year averaged $\text{NAO}_{\text{Ortega}}$ reconstruction (see Supplementary Material Fig. S3), yielding $\alpha_{pr} = 0.46$ and $\alpha_C = 0.0104$. The resulting spatial clusters of archives used in the analysis are displayed in Fig. 3.

25   Figure 4 shows the simplified networks and dominating cross-cluster links for some exemplary time windows, using a lower threshold value of $\alpha_{pr} = 0.1$ to better highlight the strongest correlations (for illustrative purposes only). The most informative clusters are those with the highest regression coefficients. These are located in Southern Greenland (SG), Fennoscandia (FS) and Central Europe (CEU) and cover all of the Common Era.

    During the first millennium CE (of the Common Era), we can distinguish two common, qualitatively different states of

30 the network, one being dominated by connections between FS and the other two clusters (Fig. 4a,c) and another exhibiting strong correlations between archives from SG and CEU (Fig. 4b,d). During the second millennium CE (with considerably more archives available), we more clearly identify periods during which mainly West-East connections between Greenland (G), Svalbard (S) and FS are present (Fig. 4f,i). During other times, North-South connections involving CEU are more strongly expressed (Fig. 4 e,g,h). The latter is commonly the case during time intervals for which $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ indicates a negative mean reconstructed NAO index, while cross-cluster links are more concentrated within the northern North Atlantic sector during positive NAO phases.

    ~~To this end~~The time intervals presented in Fig. 4 and the aforementioned general patterns are still subjective and thus call

5 for a more objective approach. For this reason, we employ the linear model (Eq. 5) for $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$. The resulting mean regression coefficients of this model support the general patterns discussed before. The respective strengths and signs of the

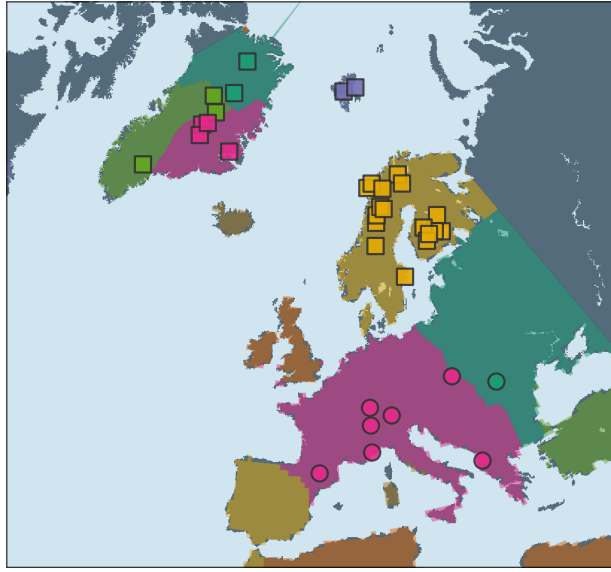



Figure 3. Division of the study area as obtained by cluster analysis of the ERA-20C summer mean temperatures, together with the pale-oclimate archives used in this study ($\alpha_C \approx 0.01$). Non-adjacent regions of the same color represent different clusters, as are indicated by different symbols (squares vs. circles) showing the spatial locations of the considered archives.

most relevant regression coefficients are illustrated in Fig. 5 and summarized in Supplementary Material Tab. S3. In that figure, thick red lines indicate that positive (negative) NAO phases coincide with relatively many (few) cross-cluster links, while the relation is just the opposite in case of linkages represented by blue lines. The time-evolution of the six CLDs associated with the largest coefficients are shown in Supplementary Material Fig. S4. The corresponding results further demonstrate that the presence of strong West-East connections is indicative of a positive NAO phase, while North-South connections, especially between CEU and the rest of the network, point toward negative NAO phases. The largest regression coefficients correspond to CLDs between SG and FS and CEU.

As an additional test for the validity of our estimated linear model describing multi-decadal NAO variability, we split the $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ reconstruction into two parts of equal size, using one part as training period and the other for validation. We apply OLS regression of the cross-link densities to $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ during the training period and then compare the values predicted by the obtained model for the validation period with the actual values of $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$. Using both parts as respective training and validation periods, the resulting r^2 values are very low (0.15 and 0.28, respectively). Hence, the linear model can scarcely explain the amplitude of the supposed long-term average NAO variability as expressed by $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$. Nevertheless, the obtained sign of the NAO phase is identified correctly in 68% and 71% of the considered time windows covered by $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$, respectively. In the following, we will refer to this quantity as the true sign ratio (TSR). Notably, taking the second (more recent) half of $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ as regression period (which corresponds to a period with more records than the

first one) results in higher values of both, r^2 and TSR. This finding suggests, that using additional records, which do not have data outside of the regression period, can still lead to a better performance of our model due to a better interpretation of the existing links. Future work along the lines of the present paper might explicitly utilize this observation in a Bayesian analysis

10   framework.

While the observed TSR provided by our model is markedly larger than that of a random guess (TSR $\approx 50\%$), it is still rather low in comparison with common requirements if using such models for predictive purposes. To better understand the $\sim 30\%$ of time intervals during which the sign of $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ is not correctly represented by our model better, we perform an additional type of cross-validation, this time independently leaving out consecutive 50-year windows as validation periods and taking the

15   rest of the data for model estimation. The results of this analysis are shown in Supplementary Material Fig. S5. We find, that whenever the TSR in a given validation period is clearly lower than the mean value of 0.69, $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ is either consistently close to zero or exhibits a transition between positive and negative NAO phases. Nevertheless, there are some periods during which our model differs markedly from $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$, e.g. in the 17$^{\text{th}}$ century. Hence, we conclude that if our model is used for the purpose of hindcasting the (themselves statistically reconstructed) NAO values according to Ortega et al. (2015), this

20   might result in incorrect identifications of the mean NAO phase at values where $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ is close to zero, since under these conditions, our ensemble of equally likely NAO "trajectories" obtained from MCMC regression includes both, positive and negative estimates for the corresponding time window. In turn, the model performs well in correctly identifying strong and persistent positive and negative NAO phases. However, we observe that the actual timing of transitions between distinct NAO phases can differ between $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ and our model (Supplementary Material Fig. S5). Thus, the relatively low predictive

25   skill of our reconstruction method might in part be due to a possibly lagged representation of NAO shifts in the connectivity between regional clusters of proxies. Further possible reasons and suggestions for improvements are briefly mentioned in Sec. 6.

Despite the fact that only four geographical clusters of paleoclimate archives cover the full Common Era, our model allows to qualitatively expand the existing "smoothed" NAO reconstruction by Ortega et al. (2015) over the last two millenia and thus

30   obtain relevant information about the dominant NAO phase at multi-decadal time scales. To do so, we draw 10,000 realizations of the regression coefficient distributions of our model and calculate the corresponding NAO values for each point in time based on the available CLDs. The probability that during a specific time interval, the NAO was in a positive (negative) phase is then estimated by the percentage of values above (below) zero in the ensemble of realizations. This value is subject to several proxy and model uncertainties and thus called the *degree of belief* that a certain NAO phase was present. The results of this

35   probabilistic description are shown in Fig. 6. We find that during the Common Era, there have been several phases during which the multi-decadal NAO variability was preferentially in a positive phase (e.g. during the migration period and the late medieval times), which alternated with strong negative phases (e.g. during the Little Ice Age, LIA) or intervals with generally more variability (e.g. the late Roman period or the centuries around 1000 CE) indicating more unstable conditions of the large-scale atmospheric circulation over the North Atlantic region.

5   Finally, we have additionally tested the robustness of the estimated regression model by varying α_C over a reasonable range (similar variations of α_{pr} were found not to alter the obtained results markedly, which is not explicitly shown here).

Supplementary Material Fig. S6 shows the corresponding results in terms of OLS-based regression models obtained with different parameter values. While most parameter sets close to the selected optimal one yield very similar results, there are few exceptions demonstrating the importance of this sensitivity analysis. A particularly remarkable example is found in the second half of the 5th century, where a transition from a predominantly negative NAO phase to a positive one is observed, the exact timing of which, however, differs significantly among the different regression models. This observation underlines, that our model has some uncertainty – not only in terms of reconstructed values, but also the timing of changes – which cannot be accounted for outside an ensemble-based approach.

5   Discussion

5.1   Climatological interpretation

In our analysis, we have related evolving functional paleoclimate networks to the dominant mode of multi-decadal variability of the atmospheric circulation in the North Atlantic region as reported by Ortega et al. (2015), which has been associated with long-term changes of the NAO. As seen from Figs. 4 and 5, the number of statistically relevant connections between Greenland, Svalbard and Fennoscandia is enhanced during positive phases of the NAO (according to the reconstruction by Ortega et al. (2015)), whereas links involving Central Europe are more pronounced during negative NAO phases. Thus, we interpret strong West-East correlations in the study region as indicators of a positive NAO phase, while North-South connections point to negative NAO phases. This observation can be qualitatively addressed by visualizing the obtained networks as in Figs. 4 and 5. More detailed statements can be made based upon a systematic inter-comparison between the time-dependent CLDs associated with distinct cluster pairs, as well as the evaluation of the regression to the existing reconstruction $\overline{\mathrm{NAO}}_{\mathrm{Ortega,50\,yr}}$ in terms of the linear model (Eq. 5). The latter analysis helps relating different NAO phases to certain cross-cluster links in a more objective fashion.

The interpretation of a preferred presence or absence of certain CLDs during specific NAO phases agrees well with the known NAO impact on European climate variability during the instrumental period. A positive NAO phase is commonly related to a northward shift of the westerlies, which causes milder temperatures and stronger precipitation in Northern Europe during boreal winter. Thus, from the observation that tree ring chronologies are strongly influenced by intense winter precipitation, it is reasonable that the considered archives from Fennoscandia (Central Europe) are particularly affected by positive (negative) NAO phases. The non-stationary influence of winter conditions is further illustrated in Supplementary Material Fig. S7. While many of the ice core records from Greenland, which were instrumental in obtaining the $\mathrm{NAO}_{\mathrm{Ortega}}$ reconstruction, exhibit strong negative correlations with that reconstruction throughout the last millennium, there is much more variability in the correlations with $\mathrm{NAO}_{\mathrm{Ortega}}$ for all the other records. More important than variability in magnitude is the fact, that a linear relationship is absent for most records at most times. In turn, they are co-varying with the NAO at certain times, but are unaffected at others.

Suppose that we are given a reference time series which exhibits a stationary relationship with the "true" NAO (in our case, the aforementioned Greenland ice cores). If the variability of any particular record shows a strong similarity with this reference series during a specific time, we expect that this record carries significant information about the NAO phase. As the actual

imprint of the NAO is different for different regions, this information is present in different proxy groups at different times. In our case, the Greenland ice cores act as a filter to indicate which regions are strongly influenced by the NAO and thus which NAO phase is more probable.

The median values of the regression model correlate well with $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ (squared Pearson correlation of $r^2 = 0.58$). However, this value could simply result from overfitting, as indicated by our cross-validation. While the obtained quantitative values of the reconstructed NAO index of Ortega et al. (2015) are thus not reliably described by our model, the latter performs well in resolving the dominant NAO phase. Drawing upon the knowledge about relationships between certain cross-cluster links and the NAO phase as discussed above, it is generally possible to extend the existing (smoothed) NAO index reconstruction $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ to the entire first millennium. However, since our linear model is not capable of describing the amplitude variability of $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ adequately, this should only be considered as qualitative information about the likely NAO phase.

While there is an influence of the NAO on regional temperatures at multi-decadal time scales, strong low-frequency temperature variations could be associated with both, multiple modes of internal variability as well as external factors like solar activity changes and explosive volcanism (Crowley, 2000), to different degrees that are subject of ongoing studies (see, e.g., Schurer et al. (2014); Bothe et al. (2015)). Thus, the observation of elevated temperatures reported for most of the Roman Warm Period (RWP) and Medieval Climate Anomaly (MCA) as opposed to lower temperatures during the Late Antique (Luterbacher et al., 2016; PAGES 2k Consortium, 2013) does not contradict our qualitative reconstruction of the predominant NAO phase. Instead, there might be common causes for such apparently contradictory observations. For example, explosive volcanism has been discussed as a major driver of the Late Antique Little Ice Age climate (Büntgen et al., 2016), but is also known to frequently trigger positive NAO-like atmospheric dynamics during the years following strong eruptions (Robock, 2000). Even though this is mostly a short-term effect, a high frequency of strong eruptions, as present during the Late Antique (Sigl et al., 2015), might have had a more persistent influence.

We emphasize that there are certain limitations to the usage of CLDs and our linear model to draw conclusions about the predominant NAO phase. First, most of the CLDs exhibit downward trends and progressively decreasing variance throughout the Common Era (Supplementary Material Fig. S4), which is probably related to the lower number of records as one goes back in time. This might add a considerable bias to any application of our methodological framework extending further back in time than the last millennium. In our case, this effect might favor positive NAO phases, since the regression coefficient for the connection between Southeast Greenland and Fennoscandia is by far the largest among all coefficient. Furthermore, our regression is based on a proxy-based reconstruction, which contains large uncertainties itself and therefore has limited value as a "ground truth". Ortega et al. (2015) reported, that their reconstruction explains only about 40% of the variance of the observed NAO. In turn, the explanatory value of the time series ($\overline{\text{NAO}}_{\text{Ortega,50 yr}}$) upon which our regression analysis is based, is a key assumption beyond our procedure. Moreover, our cross-validation showed that the linear model can disagree with $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ especially in cases in which that reconstruction has values close to zero, as well as in the timings of some transitions between positive and negative NAO phases. This intrinsic uncertainty of our qualitative reconstruction has been addressed by using MCMC regression to take a probabilistic view on the NAO phase. At time periods where the reconstructed $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ is close to zero, no particular phase is preferred in general (Fig. 6).

Following upon the considerable uncertainties in using our linear model to obtain qualitative estimates of the NAO phase during the entire Common Era, we will next compare our corresponding results to other long-term NAO-related climate reconstructions. Moreover, we will utilize further independent information in terms of documented drought periods for an independent validation of our reconstruction. Finally, we will discuss how long-term NAO variability might have affected European societies during the first millennium CE.

5.2 Comparison with other NAO reconstructions

Our model is able to reproduce most features of the $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$ reconstruction, including a dominant positive NAO phase during the late MCA, generally stronger variability during the LIA, with a tendency towards a more negative NAO phase, and another strongly positive phase during the 20$^{\text{th}}$ century, which is also in accordance with instrumental records (Vinther et al., 2003).

In contrast to other previous findings by Trouet et al. (2009), we do not observe strong West-East correlations during the early MCA, which suggests that this time interval has not been characterized by a strongly positive NAO.

Another recent NAO reconstruction by Olsen et al. (2012) used a more than 5000 years-long lake record from Southern Greenland to trace the predominant NAO phase during the entire Late Holocene. Our mean model correlates only extremely weakly with their reconstruction during the common period ($r^2 = 0.04$). One probable reason for this disagreement could be time uncertainty in their very long-term reconstruction, which Olsen et al. (2012) report as being of multi-decadal order during the first millennium CE. Furthermore, their record has been adjusted to the reconstruction by Trouet et al. (2009), which disagrees with NAO$_{\text{Ortega}}$ at many times.

Finally, it is worth considering a recent study by Deininger et al. (2016) who used 11 European speleothem records and analyzed their mutually coherent dynamics, which can be connected with changes in the North Atlantic circulation regimes reflecting long-term NAO variability. Their results indicate a strong, persistent positive NAO phase during the entire MCA and a tendency towards a negative NAO phase during the LIA, the first being in partial disagreement and the second in accordance with the NAO$_{\text{Ortega}}$ reconstruction. In addition, Deininger et al. (2016) report a dominant negative NAO phase between about 250 and 500 CE and a neutral-to-positive NAO phase thereafter. This observation agrees with our qualitative extension of $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$.

5.3 Comparison with historical droughts

Winter NAO and the corresponding precipitation anomalies exhibit known linkages to droughts in many parts of Europe, most significantly in the Western Mediterranean region (López-Moreno and Vicente-Serrano, 2008; Cook et al., 2016). Due to their severe impacts on agricultural productivity, droughts are some of the best documented weather extremes across historical times. Thus, existing reports of historical drought periods can be used as an independent source of information to test the consistency of our qualitative NAO reconstruction. For this purpose, we use three accounts of droughts during the Common Era. McCormick et al. (2012a) collected climatic evidence from the period of the Roman Empire (up to 800 CE) and reported 8 large droughts in the Western Empire (accessible through McCormick et al., 2012b). Domínguez-Castro et al. (2014) summarized

historical evidence from Muslim sources for Southern Spain, a region exceptionally vulnerable to NAO-related droughts (Cook
et al., 2016), from 711 to 1010 CE. They identified three major drought periods during this time. In addition, Cook et al. (2016)
discussed droughts during the last millennium in the Mediterranean region based upon the Old World Drought Atlas (OWDA)
(Cook et al., 2015). They reported that the drought index constructed for the Western Mediterranean correlates well with
NAO_{Ortega}. Therefore, we consider here only the five strongest drought events as discussed in Fig. 5 of their paper.

It has to be noted that although droughts are strongly related to precipitation deficits potentially associated with the dominant
NAO phase, they are complex phenomena with multiple causing factors. Thus, we do not expect the timing of all droughts in
the considered region to be fully explained by any particular NAO reconstruction.

In Fig. 6, the major drought periods discussed in the aforementioned publications are marked as gray bars. Among the 20
drought events, there is a clear tendency towards a positive NAO (P(NAO+) > 0.5) (17 cases, with 12 of these showing P(NAO+)
> 0.66). Thus, most droughts indeed coincide with positive NAO phases. However, a larger number of reported droughts during
a specific time period does not necessary imply a higher frequency of droughts. In case of historical documents, an increase in
reported events could also indicate that a society was more vulnerable to the impacts of droughts and, thus, found them more
worth reporting.

6 Conclusions

In this study, we have demonstrated that functional networks based on paleoclimate proxy records from multiple, spatially
distributed archives offer great potentials for identifying spatial patterns of atmospheric circulation in the European North
Atlantic sector together with information on their associated long-term variability.

Specifically, we have obtained a new 2000 year-long qualitative reconstruction of the leading mode of regional inter-annual
temperature variability probably associated with multi-decadal NAO variability. By combining visual inspection of changing
patterns in the (coarse-grained) network representations with a simple linear regression model, we have presented a clima-
tologically consistent interpretation of the time-dependent strength of correlations between groups of proxies from different
parts of Europe and the northern North Atlantic as indicators of different NAO phases. In general, we relate strong East-West
connections with a positive NAO phase and North-South connections with a negative phase. While the linear model does not
trace the exact variability of the reconstructed NAO index by Ortega et al. (2015) very well, it still provides a good qualitative
explanation of the succession of different phases at multi-decadal time scales.

The relatively low skill of our probabilistic reconstruction (true sign rate of $\sim 70\%$ when taking the reconstruction by Ortega
et al. (2015) as a reference) indicates that the method proposed in this work provides only a first step towards establishing a
novel tool, rather than being a conclusive framework to reconstruct the NAO from functional networks. A more restrictive proxy
selection, inclusion of additional proxies, consideration of more general methods of similarity assessment (e.g., replacing linear
correlations by non-linear mutual information) or applying different clustering schemes are possible future steps to further
improve the performance of the general scheme introduced in this work.

Uncertainties of the obtained NAO reconstruction arise mainly from an insufficient description of the observations by the linear model, a possible bias induced by a decreasing number of records when going further back in time, and existing uncertainties in the $\overline{\mathrm{NAO}}_{\mathrm{Ortega,50\,yr}}$ time series, upon which the regression is based. Thus, future consideration of additional high-resolution paleoclimate records from the North Atlantic region, especially from regions like Svalbard, Greenland and Eastern Europe, might further improve the model fit substantially. Notably, using standard similarity measures like the Pearson correlation coefficient, it is not feasible to use shorter time windows than the 50-year windows used in the present analysis. Thus, our approach cannot yet be directly applied to the instrumental record as regression target. Since the longest instrumental record of the NAO goes back to 1821 only (Vinther et al., 2003), there are not enough independent data points available for proper calibration and validation when using multi-decadal time windows.

Our qualitative expansion of the $\overline{\mathrm{NAO}}_{\mathrm{Ortega,50\,yr}}$ reconstruction demonstrates, that the Common Era has been characterized by time periods with different behavior of the NAO. In general, multi-decadal changes of the predominant NAO phase occurred relatively frequent during Roman and early medieval times, while there have been other periods characterized by a persistent positive or negative NAO phase. The first is the case for most of the migration period, the late medieval times and the 20$^{\mathrm{th}}$ century, while the latter is found during the late Roman times and the Little Ice Age. These long-term changes in the NAO phase might have had a considerable impact on European societies, as the NAO phase is associated with the likelihood of regional droughts as well as precipitation and temperature extremes, which may have directly affected agricultural productivity. In this spirit, specific phases might have supported some European societies while negatively affecting others.

In general, the procedure introduced in this study could also be applied to other large-scale climate variability patterns like the Atlantic Meridional Oscillation (Knight et al., 2006). Besides the NAO reconstruction, we have therefore performed the same analysis based upon the same set of proxies and a reconstruction of the Atlantic Multidecadal Variability (Mann et al., 2009) as a reference time series. Unlike for the NAO, this endeavour did not yield a reconstruction performing substantially different from a random one. This indicates, that the proxy selection is a crucial part of the proposed type of analysis, and that the set of records used in this study do not exhibit the necessary regional non-stationary relationship with the target variable that have been successfully utilized in reconstructing the NAO.

Code availability

All calculations in this work have been based upon open source software. AAFT surrogates have been generated using the Python package `pyunicorn` (Donges et al., 2015). Cluster analysis of reanalysis data has been conducted using the `community` package. MCMC regression has been performed with the `pyMC3` package (Salvatier et al., 2016).

Data availability

Most of the time series used in this study are available through the PAGES 2k project. The cooresponding references and data citations are provided in Supplementary Material Tab. S2. The obtained qualitative NAO reconstruction is provided as part of

the Supplementary Material accompanying this paper and additionally available via the data repository PANGAEA (Franke et al., 2017).

5 *Author contributions.* JGF designed and conducted the analysis and prepared the manuscript. RVD designed and supervised the analysis. JPW advised the data selection. RVD and JPW critically revised the manuscript and the interpretation of the obtained results.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. This work has been financially supported by the German Federal Ministry for Education and Research (BMBF) via the BMBF Young Investigators Group "CoSy-CC^2 - Complex Systems Approaches to Understanding Causes and Consequences of Past, Present
10 and Future Climate Change" (grant no. 01LN1306A), and by the joint German-Norwegian project "Nonlinear variability and regime shifts in Late Holocene climate: regional patterns and inter-regional linkages in multi-proxy networks and climate simulations" jointly funded by the German Academic Exchange Service (DAAD project no. 57245873) and the Research Council of Norway. RVD acknowledges additional support by a Bjerknes Visiting Fellow grant. The authors thank Dmitry Divine for fruitful discussions stimulating the developments described in this work. The data used in this study have been kindly provided by the Arctic2k group of the PAGES2k initiative and Saija Saarni.

15 References

Appenzeller, C., Stocker, T. F., and Anklin, M.: North Atlantic Oscillation Dynamics Recorded in Greenland Ice Cores, Science, 282, 446–449, doi:10.1126/science.282.5388.446, 1998.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E.: Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment, 2008, P10 008, doi:10.1088/1742-5468/2008/10/P10008, 2008.

Bothe, O., Evans, M., Donado, L. F., Bustamante, E. G., Gergis, J., Gonzalez-Rouco, J., Goosse, H., Hegerl, G., Hind, A., Jungclaus, J. H., et al.: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium, Climate of the Past, 11, 1673–1699, 2015.

Büntgen, U., Myglan, V. S., Ljungqvist, F. C., McCormick, M., Di Cosmo, N., Sigl, M., Jungclaus, J., Wagner, S., Krusic, P. J., Esper, J., Kaplan, J. O., de Vaan, M. A. C., Luterbacher, J., Wacker, L., Tegel, W., and Kirdyanov, A. V.: Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD, Nature Geoscience, 9, 231–236, doi:10.1038/ngeo2652, 2016.

Christiansen, B.: Straight line fitting and predictions: On a marginal likelihood approach to linear regression and errors-in-variables models, Journal of Climate, 27, 2014–2031, 2014.

Christiansen, B. and Ljungqvist, F. C.: Challenges and perspectives for large-scale temperature reconstructions of the past two millennia, Reviews of Geophysics, 55, 40–96, doi:10.1002/2016RG000521, http://dx.doi.org/10.1002/2016RG000521, 2016RG000521, 2017.

Cook, B. I., Anchukaitis, K. J., Touchan, R., Meko, D. M., and Cook, E. R.: Spatiotemporal drought variability in the Mediterranean over the last 900 years, Journal of Geophysical Research: Atmospheres, 121, 2015JD023 929, doi:10.1002/2015JD023929, 2016.

Cook, E. R., D'Arrigo, R. D., and Briffa, K. R.: A reconstruction of the North Atlantic Oscillation using tree-ring chronologies from North America and Europe, The Holocene, 8, 9–17, doi:10.1191/095968398677793725, 1998.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., Schrier, G. v. d., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., Esper, J., Griggs, C., Gunnarson, B., Günther, B., Gutierrez, E., Haneca, K., Helama, S., Herzig, F., Heussner, K.-U., Hofmann, J., Janda, P., Kontic, R., Köse, N., Kyncl, T., Levanič, T., Linderholm, H., Manning, S., Melvin, T. M., Miles, D., Neuwirth, B., Nicolussi, K., Nola, P., Panayotov, M., Popa, I., Rothe, A., Seftigen, K., Seim, A., Svarva, H., Svoboda, M., Thun, T., Timonen, M., Touchan, R., Trotsiuk, V., Trouet, V., Walder, F., Ważny, T., Wilson, R., and Zang, C.: Old World megadroughts and pluvials during the Common Era, Science Advances, 1, e1500 561, doi:10.1126/sciadv.1500561, 2015.

Crowley, T. J.: Causes of Climate Change Over the Past 1000 Years, Science, 289, 270–277, doi:10.1126/science.289.5477.270, 2000.

Deininger, M., McDermott, F., Mudelsee, M., Werner, M., Frank, N., and Mangini, A.: Coherency of late Holocene European speleothem $\delta 18O$ records linked to North Atlantic Ocean circulation, Climate Dynamics, pp. 1–24, doi:10.1007/s00382-016-3360-8, 2016.

Delworth, T. L., Zeng, F., Vecchi, G. A., Yang, X., Zhang, L., and Zhang, R.: The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere, Nature Geoscience, 9, 509–512, doi:10.1038/ngeo2738, 2016.

Domínguez-Castro, F., de Miguel, J. C., Vaquero, J. M., Gallego, M. C., and García-Herrera, R.: Climatic potential of Islamic chronicles in Iberia: Extreme droughts (ad 711–1010), The Holocene, 24, 370–374, doi:10.1177/0959683613518591, 2014.

Donges, J. F., Schultz, H. C. H., Marwan, N., Zou, Y., and Kurths, J.: Investigating the topology of interacting networks, The European Physical Journal B, 84, 635–651, doi:10.1140/epjb/e2011-10795-8, 2011.

15   Donges, J. F., Heitzig, J., Beronov, B., Wiedermann, M., Runge, J., Feng, Q. Y., Tupikina, L., Stolbova, V., Donner, R. V., Marwan, N., Dijkstra, H. A., and Kurths, J.: Unified functional network and nonlinear time series analysis for complex systems science: The pyunicorn package, Chaos: An Interdisciplinary Journal of Nonlinear Science, 25, 113 101, doi:10.1063/1.4934554, 2015.

Donner, R. V., Wiedermann, M., and Donges, J. F.: Complex network techniques for climatological data analysis, in: Nonlinear and Stochastic Climate Dynamics, edited by Franzke, C. L. E. and O'Kane, T. J., pp. 159–183, Cambridge University Press, Cambridge, 2017.

20   Folland, C. K., Knight, J., Linderholm, H. W., Fereday, D., Ineson, S., and Hurrell, J. W.: The Summer North Atlantic Oscillation: Past, Present, and Future, Journal of Climate, 22, 1082–1103, doi:10.1175/2008JCLI2459.1, 2009.

Franke, J. G., Werner, J. P., and Donner, R. V.: Reconstructing Late Holocene North Atlantic atmospheric circulation changes using functional paleoclimate networks, supplementary material, doi:10.1594/PANGAEA.875881, https://doi.pangaea.de/10.1594/PANGAEA.875881, 2017.

25   Gilks, W. R., Richardson, S., and Spiegelhalter, D.: Markov chain Monte Carlo in practice, CRC press, London, 1995.

Gouirand, I., Linderholm, H. W., Moberg, A., and Wohlfarth, B.: On the spatiotemporal characteristics of Fennoscandian tree-ring based summer temperature reconstructions, Theoretical and Applied Climatology, 91, 1–25, doi:10.1007/s00704-007-0311-7, 2008.

Guez, O. C., Gozolchiani, A., and Havlin, S.: Influence of autocorrelation on the topology of the climate network, Physical Review E, 90, 062 814, doi:10.1103/PhysRevE.90.062814, 2014.

30   Hoffman, M. D. and Gelman, A.: The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo., Journal of Machine Learning Research, 15, 1593–1623, 2014.

Hurrell, J. W. and Deser, C.: North Atlantic climate variability: The role of the North Atlantic Oscillation, Journal of Marine Systems, 79, 231–244, doi:10.1016/j.jmarsys.2009.11.002, 2010.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M.: An Overview of the North Atlantic Oscillation, in: The North Atlantic Oscillation:
35   Climatic Significance and Environmental Impact, edited by Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, r., pp. 1–35, American Geophysical Union, Washington, D. C., dOI: 10.1029/134GM01, 2003.

Knight, J. R., Folland, C. K., and Scaife, A. A.: Climate impacts of the Atlantic Multidecadal Oscillation, Geophysical Research Letters, 33, L17 706, doi:10.1029/2006GL026242, 2006.

Lehner, F., Raible, C. C., and Stocker, T. F.: Testing the robustness of a precipitation proxy-based North Atlantic Oscillation reconstruction, Quaternary Science Reviews, 45, 85–94, doi:10.1016/j.quascirev.2012.04.025, 2012.

Linderholm, H. W. and Chen, D.: Central Scandinavian winter precipitation variability during the past five centuries reconstructed from Pinus sylvestris tree rings, Boreas, 34, 43–52, doi:10.1111/j.1502-3885.2005.tb01003.x, 2005.

5   Lindholm, M., Eggertsson, Ó., Lovelius, N., Raspopov, O., Shumilov, O., and Laanelaid, A.: Growth indices of north European Scots pine record the seasonal North Atlantic Oscillation, Boreal Environment Research, 6, 275–284, 2001.

Ljungqvist, F. C., Krusic, P. J., Brattström, G., and Sundqvist, H. S.: Northern Hemisphere temperature patterns in the last 12 centuries, Clim. Past, 8, 227–249, doi:10.5194/cp-8-227-2012, 2012.

López-Moreno, J. I. and Vicente-Serrano, S. M.: Positive and Negative Phases of the Wintertime North Atlantic Oscillation and Drought
10   Occurrence over Europe: A Multitemporal-Scale Approach, Journal of Climate, 21, 1220–1243, doi:10.1175/2007JCLI1739.1, 2008.

Luterbacher, J., Werner, J. P., Smerdon, J. E., Fernández-Donado, L., González-Rouco, F. J., Barriopedro, D., Ljungqvist, F. C., Büntgen, U., Zorita, E., Wagner, S., Esper, J., McCarroll, D., Toreti, A., Frank, D., Jungclaus, J. H., M Barriendos, Bertolin, C., Bothe, O., Brázdil, R., Camuffo, D., Dobrovolný, P., Gagen, M., García-Bustamante, E., Ge, Q., Gómez-Navarro, J. J., Guiot, J., Hao, Z., Hegerl, G. C., Holmgren, K., Klimenko, V. V., Martín-Chivelet, J., Pfister, C., N Roberts, Schindler, A., Schurer, A., Solomina, O., Gunten, L. v., Wahl,

E., Wanner, H., Wetter, O., Xoplaki, E., Yuan, N., D Zanchettin, Zhang, H., and Zerefos, C.: European summer temperatures since Roman times, Environmental Research Letters, 11, 024 001, doi:10.1088/1748-9326/11/2/024001, 2016.

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Global-scale temperature patterns and climate forcing over the past six centuries, Nature, 392, 779–787, doi:10.1038/33859, 1998.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F.: Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia, Proceedings of the National Academy of Sciences, 105, 13 252–13 257, doi:10.1073/pnas.0805721105, 2008.

Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Ammann, C., Faluvegi, G., and Ni, F.: Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly, Science, 326, 1256–1260, 2009.

McCormick, M., Büntgen, U., Cane, M. A., Cook, E. R., Harper, K., Huybers, P., Litt, T., Manning, S. W., Mayewski, P. A., More, A. F. M., Nicolussi, K., and Tegel, W.: Climate Change during and after the Roman Empire: Reconstructing the Past from Scientific and Historical Evidence, Journal of Interdisciplinary History, 43, 169–220, doi:10.1162/JINH_a_00379, 2012a.

McCormick, M., Harper, K., More, A. M., and Gibson, K.: Geodatabase of historical evidence on Roman and post-Roman climate, DARMC Scholarly Data Series, DataContributionSeries, 1, 2012b.

McRobie, F. H., Stemler, T., and Wyrwoll, K. H.: Transient coupling relationships of the Holocene Australian monsoon, Quaternary Science Reviews, 121, 120–131, doi:10.1016/j.quascirev.2015.05.011, 2015.

Ogi, M., Tachibana, Y., and Yamazaki, K.: Impact of the wintertime North Atlantic Oscillation (NAO) on the summertime atmospheric circulation, Geophysical Research Letters, 30, 1704, doi:10.1029/2003GL017280, 2003.

Olsen, J., Anderson, N. J., and Knudsen, M. F.: Variability of the North Atlantic Oscillation over the past 5,200 years, Nature Geoscience, 5, 808–812, doi:10.1038/ngeo1589, 2012.

Ortega, P., Lehner, F., Swingedouw, D., Masson-Delmotte, V., Raible, C. C., Casado, M., and Yiou, P.: A model-tested North Atlantic Oscillation reconstruction for the past millennium, Nature, 523, 71–74, doi:10.1038/nature14518, 2015.

Oster, J. L. and Kelley, N. P.: Tracking regional and global teleconnections recorded by western North American speleothem records, Quaternary Science Reviews, 149, 18–33, doi:10.1016/j.quascirev.2016.07.009, 2016.

PAGES 2k Consortium: Continental-scale temperature variability during the past two millennia, Nature Geoscience, 6, 339–346, doi:10.1038/ngeo1797, 2013.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, Journal of Climate, 29, 4083–4097, doi:10.1175/JCLI-D-15-0556.1, 2016.

Radebach, A., Donner, R. V., Runge, J., Donges, J. F., and Kurths, J.: Disentangling different types of El Ni\~no episodes by evolving climate network analysis, Physical Review E, 88, 052 807, doi:10.1103/PhysRevE.88.052807, 2013.

Rehfeld, K. and Kurths, J.: Similarity estimators for irregular and age-uncertain time series, Clim. Past, 10, 107–122, doi:10.5194/cp-10-107-2014, 2014.

Rehfeld, K., Marwan, N., Heitzig, J., and Kurths, J.: Comparison of correlation analysis techniques for irregularly sampled time series, Nonlin. Processes Geophys., 18, 389–404, doi:10.5194/npg-18-389-2011, 2011.

Rehfeld, K., Marwan, N., Breitenbach, S. F. M., and Kurths, J.: Late Holocene Asian summer monsoon dynamics from small but complex networks of paleoclimate data, Climate Dynamics, 41, 3–19, doi:10.1007/s00382-012-1448-3, 2013.

Robock, A.: Volcanic eruptions and climate, Reviews of Geophysics, 38, 191–219, doi:10.1029/1998RG000054, 2000.

15 Salvatier, J., Wiecki, T. V., and Fonnesbeck, C.: Probabilistic programming in Python using PyMC3, PeerJ Computer Science, 2, e55, doi:10.7717/peerj-cs.55, 2016.

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., and Knight, J. R.: European Climate Extremes and the North Atlantic Oscillation, Journal of Climate, 21, 72–83, doi:10.1175/2007JCLI1631.1, 2008.

Schmutz, C., Luterbacher, J., Gyalistras, D., Xoplaki, E., and Wanner, H.: Can we trust proxy-based NAO index reconstructions?, Geophysical
20 Research Letters, 27, 1135–1138, doi:10.1029/1999GL011045, 2000.

Schreiber, T. and Schmitz, A.: Surrogate time series, Physica D: Nonlinear Phenomena, 142, 346–382, doi:10.1016/S0167-2789(00)00043-9, 2000.

Schurer, A. P., Tett, S. F., and Hegerl, G. C.: Small influence of solar variability on climate over the past millennium, Nature Geoscience, 7, 104, 2014.

25 Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schüpbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years, Nature, 523, 543–549, doi:10.1038/nature14565, 2015.

Trigo, R. M., Osborn, T. J., and CorteReal, J. M.: The North Atlantic Oscillation influence on Europe: climate impacts and associated physical
30 mechanisms, Climate Research, 20, 9–17, doi:10.3354/cr020009, 2002.

Trouet, V., Esper, J., Graham, N. E., Baker, A., Scourse, J. D., and Frank, D. C.: Persistent Positive North Atlantic Oscillation Mode Dominated the Medieval Climate Anomaly, Science, 324, 78–80, doi:10.1126/science.1166349, 2009.

Tsonis, A. A., Swanson, K. L., and Roebber, P. J.: What Do Networks Have to Do with Climate?, Bulletin of the American Meteorological Society, 87, 585–595, doi:10.1175/BAMS-87-5-585, 2006.

35 Vaganov, E. A., Hughes, M. K., Kirdyanov, A. V., Schweingruber, F. H., and Silkin, P. P.: Influence of snowfall and melt timing on tree growth in subarctic Eurasia, Nature, 400, 149–151, doi:10.1038/22087, 1999.

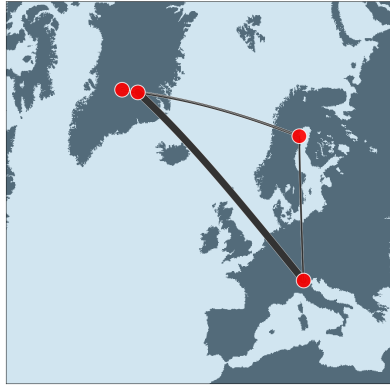
Vinther, B. M., Andersen, K. K., Hansen, A. W., Schmith, T., and Jones, P. D.: Improving the Gibraltar/Reykjavik NAO index, Geophysical Research Letters, 30, 2222, doi:10.1029/2003GL018220, 2003.

Vinther, B. M., Jones, P. D., Briffa, K. R., Clausen, H. B., Andersen, K. K., Dahl-Jensen, D., and Johnsen, S. J.: Climatic signals in multiple highly resolved stable isotope records from Greenland, Quaternary Science Reviews, 29, 522–538, doi:10.1016/j.quascirev.2009.11.002,
675 2010.

Werner, J. P., Divine, D. V., Charpentier Ljungqvist, F., Nilsen, T., and Francus, P.: Spatio-temporal variability of Arctic summer temperatures over the past two millennia: an overview of the last major climate anomalies, Climate of the Past Discussions, 2017, 1–43, doi:10.5194/cp-2017-29, https://www.clim-past-discuss.net/cp-2017-29/, 2017a.

Werner, J. P., Divine, D. V., Ljungqvist, F. C., Nilsen, T., and Francus, P.: Spatio-temporal variability of Arctic Summer Temperatures over
680 the past two Millennia: an overview of the last major climate anomalies, in revision, 2017b.

Wiedermann, M., Donges, J. F., Handorf, D., Kurths, J., and Donner, R. V.: Hierarchical structures in Northern Hemispheric extratropical winter ocean–atmosphere interactions, International Journal of Climatology, pp. n/a–n/a, doi:10.1002/joc.4956, 2016.

Zorita, E. and González-Rouco, F.: Are temperature-sensitive proxies adequate for North Atlantic Oscillation reconstructions?, Geophysical Research Letters, 29, 48–1, doi:10.1029/2002GL015404, 2002.

(a) 150 to 200 CE   (b) 350 to 400 CE   (c) 650 to 700 CE

(d) 710 to 760 CE   (e) 1100 to 1150 CE   (f) 1200 to 1250 CE

(g) 1400 to 1450 CE   (h) 1620 to 1670 CE   (i) 1896 to 1946 CE

Figure 4. Simplified functional paleoclimate networks for different exemplary time windows illustrating the great variety of spatial connectivity patterns during the time interval coverd by this study. The red circles indicate the center of each group of records (only shown if any record of a cluster has values at the specific time window). The thickness of each link is proportional to the CLD of that connection. The time intervals have been chosen such that they demonstrate the general patterns of zonal and meridional connectivity. For a more objective analysis we employ a linear model.

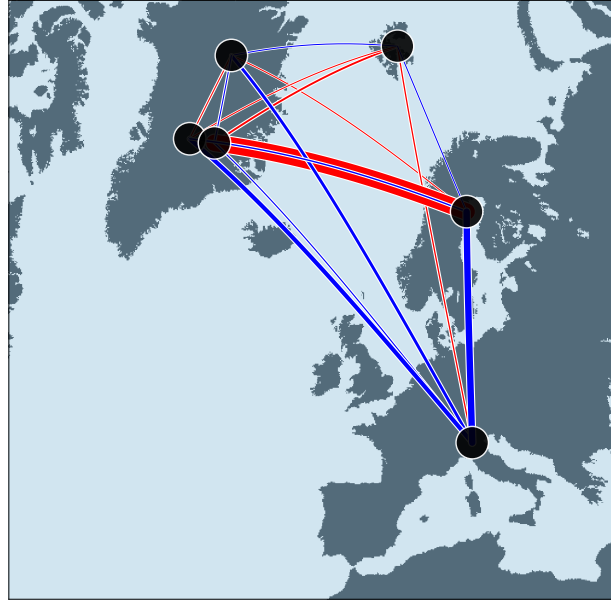



Figure 5. Regression coefficients between the 15 cross-link densities (CLD) among the spatial clusters of records and $\overline{\text{NAO}}_{\text{Ortega,50 yr}}$. Note that these linkages represent statistical relations and do *not* necessarily relate to (temperature) teleconnections between different regions, but may also reflect common factors influencing the respective regional climate dynamics as recorded by the proxies in a similar way. Black circles mark the centers of each group of records. Each line indicates the correlation between the CLD for two groups of records and the NAO reconstruction by Ortega et al. (2015). Red (blue) colors indicate positive (negative) sign of the coefficient, whereas the width of the drawn links is proportional to the mean coefficient value as given in Supplementary Material Tab. S3.

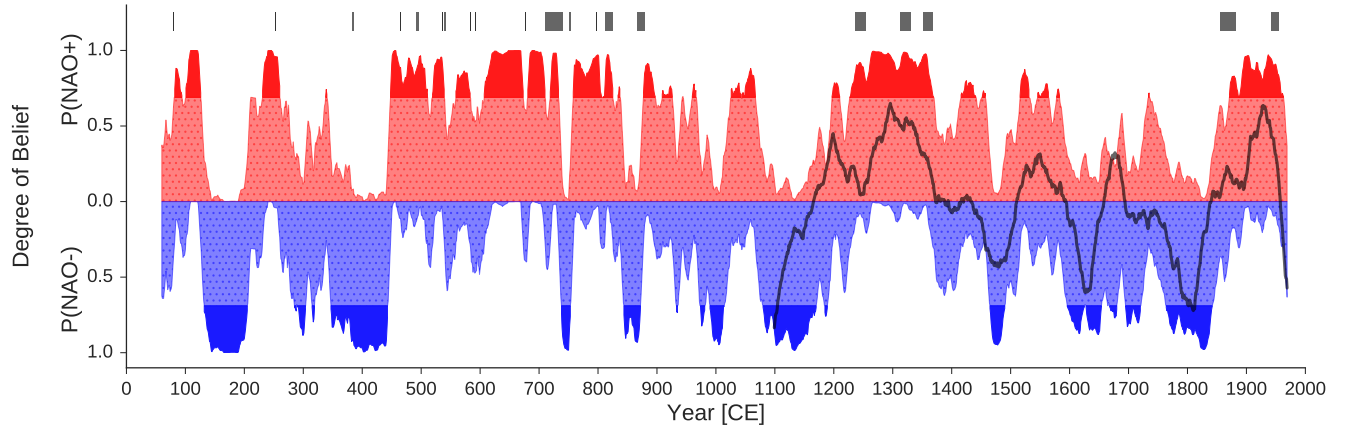



Figure 6. Degree of belief (probability) that the NAO is in a specific (positive vs. negative) phase during each 50-year time window. The figure has been smoothed by a 10-year moving-average filter to enhance its readability. The brighter areas indicate time intervals for which less than 66% of all considered MCMC ensemble members agree upon the sign of the reconstructed NAO variable. Gray bars correspond to known major drought episodes in the Western Mediterranean as discussed in Sec. 5. For better comparison, the NAO reconstruction by Ortega et al. (2015) is shown as a black line, indicating a general agreement with our probabilistic reconstruction over the common period as expected.