Can we trust proxy-based NAO index reconstructions?

C. Schmutz, J. Luterbacher, D. Gyalistras, E. Xoplaki, H. Wanner

Institute of Geography, University of Bern, Bern, Switzerland

Abstract. Three existing cold season North Atlantic Oscillation (NAO) indices which were reconstructed from proxy data were assessed for their consistency and robustness at the interannual and decadal timescales. All three indices were found to be inconsistent with an observed NAO index, whereas two other, mainly instrumentally based indices, showed good agreement with observations. The combination of the proxy-based indices into one new index by means of multivariate regression yielded some improvement, such that the skill of the new index was comparable or higher than the individual indices. However, the new index still did not correlate significantly with the instrumentally based reference time series prior to 1850. Our results suggest non-stationarities in the statistical relationships between the proxy-based indices and the NAO. This has important implications with regard to the choice of verification periods and predictor data-base for upcoming reconstruction attempts.

Introduction

The most important variation of atmospheric mass, energy and momentum in the North Atlantic-European sector in all seasons is associated with the North Atlantic Oscillation (NAO) [*Walker*, 1924; *Lamb and Peppler*, 1987; *Wallace and Gutzler*, 1981; *Barnston and Livezey*, 1987; *Kushnir and Wallace*, 1989]. Its variability is an important source of regional climate anomalies at seasonal to decadal timescales [*van Loon and Rogers*, 1978; *Rogers*, 1984; *Hurrell*, 1995, 1996; *Hurrell and van Loon*, 1997].

NAO indices defined by using two geographically fixed locations capture the main seasonal to interannual variability of the NAO. Such indices have been calculated based on sea level pressure (SLP) data back to 1865 [*Hurrell*, 1995] and 1821 [*Jones et al.*, 1997] (the latter denoted as J).

In order to gain a better understanding of the long-term behaviour of the NAO, several proxy NAO indices have been reconstructed based on paleoenvironmental data. *Cook et al.,* [1998] used tree-rings from Northern Fennoscandia and Eastern United States for reconstructions of the *Rogers* [1984] winter NAO (December to February) index back to 1701 (hereafter denoted as C). Based on ice accumulation rates in Greenland, a yearly mean NAO index (April to March) with a strong bias to the winter signal for the last 350 years has been provided by *Appenzeller et al.,* [1998] (hereafter denoted as A). *Stockton and Glueck,* [1999] used tree-ring data from Morocco and Finland as well as ice core data to reconstruct the winter NAO index [*Hurrell*, 1995] back to 1429 (hereafter denoted as SG). Recently, *Luterbacher et al.,* [1999] used early instrumental and documentary data for monthly NAO reconstructions back to 1675 (denoted as L).

Due to the use of different methods and time intervals to assess the skill of these reconstructions, a direct comparison of the different indices based on the literature is not possible. Luterbacher et al. [1999] argue to have provided the most reliable reconstruction back to 1675. However, several studies suggested a varying quality for different reconstructions [*Cook et al.*, 1998; *Osborn et al.*, 1999; *Luterbacher et al.*, 1999], and to our knowledge, no rigorous comparison of the various reconstructions has been undertaken up to now. Therefore, it is not clear which index should preferably be used for climate studies and testing of global climate models. In particular, prior to 1675 only proxy-based indices are available, so that it is important to know how far one can trust the proxy-based index reconstructions.

Mann et al., [1998] showed how the complementary seasonal information contained in different proxies may be used to estimate robust annual mean large-scale temperature patterns. At present, it is not clear whether a combination of different proxy data-bases may lead to an improvement of NAO index reconstructions for the pre-instrumental period.

Here we address the following questions: i.) Are the different NAO index reconstructions consistent? ii.) How far can we trust proxy-based NAO indices? iii.) Is it possible to improve the proxy-based reconstructions by combining different proxy-based NAO indices?

Data and Methods

For reasons of data availability we restricted our analysis to mean winter indices and an annual index with a strong bias to the winter. Figure 1 shows the five considered index time series and the respective calibration and reconstruction periods.

To test L we calculated two new NAO indices with fixed data-bases according to the *Luterbacher et al.,* [1999] method (based on canonical correlation analysis). Both indices were calibrated onto J: i.) L1 reflected the data availability of the early eighteenth century (no station SLP data). ii.) L2, derived from an intermediate data base, typical for the late eighteenth century (few SLP station series). From the monthly reconstructions, we computed seasonal averages for December through March (DJFM). No SLP data from stations at the Azores, Portugal or Iceland were used for either reconstruction.

The following strategy was adopted: a.) In order to compare the performance of the five considered indices back to 1821, J was used as the reference time series. b.) The proxy NAO indices were calibrated against J to reconstruct a combined proxy NAO index (C+SG+A) with a least-square fit. c.) In order to analyse the proxy-based reconstructions back

Copyright 2000 by the American Geophysical Union.

Paper number 1999GL011045.
0094-8276/00/1999GL011045$05.00

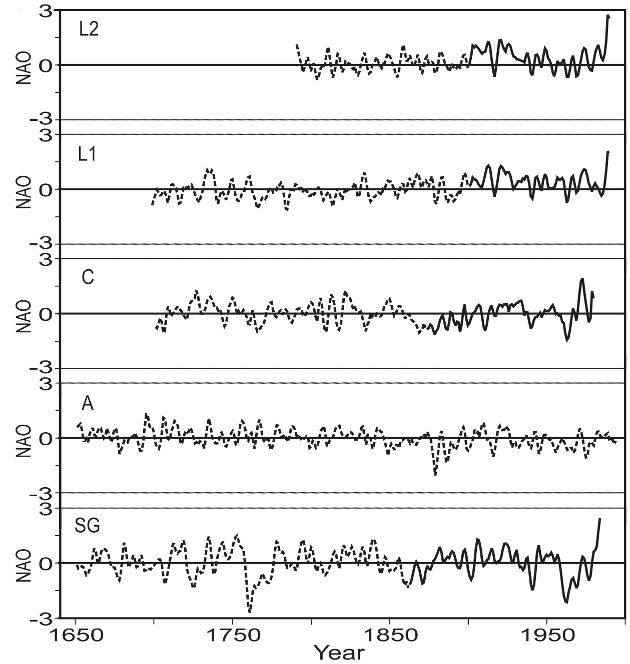

Figure 1. Overview of the considered NAO index reconstructions between 1650 and 1990. L1 and L2 were reconstructed with the methods of *Luterbacher et al.* [1999] but using smaller predictor data bases. C is the index by *Cook et al.* [1998], A by *Appenzeller et al.* [1998] and SG by *Stockton and Glueck* [1999]. SG originally extends to 1429, here it is standardized for optical reasons. Solid lines denote the final calibration periods, the reconstruction periods are dotted.

to 1701, they were correlated to the reference time series L back to 1701.

To investigate the decadal scale covariability of the various indices, they were filtered with a triangular 5-point low-pass filter (1-3-5-3-1) and correlated with the product-moment coefficient (r_{xy}). Confidence intervals for r_{xy} were calculated with Monte-Carlo experiments in which we generated pairs of synthetic time series having the same length, mean, variance and first order autocorrelation coefficients as each of the tested index time series.

The time-dependent development of the reconstruction performance of the various NAO indices was tested with 30-year window running correlations. A combined proxy NAO index (denoted as C+SG+A) was constructed through linear combination based of C, SG and A. The corresponding linear regression model was calibrated to J using ordinary least-square fits for the period 1901-1980. C+SG+A was tested against independent data from the periods 1824-1864 and 1865-1900.

Results

Figure 2 shows the running correlations of the low-pass filtered indices with J. L1 and L2 were highly significantly correlated with J at the decadal (Figure 2) and also at the interannual timescales (not shown). For the first half of the nineteenth century, only weak correlations were found between C, SG, A and J. Significance testing of the running correlations suggested that the proxy-based reconstructions were accurate only in the late nineteenth and mid-twentieth century.

The mutual interannual (unfiltered) and decadal (low-pass filtered) correlations of L1, C, SG and A in the 100-year period (1716-1815) were also calculated (not shown). Except for the unfiltered C and SG, no significant correlations were found among the different reconstructions.

The performance of C+SG+A was verified in two independent periods. A variance in common of 0.04 (0.38) with J was found for the verfication periods 1824-1864 (1865-1900), respectively.

Figure 3 presents the 30-year running correlations of L with C+SG+A, C, SG and A. All correlations showed decreasing values in the eighteenth century to a minimum around 1800 and then a strong increase up to the present, with a marked secondary drop around 1920. The most pronounced departure from this trend was observed for the correlation between L and A in the second half of the twentieth century.

The 30-year window running correlations of C+SG+A with L were generally higher or of the same order of magnitude as the best single proxy NAO reconstruction, except for the early nineteenth century and the second half of the eighteenth century, indicating a strong time-dependency in the accuracy of the C+SG+A time series.

Discussion

If all the reconstructed indices were expressions of the same atmospheric phenomenon (NAO), one would expect that they were significantly correlated to the reference time series J in an independent period (mid-nineteenth century). Results clearly showed that this was not the case. Moreover, only L1 and L2 were highly significantly correlated to J, whereas C, SG and A behaved like random time series. However, for the low-pass filtered proxy index time series, weak correlations to J were observed, thus indicating some modest reconstruction skill at the decadal timescale (Fig. 2). Our results do not support the claim of *Cook et al.,* [1998] and *Stockton and Glueck,* [1999], that their reconstructions are reliable with regard to the low-frequency variations of the NAO.

This is also suggested by the lack of significant mutual correlation on the interannual and decadal timescale be-

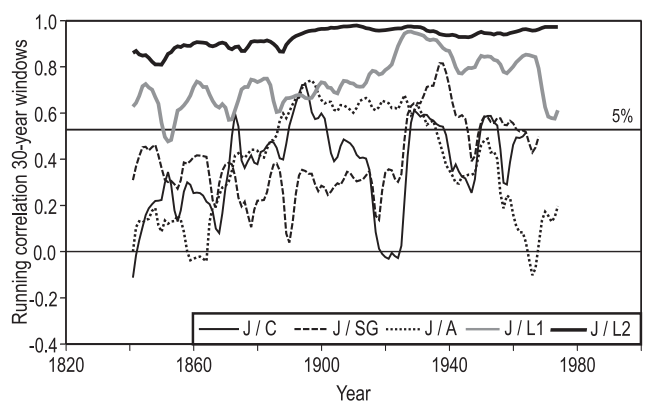

Figure 2. 30-year window running correlations between J and C, SG, A, L1, L2. The time series were low-pass filtered prior to the analysis, using a 5-point triangular filter. The two-tailed 5% significance level of 0.53 is estimated with a Monte-Carlo experiment, which is valuable for all five running correlation developments.

tween L1, C, SG and A in the eighteenth century and by the results shown in Figure 2. The significant correlation on the interannual timescale between C and SG might be due to a similar proxy data-base in northern Europe capturing common parts of the NAO variability.

The results of *Jones et al.,* [1998] showed that proxy variables have spatial and seasonal limitations, i.e. different proxies are potentially limited in their ability to represent climatic variations over a range of different timescales. Moreover, our results suggest that the proxy-based NAO indices are not able to capture the NAO variability in a consistent manner in the eighteenth and nineteenth centuries nor presumably further back in time. Possibly, the NAO phases include several different circulation patterns with specific impacts on a given proxy, resulting in low or varying covariability.

The C+SG+A index makes use of the time-dependent and complementary information content of the various proxies, such that the correlation to L is comparable or better than the respective best single proxy-index. Even though the correlations are below the 5% significance level prior to 1850, our results indicate that the integration of different proxy data-bases could yield improved or at least more robust reconstructions.

Figures 2 and 3 reveal that the reconstruction skill of proxy-based NAO index reconstructions varies considerably with time. This could have an important impact on the design of the calibration-validation experiments. For instance, *Cook et al.,* [1998] and *Stockton and Glueck,* [1999] validated their reconstructions in different periods after 1863. With a verification period in the second half of the nineteenth century and from 1920 to 1970, a satisfactory reconstruction skill can be achieved, while this is not the case for the first half of the nineteenth century or shortly after the turn of the twentieth century.

The reasons for the these non-stationarities are not obvious. One possibility is that the variance in common between the northern hemisphere temperature and the NAO is not always in the same order of magnitude [*Osborn et al.*, 1999]. From the turn of the nineteenth century to 1920, this variance in common is around zero, while before and after this period it was highly significant. In fact, the low correlations of the reference time series (J, L) to the proxy-based indices C and SG in the early twentieth century, indicate the strong temperature dependency of these NAO proxy indices. This might also explain the results of *Cook et al.,* [1998]. They calculated the correlations of Stykkisholmur and Bermuda SLP with C. In the periods 1838-1873 (Stykkisholmur) and 1837-1873 (Bermuda) the correlations were not significant (5% significance level), whereas in the period (1874-1980) the correlations were highly significant (1% significance level). The NAO index A showed a clearly different development of the 30-year window running correlations, because it is based on ice accumulation rates (precipitation dependency).

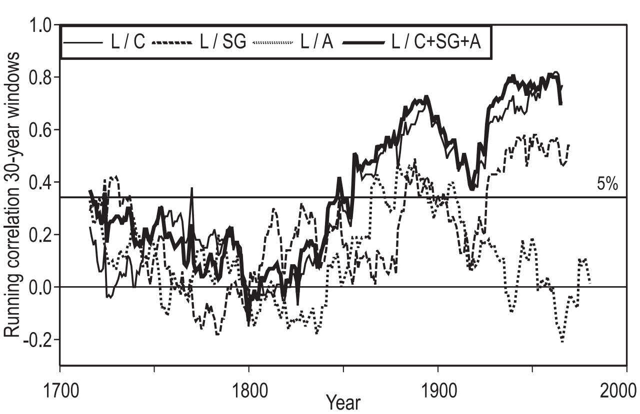

Figure 3. 30-year window running correlations between L and C, SG, A, C+SG+A. The two-tailed 5% significance level of 0.36 is estimated with a Monte-Carlo experiment, which is valuable for all four running correlation developments.

Conclusions

Prior to the twentieth century the various existing NAO index reconstructions show an inconsistent picture. This is the case for interannual as well as for decadal scale variability.

Our analysis clearly showed, that the most reliable reconstruction back to 1675 available to date is the one by *Luterbacher et al.,* [1999]. This index can be used for long-term testing of proxy-based indices on the monthly to decadal timescales.

The considered proxy-based indices only have modest reconstruction skill on the decadal timescale prior to 1850, while on the interannual timescale they seem to have no skill at all. The reconstruction skill varies considerably with time. This is of great importance for the choice of the validation period.

There are indications that a combination of complementary proxy data bases can improve the reconstruction skill. An intelligent choice of a diversified proxy data set which is sensitive to different climate parameters (e.g. temperature *and* precipitation dependencies) helps to overcome the drawbacks of the time-dependent reconstruction skill of the various proxies.

Acknowledgments. The authors wish to thank C. W. Stockton, M. F. Glueck and C. Appenzeller for providing their NAO index reconstructions. Many thanks to Phil Jones and another (anonymous) reviewer for their helpful comments. We would also like to thank Mary Brown for proofreading the English text. This work was made possible by the Swiss National Science Foundation.

References

Appenzeller, C., T. F. Stocker, and M. Anklin, North Atlantic Oscillation dynamics recorded in Greenland ice cores, *Science, 282,* 446-449, 1998.

Barnston, A. G., and R. E. Livezey, Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Mon. Wea. Rev., 115,* 1083-1126, 1987.

Cook, E. R., R. D. DArrigo, and K. R. Briffa, The North Atlantic Oscillation and its expression in circum-Atlantic tree-ring chronologies from North America and Europe, *The Holocene, 8,* 9-17, 1998.

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

Hurrell, J. W., Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperature, *Geophys. Res. Lett., 23,* 665-668, 1996.

Hurrell, J. W., and H. van Loon, Decadal Variations in Climate associated with the North Atlantic Oscillation, *Clim. Change, 36,* 301-326, 1997.

Jones, P. D., T. Jonsson, and D. Wheeler, Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland, *Int. J. Climatol., 17,* 1433-1450, 1997.

Jones, P. D., K. R. Briffa, T. P. Barnett and S. F. B. Tett, High-resolution palaeoclimatic records for the last millenium: interpretation, integration and comparison with General Circulation Model control-run temperatures, *The Holocene, 8,* 455-471, 1998.

Kushnir, Y., and J. M. Wallace, Low-frequency variability in the Northern Hemisphere winter, *J. Atmos. Sci., 46,* 3122-3142, 1989.

Lamb, P.J., and R.A. Peppler, North Atlantic Oscillation: concept and application, *Bull. Amer. Meteor. Soc., 68,* 1217-1225, 1987.

Luterbacher, J., C. Schmutz, D., Gyalistras, E. Xoplaki and H. Wanner, Reconstruction of monthly NAO and EU indices back to AD 1675, *Geophys. Res. Lett., 26,* 2745-2748, 1999.

Mann, M., R. Bradley and M. Hughes, Global-scale temperature patterns and climate forcing over the past six centuries, *Nature, 392,* 779-787, 1998.

Osborn, T. J., K. R. Briffa, S. F. B. Tett, P. D. Jones, and R. M. Trigo, Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Clim. Dyn., 15,* 685-702, 1999.

Rogers, J. C., The association between the North Atlantic Oscillation and the Southern Oscillation in the Northern Hemisphere, *Mon. Wea. Rev., 112,* 1999-2015, 1984.

Stockton, C. W., and M. F. Glueck, Long-term variability of the North Atlantic Oscillation (NAO), in Preprint of the Am. Met. Soc. 10th Symposium on Global Change Studies, 10-15 Jan 1999, Dallas, Texas, 290-293, 1999.

van Loon, H., and J. Rogers, The seesaw in winter temperatures between Greenland and Northern Europe. Part I: General Description, *Mon. Wea. Rev., 106,* 296-310, 1978.

Wallace, J. M., and D. S. Gutzler, Teleconnections in the geopotential height field during the Northern Hemisphere winter, *Mon. Wea. Rev., 109,* 784-812, 1981.

Walker, G. T., Correlations in seasonal variations of weather IX, *Mem. Ind. Meteorol. Dept., 24,* 275-332, 1924.

C. Schmutz, J. Luterbacher, D. Gyalistras, E. Xoplaki and H. Wanner, Institute of Geography, University of Bern, Hallerstrasse 12, 3012 Bern, Switzerland. (e-mail: schmutz@giub.unibe.ch)

(Received September 1, 1999; revised February 17, 2000; accepted February 23, 2000.)